Dissertation for Degree of Doctor of Medicine,

University of Oxford.

THE AETIOLOGY OF TYPE 1 DIABETES MELLITUS

A prospective study of HLA-linked genetic factors, autoimmunity and viral infections.

ANDREW NICHOLAS GORSUCH

MA, BM, BCh, MRCP(UK)

Senior Registrar in Medicine, University College Hospital,
and
Honorary Clinical Lecturer, University College, London

1988

TABLE OF CONTENTS

Page

(Each chapter has its own detailed list of contents)

List of Tables iii
List of Figures vi
Abbreviations viii

I Brief abstract 1
II Introduction 2
A. Background 2
B. Principal aims of the present work 3
C. About this dissertation 4
III Literature survey
1 Introduction; heterogeneity; definition of Type 1 diabetes 5
2 Aetiology: historical background 8
3 Islet histopathology in diabetes 10
4 Epidemiology of Type 1 diabetes 31
5 Genetics (1): early work on diabetes 52
6 Genetics (2): susceptibility to Type 1 diabetes 62
7 Type 1 diabetes as an autoimmune disease 94
8 Environmental factors in aetiology of Type 1 diabetes . . . 112
9 The pre-diabetic phase of Type 1 diabetes 123
10 Heterogeneity within Type 1 diabetes 127
11 Actions of susceptibility genes 137
IV Indications for present research 143
V Work done 147
VI Methods
1 Fieldwork 151
2 Laboratory methods 160
3 Data handling & statistical analysis 170
VII Subjects studied 193
VIII Results 203
1 Basic genetic and immunological observations. (Contents) . 203
A. Introduction 204
B. Chromosome-6 genetic markers in the BWFS 204
i Phenotype and gene frequencies in BWFS probands . 204
ii Linkage disequilibrium in the BWFS 211
iii Genetic recombination in the BWFS 213
iv Zygotic assortment of HLA-B8: random or not? . . . 215
v Siblings grouped by haplotype concordance with probands 225
C. Basic immunological observations 228
2 HLA and Type 1 diabetes: mode of inheritance. (Contents) . 236
A. Introduction 236
B. Observations and Analyses 237
i Linkage: combined data from diabetic sib pair studies 237
ii Heterogeneity of Type 1 diabetes in unrelated individuals 239
iii Further analyses relevant to mode of inheritance . 263
C. Discussion 275
3 The prospective study: the pre-diabetic period. ICA as a marker. Environmental factors. (Contents) 280
A. Introduction 280

TABLE OF CONTENTS

VIII Results 3 (continued)
B. Observations and Analyses 281
i Prospective follow-up: numbers and duration . . . 281
ii Islet-cell antibodies after diagnosis of diabetes . 282
iii Islet-cell antibodies in non-diabetic subjects . . 284
iv Environmental events 287
v New cases of diabetes developing during follow-up . 289
vi Glucose intolerance short of diabetes 290
vii Islet-cell antibody as a marker of the pre-diabetic state 294
C. Discussion 296
4 What is the risk of Type 1 diabetes in siblings of an affected child? (Contents) 299
A. Introduction 299
B. Methods: further details 299
C. Observations and analyses using local control data . . . 301
D. Discussion 302
5 Aetiological mechanisms: the nature of HLA-linked susceptibility and its interaction with other factors in pathogenesis 306
A. Introduction 306
B. Observations and analyses 307
i Immune response and HLA-linked susceptibility to diabetes 307
ii Islet B-cell function and HLA-linked susceptibility to diabetes 314
C. Discussion 314
IX Summary of findings and conclusions 317
X Indications for future research 327
XI Acknowledgments 329
XII References 330
XIII Appendices 351
1 Diabetic probands in the BWFS 352
2 Tables of genetic observations 358
3 Newly-diagnosed diabetics not included in the BWFS 362
4 Availability of data from the BWFS and related studies . . 363
5 Microcomputer programs 364
6 East Berkshire: demographic and epidemiological data . . . 382
7 Papers by the author forming part of this dissertation . . 385

LIST OF TABLES

TABLE	TITLE	PAGE
III-1	Characteristic features of fully established Type 1 and Type 2 diabetes	6
III-2	Insulitis in type 1 diabetes: occurrence in relation to age at diagnosis	23
III-3	Insulitis in acute Type 1 diabetes	24
III-4	Male:female ratios for new cases of Type 1 diabetes	36
III-5	Sources of data used in Figure 1	39
III-6	Linkage disequilibria in the HLA system	70
VIII-1	Some HLA-A,B,Cw phenotype and gene frequencies in 165 unrelated probands of the BWFS, with control observations	205
VIII-2	Complement factor B (Bf) gene-counts in unrelated probands and in 'control' haplotypes from the BWFS	209
VIII-3	Complement C2 phenotype and gene frequencies in Type 1 diabetic probands of the BWFS, with control observations.	210
VIII-4	Glyoxalase (GLO) allotypes in unrelated Type 1 diabetic BWFS probands and in London control subjects	210
VIII-5	First-order linkage disequilibria in pooled haplotypes from the BWFS	212
VIII-6	The three-locus association B18.Cw5.BfF1: contingency table.	212
VIII-7	Distribution of B8+ and B8- haplotypes among offspring of 110 parents heterozygous for B8, categorized by offspring's sex and presence or absence of Type 1 diabetes (DM1) at entry, and by parent originating haplotype.	219
VIII-8	Log-linear analysis of data in Table VIII-7, using GLIM.	221
VIII-9	Parameter estimates for Model 5 of Table VIII-8.	221
VIII-10	Distribution of B8+ and B8- haplotypes among offspring of 110 parents heterozygous for B8, categorized by offspring's sex and presence or absence of Type 1 diabetes (DM1) at entry, irrespective of parent originating haplotype.	223
VIII-11	HLA-typed siblings of probands in the BWFS by haplotype concordance with probands.	225
VIII-12	Age at entry of siblings in the BWFS, grouped by haplotype concordance with probands: Analysis of variance.	227
VIII-13	Diabetes in relatives of BWFS probands.	229
VIII-14	Certain diseases other than diabetes in relatives of BWFS probands.	230
VIII-15	Islet-cell and other organ-specific auto-antibodies in diabetic subjects at entry to the BWFS and supplementary study.	231
VIII-16	Islet-cell and other organ-specific auto-antibodies in non-diabetic subjects at entry to the BWFS	233
VIII-17	Islet-cell antibody in subjects at entry to the BWFS: cross-classification by technique used (conventional versus complement-fixing immunofluorescence tests)	234
VIII-18.	HLA haplotype concordance in pairs of Type 1 diabetic siblings. Combined data, 1981.	237

LIST OF TABLES

VIII-19 Variables considered in discrete multivariate analysis for heterogeneity within Type 1 diabetes. 240

VIII-20 Thyrogastric antibody status (T), Age at diagnosis of diabetes (A) and Age at entry to study (E) in 308 newly-diagnosed cases and probands of the Barts-Windsor and augmented BDA Multiplex family studies. 241

VIII-21 Season at diagnosis of diabetes (Q), Number of affected siblings (N), Age at diagnosis of diabetes (A) and Age at entry to study (E) in 308 newly-diagnosed cases and probands of the Barts-Windsor and augmented BDA Multiplex family studies. 242

VIII-22 HLA-B8 (B), Age at diagnosis (A), Sex (G), Number of affected siblings (N) and Source of data (S) in 559 unrelated Type 1 diabetic probands from the Barts-Windsor, augmented BDA multiplex and Liverpool studies. 244

VIII-23. HLA-B8 (B), Age at diagnosis (A), Sex (G) and Number of affected siblings (N): Pooled data from three sources (from Table VIII-22, collapsed over S). 246

VIII-24. HLA-B15 (B), Age at diagnosis (A), Sex (G), Number of affected siblings (N) and Source of data (S) in 559 unrelated Type 1 diabetic probands from the Barts-Windsor, augmented BDA multiplex and Liverpool studies. 247

VIII-25 Type 1 diabetic probands categorized by Sex, Number of siblings affected, HLA-B8 status and tranche of Age at diagnosis. Log-linear analysis of frequency data, separately according to source (data of Table VIII-24-A) 249

VIII-26 HLA-B15 (B), Age at diagnosis (A), Sex (G), Number of affected siblings (N) and Source of data (S) in 559 unrelated Type 1 diabetic probands. Data of Table VIII-24, simplified by amalgamation of multiple categories of A and S. 250

VIII-27 HLA-B8,15 (B), Age at diagnosis (A; 5 categories), Sex (G), Number of affected siblings (N) and Source of data (S) in 559 unrelated Type 1 diabetic probands. 251

VIII-28 HLA-B8 (B), Age at diagnosis (A; 4 and then 3 categories), Quarter at diagnosis (Q), Sex (G) and Thyrogastric antibody status (T) in 300 unrelated Type 1 diabetic probands. 252

VIII-29 HLA-B8 (B), Age at diagnosis (A; 2 categories), Quarter at diagnosis (Q), Sex (G) and Number of affected siblings (N) in 308 unrelated Type 1 diabetic probands from the pooled BDA multiplex and Barts-Windsor studies (with newly-diagnosed cases). 254

VIII-30. HLA-B8 (B), Quarter at diagnosis (Q), Sex (G) and Number of affected siblings (N). A, B: Log-linear analyses, separately for 73 probands aged <5 years and for 235 probands aged 5 or more at diagnosis. 255

VIII-31. HLA-B8 (B), Quarter at diagnosis (Q), Sex (G) and Age at diagnosis (A). A, B: Log-linear analyses, separately for 181 simplex and 127 multiplex probands. 258

VIII-32. HLA-B15 (B), Age at diagnosis (A; 3 categories), Quarter at diagnosis (Q), Sex (G) and Thyrogastric antibody status (T) in 300 unrelated Type 1 diabetic probands. 260

VIII-33. HLA-B15 (B), Age at diagnosis (A; 2 categories), Quarter at diagnosis (Q), Sex (G) and Number of affected siblings (N) in 308 unrelated Type 1 diabetic probands. 261

VIII-34. HLA-B8,15 (B), age at diagnosis (A; 3 categories), quarter at diagnosis (Q), sex (G) and thyrogastric antibody status (T) in 300 unrelated Type 1 diabetic probands. 262

VIII-35. HLA-B8,15 (B), age at diagnosis (A; 2 categories), quarter at diagnosis (Q), sex (G) and number of affected siblings (N) in 308 unrelated Type 1 diabetic probands. 263

VIII-36. Hardy-Weinberg equilibrium of HLA-A, -B and -C alleles. A. In 165 unrelated probands of the Barts-Windsor family study. B. HLA-B8: pooled data. 264

LIST OF TABLES

VIII-37. Hardy-Weinberg equilibrium of complement factor B (Bf) allotypes in 53 unrelated probands of the Barts-Windsor Family Study, partially selected for presence of HLA-B18. 265

VIII-38. Hardy-Weinberg equilibrium of complement factor C2 allotypes B and C in 106 unrelated probands of the Barts-Windsor Family Study. 265

VIII-39. Hardy-Weinberg equilibrium of erythrocyte glyoxalase-I (GLO) allotypes 1 and 2 in 155 unrelated probands of the Barts-Windsor Family Study. 265

VIII-40. Genotypic combinations of the HLA-B 'alleles' 8, 15 and X: frequencies in 310 Type-1 diabetic probands. 267

VIII-41. Linkage disequilibria within the HLA region: comparison between 'diabetic' and 'control' haplotypes in the Barts-Windsor Family Study. 271

VIII-42. Haplotype sharing and HLA-B specificities in duplex sibships: segregation to diabetic children of parental haplotypes categorized by HLA-B allele. 273

VIII-43. Haplotype sharing and HLA-B specificities in duplex sibships. 273

VIII-44 Persistent islet-cell antibody versus HLA-B8 in Type 1 diabetes in the Barts-Windsor Family Study. 283

VIII-45. Islet-cell antibodies in non-diabetic participants in the Barts-Windsor Family Study: prospective follow-up to 31st October 1980, or to development of diabetes if sooner. 285

VIII-46. Glycosylated haemoglobin in non-diabetic participants in the Barts-Windsor family study. Distribution of 'high' results with respect to HLA haplotype concordance with proband. 292

VIII-47. Frequency of probable viral infections shortly before 'high' and 'normal' HbA1 results in non-diabetic BWFS participants. 293

VIII-48. Islet-cell antibodies and development of diabetes in 584 initially non-diabetic relatives of type 1 diabetic probands followed up for up to 2.5 years. 295

VIII-49. Prevalence of Type 1 diabetes on 30:10:1980 in HLA-typed siblings of probands in the BWFS, grouped by haplotype concordance with proband, and relative risk with respect to control population 302

VIII-50. Absolute risk of Type 1 diabetes in siblings grouped by haplotype concordance with proband, in East Berkshire, 1978-80. 303

VIII-51. Occurrence of antinuclear antibodies (ANA) in HLA-typed siblings of probands in the BWFS, grouped by haplotype concordance with proband. 313

VIII-52. Antinuclear antibodies (ANA) in relation to thyroid and islet-cell antibodies in the BWFS 314

LIST OF FIGURES

Figure III-1.	Relative cumulative frequency of Type 1 diabetes, related to age.	40
Figure III-2.	Age-specific cumulative frequency of diabetes, both sexes	42
Figure III-3.	Potential prevalence of Type 1 diabetes	44
Figure III-4.	The HLA system	66
Figure III-5.	Determination of HLA haplotypes	67
Figure III-6.	HLA phenotype associations with Type 1 diabetes	76
Figure III-7.	HLA genotypes in a diabetic multiplex family	78
Figure III-8.	Expected haplotype concordance for diabetic sib pairs, assuming no linkage.	79
Figure VI-1.	Map of East Berkshire and surrounding districts.	153
Figure VI-2.	Complement-fixing islet-cell antibody	165
Figure VI-3.	Gastric parietal-cell and adrenal steroidal-cell antibodies	166
Figure VI-4.	Anti-nuclear antibodies.	167
Figure VI-5.	Genobase: overall structure and facilities	176
Figure VI-6.	Genobase: schematic representation of family data array.	177
Figure VI-7.	Genobase: data input document format.	183
Figure VI-8.	Genobase: a command file for the data analysis system.	186
Figure VI-9.	Genobase: the volume of data.	188
Figure VII-1.	Age of 166 probands at entry to the BWFS	195
Figure VII-2.	Age of 289 siblings at entry to the BWFS	195
Figure VII-3.	Age of 154 fathers (a) and 159 mothers (b) at entry to the BWFS	196
Figure VII-4.	Age at diagnosis of diabetes in 166 BWFS probands	198
Figure VII-5.	Duration of diabetic symptoms reported by 151 BWFS probands.	198
Figure VII-6.	Duration of diabetes in 166 probands at entry to BWFS.	198
Figure VII-7.	Age at diagnosis of diabetes in 125 probands in the UK (BDA) Multiplex Study	201
Figure VII-8.	Age at diagnosis of diabetes in 288 probands in the Liverpool study.	201
Figure VIII-1.	Intra-HLA recombinants in the BWFS	214

FIGURES

Figure VIII-2.	HLA haplotype concordance in pairs of Type 1 diabetic siblings from 376 multiplex sibships	238
Figure VIII-3.	Season at diagnosis of diabetes versus HLA-B8 and age at diagnosis.	256
Figure VIII-4.	Season at diagnosis of diabetes versus age at diagnosis and number of affected siblings	259
Figure VIII-5.	Serological follow-up of 155 probands, 271 siblings, 297 parents by 31st October 1980.	281
Figure VIII-6.	Islet-cell antibody: frequency of positive results versus duration of diabetes	283
Figure VIII-7.	Frequency of urine-testing by 204 non-diabetic siblings in the Barts-Windsor Family Study	288
Figure VIII-8.	Rate of occurrence of illnesses attributable to viruses in 110 male and 94 female non-diabetic siblings in the Barts-Windsor Family Study.	288
Figure VIII-9.	Polio titres in previously-vaccinated children in the BWFS	308
Figure VIII-10.	Proportions (%) of "responders" and "non-responders" to Poliovirus Types 1, 2 and 3 among vaccinated BWFS siblings grouped by HLA haplotype concordance with probands	309
Figure VIII-11.	Mean of log titres to Poliovirus Types 1, 2 and 3, in vaccinated, non-diabetic BWFS siblings grouped by HLA-haplotype concordance with probands	310
Figure VIII-12.	Poliovirus Type 1 titres in vaccinated non-diabetic BWFS siblings, grouped by HLA haplotype concordance with probands.	311
Figure VIII-13.	Rubella titres in BWFS siblings grouped by HLA-haplotype concordance with probands	312

LIST OF ABBREVIATIONS

Δ	(a measure of linkage disequilibrium)
δ	Population attributable risk, or aetiologic fraction
Ab	Antibody
AI	Autoimmune, autoimmunity
ANA	Antinuclear antibody
BB rat	(originally) Bio-Breeding rat strain
BDA	British Diabetic Association
Bf	Factor B; properdin
C2	Second component of complement
C3	Third component of complement
C4	Fourth component of complement
CF-ICA	Complement-fixing ICA
CMI	Cell-mediated immunity
Conv-ICA	Cytoplasmic ICA detected by conventional IFT
CPAF	Chlorpropamide-alcohol flushing
df	Degrees of freedom
E-RFC	Erythrocyte rosette-forming cell(s)
EDTA	Ethylene diamine tetra-acetic acid
EMCV	Encephalomyocarditis virus
GLO	Glyoxalase-I
HbA_1	Glyc(osyl)ated haemoglobin
H0, H1, H2	(siblings) Sharing 0, 1, 2 haplotypes with diabetic proband
HLA	The human major histocompatibility complex
ICA	Islet-cell antibody/antibodies
ICSA	Islet-cell surface antibody
IDD	Insulin-dependent diabetes
IFT	Immunofluorescence test
JOD	Juvenile-onset diabetes
LMIT	Leucocyte migration inhibition test
MHC	Major histocompatibility complex
MIF	Migration inhibitory factor
MLC	Mixed lymphocyte culture
MOD	Maturity-onset diabetes
MODY	Maturity-onset diabetes in youth
NIDD	Non-insulin-dependent diabetes
NS	Not (statistically) significant
OSAb	Organ-specific autoantibody/antibodies
OSAI	Organ-specific autoimmunity
OSAID	Organ-specific autoimmune disease
PAR	Population attributable risk, or aetiologic fraction
PNU	N-3-pyridylmethyl N-p-nitrophenyl (Vacor, a rodenticide)
PP	Pancreatic polypeptide
R	Hypothetical diabetes resistance gene associated with HLA-DR2
rΔ	Relative Δ
RR	(Approximate) relative risk
S1	Hypothetical diabetes susceptibility gene associated with HLA-DR3
S2	Hypothetical diabetes susceptibility gene associated with HLA-DR4
SD	Standard deviation
SEM	Standard error of mean
TSH	Thyroid stimulating hormone
WHO	World Health Authority

PART I: BRIEF ABSTRACT

TITLE: Aetiology of Type 1 diabetes mellitus: a prospective investigation of HLA-linked genetic factors, autoimmunity and viral infections.

HLA genotypes were determined and prospective investigations performed on members of 166 families, each including a Type 1 diabetic proband. A computer-based system (Genobase) was created for data storage and analysis. Newly-diagnosed cases ascertained in East Berkshire (incidence rate: 17.8 x 10^{-5}/year) and additional data from two other studies contribute to some analyses.

Major findings include non-random assortment of HLA-B8 within families and a tendency for HLA-identical siblings of diabetic probands to be younger than expected. Discrete multivariate analysis and other methods show no major genetic heterogeneity within Type 1 diabetes. However, there is minor heterogeneity affecting the interaction of age and seasonal variation in incidence of the diabetes. Furthermore, the HLA genotype B8,15 is here shown to be significantly more frequent in Type 1 diabetic probands than expected assuming Hardy-Weinberg equilibrium. With support from other internal and published data, it is concluded that susceptibility to Type 1 diabetes is genetically multifactorial.

Increased prevalence of organ-specific autoimmunity is confirmed in the families, and differences in distribution of islet-cell antibodies (ICA) and thyroid/gastric antibodies demonstrated. Five subjects, ICA-positive from the first test, developed diabetes after up to 30 months, demonstrating prolonged latency. Quantitative risk of diabetes is estimated (separately) in terms of ICA results and of HLA genotypes. Non-diabetic siblings of affected probands are investigated for evidence that an HLA-linked immune response gene might contribute to susceptibility. Response patterns to previous polio vaccination give some support, but the excess of thyroid/gastric antibodies in the siblings is not due to such a gene. The dissertation concludes with discussion of aetiological hypotheses and indications for further work.

PART II — INTRODUCTION

CONTENTS

A. Background
B. Principal aims of the present work
C. About this dissertation

A. BACKGROUND

The aetiology of Type 1 (insulin-dependent) diabetes involves both genetic and environmental factors. It is well established that the genetic susceptibility is HLA-linked, so that a sibling who has inherited from his parents the same two HLA haplotypes as a diabetic child (irrespective of the HLA specificities included) is likely to be susceptible, whereas a sibling who has neither of these is not. The mode of inheritance, mechanism of action and 'penetrance' of these susceptibility genes are not known. By analogy with other genes linked to the major histocompatibility system in man and animals, it is likely that they modify some immunological function.

Histological and clinical observations suggest that the pathological process leading to islet B-cell damage may begin several months, or perhaps longer, before onset of clinical diabetes. This means that investigation of subjects who are already diabetic in an attempt to identify the mechanism of islet B-cell damage is necessarily retrospective. There is indirect evidence from such studies suggesting that viruses may constitute an environmental factor in aetiology. However, this evidence is circumstantial. A prospective study of initially healthy children starting before initiation of pathogenesis or early in the pre-diabetic period should help to elucidate the mechanism of pathogenesis and the role of environmental factors.

Islet-cell antibodies (ICA) are rare in the general population but present in most newly-diagnosed Type 1 diabetics. It seems likely that ICA first appears at some time during, and could thus act as a marker of, the pre-diabetic period. This 'marker' function has been shown in adults with organ-specific autoimmune diseases, but until the work described here has not been demonstrated in children; if confirmed in a prospective study it would help the investigation of events before onset of diabetes.

B. PRINCIPAL AIMS OF THE PRESENT WORK

i. To initiate a prospective investigation of environmental events and autoimmune status before the onset of any future Type 1 diabetes developing in siblings or parents of affected children;

ii. To investigate the use of islet cell antibodies in predicting future Type 1 diabetes;

iii. To investigate the use of HLA genotypes in assessing risk of future diabetes in siblings of affected children;

iv. To examine organ-specific autoimmunity and viral responses in HLA-genotyped families of Type 1 diabetic probands, comparing 'susceptible' with 'non-susceptible' siblings, in order to explore the hypothesis that HLA-linked genes determining susceptibility act by modifying immune response;

v. To investigate various genetic aspects of HLA generally and of Type 1 diabetes in the large cohort of HLA-genotyped families built up:
   a) Associations with Type 1 diabetes of HLA Class I-III allelotypes
   b) Linkage disequilibria within the HLA complex
   c) Whether assortment of parental HLA haplotypes is random, following Mendelian rules
   d) Whether there is genetic heterogeneity within Type 1 diabetes
   e) Mode of inheritance of HLA-linked genetic susceptibility to Type 1 diabetes.

C. ABOUT THIS DISSERTATION

The background on which this work is founded covers a wide range of biomedical knowledge, including aspects of biochemistry, histopathology, epidemiology, genetics, immunology, virology and statistics, based on an extensive literature which is often confusing and contradictory. Inevitably, the Literature Survey is long, but it is carefully structured so that topics referred to later can readily be located.

Another feature of this project is the large amount and complex nature of data generated in the first 2.5 years of prospective study, necessitating the development of a new computerized system for data storage and analysis. The analyses and results are presented in five chapters, each having its own discussion. The first two entail rigorous analyses of extensive data, and again are inescapably rather long.

Because of the problems of scale, and with commitments in clinical medicine, it has taken the author several years to complete and present the work in this form. Some findings have meanwhile been published, mainly in 1980-1982; the papers concerned (bound as Appendix 7) form an integral part of the dissertation, which should be considered in the context of that period. Relevant subsequent developments are, however, mentioned in Literature Survey and Discussion sections as appropriate.

The text was written and edited by the author using word-processing and outliner software (LEX, PC-Write and PC-Outline), printed under PC-write on a Panasonic KX-P1080 9-pin dot-matrix printer, and photocopied. Most tables were prepared and printed in the same way, using PC-Write, while many figures and a few tables were prepared and printed using the integrated microcomputer software package Ability. The title page was printed in Main and certain diagrams (e.g. Figures III-5 to III-8) were created using the same equipment controlled by the program LQ. Further details are given in VI-3.

PART III - LITERATURE SURVEY

CHAPTER 1

INTRODUCTION

HETEROGENEITY; DEFINITION OF TYPE 1 DIABETES

The syndrome of diabetes mellitus is characterized by the copious volume and sweet taste of the urine produced. Both were recognized in ancient times. Aretaeus the Cappadocian first used the term diabetes ('διαβετες', a passer through, a syphon), applying it in Alexandria in the second century AD to a dramatic but 'fortunately rather rare' condition in which there was severe weight loss ('a liquefaction of the flesh and limbs into urine') followed before very long by death (Papaspyros, 1964). His description is recognizable as the insulin dependent form of diabetes which often begins in childhood or adolescence, and usually leads to fatal ketoacidosis if not treated.

In 1877 Lancereaux emphasized the difference between this 'Diabète maigre' and the 'Diabète gras' typically developing in overweight adults (Papaspyros,1964). In the latter form, which is the more common nowadays in developed countries, excess production of ketones is not a feature unless there is severe intercurrent illness or starvation, and insulin treatment is usually unnecessary, except perhaps after many years.

The classification of diabetes mellitus and the terminology used are still somewhat controversial. In the temperate zones most patients have primary ('idiopathic') diabetes, whereas cases secondary to other conditions and the mild diabetes of some rare genetic syndromes (Rimoin 1976) are much less common. Within primary diabetes individual cases may not fall neatly into either category, but it is now clear that the two broad clinical subdivisions described above reflect true aetiopathogenic heterogeneity, rather than representing severe and mild expressions of

the same underlying disease process. The histopathological, genetic and immunological evidence for this conclusion is summarized in Table III-1, and is discussed in more detail in the appropriate sections below (Chapters 3, 5, 6 and 7).

TABLE III-1. CHARACTERISTIC FEATURES OF FULLY ESTABLISHED TYPE 1 AND TYPE 2 DIABETES

	TYPE 1	TYPE 2
CLINICAL FEATURES		
Age of onset	Any; usually < 30y	Usually > 20y
Sex	M > F (children)	M < F
Body weight	Usually lean	Often obese
Presenting symptoms	Often acute	Usually insidious
Ketonuria (fed)	Heavy	Absent
Insulin dependence (not under metabolic stress)	Yes	No, unless after many years
PANCREATIC HISTOLOGY [a]		
Size of pancreas	Small	Normal
Specific B-Cell depletion	Severe or total	No
Hyalinosis	No	Longstanding cases
Insulitis	Some acute cases	Never reported
GENETIC FEATURES [b]		
Increased family prevalence of:		
Type 1 diabetes	Yes	No
Type 2 diabetes	No	Yes
OSAID	Yes	No (?)
HLA-linked suscept'y	Yes	No
CPAF-linked suscept'y	No	Some families
IMMUNOLOGICAL PHENOMENA [c]		
Islet cell antibody	Most acute cases	No
Other OSAb	Common	Probably not increased
CMI (pancreas)	Some acute cases	Rare

ABBREVIATIONS: CMI = cell-mediated immunity. CPAF = chlorpropamide-alcohol flushing. OSAb = organ-specific auto-antibodies. OSAID = organ-specific autoimmune disease.
FOR DISCUSSION OF EVIDENCE: (a) see Chapter 3 of literature survey; (b) see Chapter 5; (c) see Chapter 6.

The terms 'juvenile onset' (JOD) and 'maturity onset diabetes' (MOD) are not entirely appropriate. Some cases developing in adult life clearly belong to the 'JOD' group (Cudworth 1978), and conversely some childhood cases do not and are termed 'MODY' (maturity onset diabetes of the young; Tattersall & Fajans, 1975). The best clinical criteria for defining that type of diabetes which most typically occurs in children, irrespective of the actual age of onset, are undoubtedly dependence on insulin within a few years after diagnosis, and development of ketosis when treatment is inadequate. Many authors therefore favour the terms 'insulin-dependent' (IDD) and 'non-insulin-dependent' or 'insulin-independent' diabetes (NIDD), which are generally satisfactory for clinical purposes (Pyke, 1980).

Nevertheless, these terms do have disadvantages. Even classical MOD may require insulin at times of severe metabolic stress or after many years of disease, whereas in some parts of the world where classical JOD is very rare the insulin-dependent but non-ketosing condition of tropical malnutrition diabetes is common in children (McMillan & Geevarghese, 1979). Furthermore some cases of mild diabetes not requiring insulin may actually represent a 'forme fruste' of the juvenile onset type; this is suggested by the observation that adults in whom oral hypoglycaemic therapy fails after a number of years of satisfactory control show the HLA-genetic and immunological features of that category (Cudworth & Woodrow, 1974). On the other hand there are many patients who are treated with insulin even though they are not truly dependent on it.

Since no single clinical or other feature can reliably identify which of the two principal aetiopathogenic types of disease process is operating in a living diabetic patient, the terms Type 1 and Type 2 diabetes are preferable in a pathological context (Hugh-Jones, 1955; Cudworth, 1978; WHO Expert Committee on Diabetes Mellitus, 1980). This Dissertation is concerned with aetiological aspects of Type 1 diabetes only, except when considering the evidence for heterogeneity within primary diabetes.

PART III — CHAPTER 2

AETIOLOGY: HISTORICAL BACKGROUND

During and after the Renaissance in Europe there were theories attributing diabetes to numerous causes. Among the less bizarre were some obviously suggested by the polyuria, such as the views that abuse of diuretic herbs or beverages, suppression of sweating or indulgence in unnecessarily frequent micturition might be responsible. Other claims may have reflected actual or apparent precipitation of the disease by acute infections, with several references to smallpox, while the current interest in dietary factors in development of Type 2 diabetes was anticipated by a tendency to blame sweet fruit and vegetable foods (reviewed by Müller, 1822).

The sweetness of the urine in what we now call diabetes mellitus was known to ancient Hindu physicians, as well as to Avicenna around 1000 AD. In Europe it seems to have become sufficiently familiar by 1650 for a mention in a Molière comedy. Even so, it is Thomas Willis who is generally credited with this discovery in 1674, and it was Willis who held with some prescience that the sweetness of the urine was secondary to a similar abnormality of the blood. Later developments in analytical chemistry allowed Matthew Dobson to show sugar in diabetic urine and serum in 1766 (Papaspyros, 1964).

Another great advance in understanding came in 1889, when von Mering and Minkowski demonstrated that pancreatectomy led to diabetes in the dog. (The same experiment had been performed and the resulting thirst and polyuria noted two centuries before, by Brünner in Amsterdam in 1683 (cited by Papaspyros, 1964).) It was already known that ligating the pancreatic duct did not have this effect, although it did cause atrophy of the gland, and Arnozant & Vaillard (1884) had shown that after duct ligation the exocrine tissue atrophied whereas there was no apparent

damage to the islets described by Langerhans (1869). Ssobolew (1902) later confirmed that this experimental procedure did indeed lead to immediate atrophy and sclerosis of the exocrine pancreas, while islet atrophy and glycosuria ensued only after one to three months. Thus it appeared that loss of islets rather than of exocrine tissue was the cause of diabetes following pancreatectomy in experimental animals.

Subsequent pathological observations in naturally occurring diabetes led to the discovery that here, too, islet abnormalities were common, one characteristic finding being absence of the B cells; and it became generally accepted that these abnormalities were fundamentally important in the human disease.

The term 'insulin' was coined independently by de Meyer in 1909, and by Sharpey-Schafer in 1916 (Sharpey-Schafer, 1926, p. 343). Banting and Best (1922a,b) later extracted the hormone from bovine pancreas before demonstrating its efficacy in diabetes, whether naturally occurring or induced by pancreatectomy. It is now well established that insulin does indeed originate in the islet B cells, and that the immediate cause of Type 1 diabetes is a marked depletion or loss of these cells (see below), resulting in grossly subnormal or undetectable circulating levels of insulin and C-peptide (Bornstein & Lawrence, 1951; Ludvigsson & Heding, 1978).

PART III — CHAPTER 3

ISLET HISTOPATHOLOGY IN DIABETES

CONTENTS

Section A. Changes in arrangement and numbers of islet cells · · 10
i. Early work · 10
ii. Qualitative: changes in islet architecture · · · 12
iii. Quantitative: changes in numbers of B cells · · 14
B. Insulitis · 18
i. Occurrence · · · · · · · · · · · · · · · · · · · 18
ii. Characteristics · · · · · · · · · · · · · · · · 20
C. Implications concerning the disease process · · · · · 21
i. Evidence for heterogeneity of diabetes · · · · · 21
ii. Specificity of lesion · · · · · · · · · · · · · 25
iii. Timing · 26
iv. Mechanism · · · · · · · · · · · · · · · · · · · 27

3-A. CHANGES IN ARRANGEMENT AND NUMBERS OF ISLET CELLS

3-A-i. Early work

During the first quarter of the present century a number of pathologists studied islet histopathology using conventional staining techniques on post-mortem material. Their work, reviewed by Kraus (1929), did not in general distinguish clearly between the patterns of disease characteristic of Type 1 and Type 2 diabetes, but it was recognized that appearances varied from case to case, and the observation of 'atrophic' islets was made at this time.

One of the most impressive of these early reports is that of Weichselbaum (1910) who described post-mortem findings in 183 cases, including 20 aged less than 20 years of whom 10 were examined within six months of onset of symptoms. A further 24 cases were in their third decade, but in only three or four was the duration of disease less than six months. The clinical details given indicate that nearly all of the 44 aged under 30 had Type 1 diabetes.

Weichselbaum noted a number of changes characteristic to these younger patients, including a diminution in the number of islets seen, 'hydropic degeneration' of the islets, and 'atrophy apparently following hydropic degeneration'. The latter was seen in 26 out of 83 aged up to 40 y but in only 5 of 100 older patients, and hydropic degeneration was reported in 67 and 16 of these two respective groups. In contrast, hyaline degeneration was present mostly in those over 50 at the time of death (Weichselbaum claimed to have described hyaline degeneration in 1901, before its independent discovery and naming by Opie). Thus, although the histological appearances described by Weichselbaum would now be interpreted somewhat differently (see below), he was one of the first to stress that they showed heterogeneity with respect to age. He also demonstrated that the weight of the pancreas was reduced most markedly in some of his younger cases.

In 1927 Warren stated that in cases of diabetes with onset in childhood, 'There is little pathologic change in either islets or acinous tissue, though commonly reported in the old, and that present does not appear sufficient to account for the disturbance in function'. This was based on ten juvenile-onset cases, most of whom had died within two years of diagnosis without benefit of insulin treatment. However, the work is subject to a number of criticisms. In particular, in only one case report is there any mention of B cells, which had been described 20 years previously by Lane (1907); and Warren mentioned neither the histological techniques used nor the sizes or sites of the necropsy specimens examined. Nevertheless he did stress that there were differences in histology between juvenile and adult-onset diabetes.

More recent qualitative work using new techniques has helped to clarify the picture, and quantitative analyses have confirmed that the total amount of islet tissue is almost invariably reduced in classical, juvenile onset (Type 1) diabetes, whereas in most maturity onset (Type 2) diabetics there is no more than a slight reduction when compared with normal subjects of similar age. However there are obstacles hindering interpretation of much of the evidence, particularly the quantitative work, as described below.

3-A-ii. Qualitative changes in islet architecture

In describing the general cyto-architecture of the islets, Gepts and LeCompte (1981) emphasize in their recent review the contrast between the two principal types of diabetes. In Type 2 there is no characteristic appearance, except in some longstanding cases with abundant hyalinosis (amyloidosis). In juvenile Type 1 diabetes, however, the islets 'always show distinctive and characteristic changes, even in cases of recent onset'. These abnormal islets fall into three categories. In all cases the majority of islets appear small and inconspicuous by conventional staining techniques. They are composed of narrow cords of small cells and were formerly considered atrophic, but immunohistochemical techniques show that the cells principally secrete glucagon (A cells) or somatostatin (D cells) (Orci et al 1976; Gepts & de Mey 1978). They have therefore been termed 'pseudo-atrophic' islets (Gepts & de Mey, 1978).

The second type of islet seen (rarely) in acute Type 1 diabetes is believed by Gepts to be that interpreted by Weichselbaum (1910) as showing 'hydropic degeneration', and consists mainly of B cells with features suggesting considerable functional hyperactivity: large, rounded, vacuolated nuclei, increased cytoplasmic RNA (Weichselbaum's 'Körnchen') and loss of secretory granules (Gepts, 1965; Gepts & Lecompte, 1981).

In long-standing Type 1 diabetes many islets show a third cytoarchitectural pattern, with cords of cylindrical cells often

apparently arising from duct epithelium (Gepts, 1965). These cells predominantly contain pancreatic polypeptide (PP) (Orci et al, 1976). It is now recognized that even in normal subjects such islets are common in certain lobules which, although probably derived from the ventral primordial pancreas, come to lie in the posterior part of the head of the organ, whereas insulin-rich islets form the majority elsewhere (Orci et al 1976; Malaisse-Lagae et al 1979; Baetens et al, 1979). It is not yet clear whether this heterogeneous distribution and the inconspicuous appearance of the majority of islets in Type 1 diabetes are alone responsible for the prominence of PP-rich islets described in some long-standing cases, or whether there is in fact an increase in absolute numbers of PP-cells.

Immunohistochemistry also shows that, in Type 1 diabetes, isolated endocrine cells may be present in the exocrine tissue as well as in the islets. This applies to any remaining B cells as well as to the more plentiful A and D cells (Gepts & De Mey, 1981) and further complicates attempts to estimate total B cell mass.

Throughout the literature there are references to islets apparently arising by regeneration from ductules, in that cords of islet cells often appear to show continuity with the epithelium. For example, Weichselbaum (1910) mentions such an appearance in 58 cases. It is likely that in many instances these were in fact PP islets, but Gepts & de Mey (1978) and Gepts & LeCompte (1981) believe that those islets consisting largely of hyperactive B cells which they have observed in some acute diabetics may also arise in this way. Studies of nucleotide incorporation into DNA have shown that regeneration of B cells in adult animals, normally at a low level, is stimulated by elevated blood glucose concentrations (Hellerström, 1976) and it is not unlikely that the same should apply in human Type 1 diabetes. If so, however, it is odd that mitotic figures are virtually never seen in such cases: Gepts (personal communication, 1982) recalls seeing only two B-cell mitoses in a single diabetic patient, who also had severe liver failure.

3-A-iii. Quantitative changes in islets and in numbers of B cells

According to the classical view, by the time of diagnosis of juvenile Type 1 diabetes islet B cells are either already absent or they disappear completely soon afterwards (Selye, 1947, p. 510). This view was based on post-mortem studies using conventional staining techniques on sections from unspecified areas of pancreas, but reliable counting of small numbers of endocrine cells is possible only with modern, hormone-specific histological techniques, which even the more sophisticated granule stains may fail to demonstrate degranulated B cells (Gepts & LeCompte, 1981).

Even when individual B cells can be identified without ambiguity, in only a few reports is there any serious attempt at quantitation. Ideally, calculation of total B cell mass should take into account not only the heterogeneous distribution of islet cells within the pancreas (see above) and the marked changes with age reported recently in normal subjects (Orci et al, 1979; Malaisse-Lagae et al, 1979; Rahier et al, 1980), but also the size of the whole organ. The pancreas is often much smaller than normal in long-standing Type 1 diabetes (Warren, 1927; Maclean & Ogilvie, 1955; Gepts 1965; Doniach & Morgan, 1973; Rahier et al, 1983) although not in newly diagnosed cases (Gepts, 1965; Doniach & Morgan, 1973, Gepts & LeCompte, 1981). Furthermore, failure to distinguish clearly between Type 1 and Type 2 diabetes is now seen as another important drawback to many of the earlier studies.

Maclean and Ogilvie (1955) used Gomori's chrome haemotoxylin phloxin method, a granule stain which may miss degranulated B cells. Unfortunately, few clinical details are given; of 30 pancreata suitable for examination only five were from cases with onset before age 20 and therefore likely to represent Type 1 diabetes. Longitudinal sections were taken from the head, body and tail of each pancreas and attempts were made to allow for variations in distribution of islets within the organ, but it is not clear whether the PP-rich, insulin-poor islets of the posterior part of the head of the pancreas would have been included

(see above). Thus the quantitative conclusions reached may not have been as accurate as appeared at the time.

Nevertheless Maclean & Ogilvie (1955) did show fairly clearly that the total B cell mass was inversely related to age at onset of the diabetes in their 30 subjects. The only child with an estimated B cell mass > 0.1 g had been diagnosed only six weeks before death, compared with 5, 14, 19 and 21 years in the other four. This is interesting when considering the question of timing of the islet lesion (see 'Implications' below). In the 25 maturity-onset diabetics (age at diagnosis >35 y) the mean B cell mass was higher at 0.25g (range 0.096 - 0.43g), with only one result less than 0.1g. This last patient had clinical features suggesting Type 1 rather than Type 2 diabetes. In normal controls the mean B cell mass came to 0.64g (range 0.36 - 1.07g).

In 1959 the same Edinburgh team reported on 41 young diabetics dying before age 40 years (Maclean & Ogilvie, 1959). Most or all probably had Type 1 diabetes, and eighteen cases were 'acute' (dying within eight weeks of onset). In all but six cases, twelve transverse sections of each pancreas were examined to compensate for heterogeneous distribution of islet tissue.

This study confirmed that in longstanding Type 1 diabetes the total weight of the pancreas, the mean size of individual islets, the proportion of islet tissue in the organ and the estimated total islet tissue mass were all significantly lower than in controls. In contrast, these changes were much less marked in the acute cases, often falling within the normal range. Indeed, the mean size of the islets was significantly greater, and unusually large islets more common, than in controls of similar age. The authors commented that large islets had been noted previously in acute childhood diabetes. Estimation of B cell mass was not attempted.

Somewhat later, again using granule stains although mostly on small blocks of tissue from routine necropsies, Gepts (1965) was unable to demonstrate any B cells at all in the islets of eight out of 12 cases

classed as 'chronic juvenile diabetes', but in three of the remainder there were substantial numbers, albeit well below the lowest seen in normal controls. One case had normal numbers of B cells, but it might be argued that this patient may not have had Type 1 diabetes in that the age of onset was 29 years and it is not clear whether there was insulin dependence. The duration of disease in these 12 patients ranged from six to 37 years, and the age at diagnosis from four to 29 years.

An important question here is the timing of the lesion in relation to onset of diabetes. The less marked changes in the acute juvenile-onset cases studied by Maclean & Ogilvie (1955, 1959) than in their chronic cases have already been mentioned. Gepts (1965) examined pancreatic specimens from 21 subjects aged under 31 years (mostly under 20 years) who had died within six months of diagnosis of primary diabetes, usually in diabetic coma. In only one specimen were no B cells seen, but with a single exception the numbers present were markedly reduced, generally to less than 10% of the normal mean. Those remaining often appeared hyperactive, as described above.

On the study of Doniach and Morgan (1973), the 13 'acute' cases had died in coma between 1907 and 1930 without receiving insulin treatment, and the recorded duration of diabetes was less than 'a few' weeks in all but four. Once again preserved routine necropsy specimens were examined. The authors state that 'such specimens are usually taken from the body of the pancreas' although in no case was the actual site recorded. Cell counts were performed within a single area of $1cm^2$ in each case. Without using immunohistochemistry, the authors considered that eleven of the 13 cases showed a reduction in number of granulated B cells to an average of about one-seventh that seen in carefully matched controls, none at all being detected in three. The remaining two cases appeared to have normal numbers of B cells.

Among their 13 long-standing juvenile-onset cases (duration of disease 6-38 years, mean 9 years) Doniach & Morgan reported that no B cells were seen in six. This was not significantly different from the proportion of acute cases lacking B cells. In only two of the chronic

cases did the number of B cells appear to be within the range seen in matched normal controls. Perhaps the major problem with this study is that the sample examined in each pancreas may not have been representative of the whole.

Gepts and De Mey (1978) applied immunohistochemical methods to the question of B cell survival in juvenile diabetes. Recognizing the variability of their pathological material (much of it the same as that used by Gepts, 1965) and the problem of irregular distribution within the pancreas, they did not attempt a detailed morphometric analysis but used instead a simple scoring system. In all cases the overall number of B cells was obviously reduced, and most islets contained none. However, among 16 recently-diagnosed cases (within six months before death) only two had no detectable B cells at all, and in four they were observed in many of the fields examined. In the remainder of this group B cells were present only in more or less restricted areas of the blocks of tissue. With more prolonged disease surviving B cells were found in smaller proportions of patients: between six months and ten years 6/14 were positive, and after ten years, 5/28. Rahier et al (1983) recently reported presence of rare B cells, isolated among exocrine cells, in three of four longstanding Type 1 diabetic pancreata examined. That persisting B cells may be functionally active is shown not only by their degranulation, open nuclei and increased cytoplasmic RNA, but also by the detection of circulating insulin and C peptide in similar patients during life (Bornstein & Lawrence, 1951; Ludvigsson & Heding, 1978).

In contrast to B cells, A, D and PP cells appeared to Gepts and De Mey to be present in normal or even perhaps increased numbers, supporting the opinions of Orci et al (1976). PP cells in particular became prominent in long-standing juvenile diabetes. In a meticulous, new, quantitative study combining wide sampling of each pancreas with immunohistochemical techniques, the total mass of A, D and PP cells in four Type 1 diabetic patients did not differ from those in eight control subjects of similar ages (Rahier et al, 1983).

To summarize this section, there is wide agreement that in juvenile

Type 1, as opposed to Type 2 diabetes, the number of B cells is already much reduced at the time of diagnosis, and becomes progressively more so thereafter. However, small numbers of functioning B cells persist for many years in a substantial minority. In contrast the other pancreatic endocrine cells are affected relatively little if at all, or may even increase in number. In addition there are characteristic, although unevenly distributed, abnormalities of islet cytoarchitecture. Whether adult onset Type 1 diabetes shows histological changes identical to those in children has yet to be confirmed.

3-B. INSULITIS

3-B-i. Occurrence

Among the early histological reports reviewed by Kraus (1929) were several describing lymphocytic infiltration of islets in the diabetic pancreas. The first is said to be that of Schmidt (1902) who noted among 23 diabetic necropsies:

> *'a nearly isolated acute interstitial inflammation of the islets of Langerhans in a 10 year old child showing 6.8% sugar'. (Author's translation)**

Weichselbaum (1910) made similar observations in seven of his 183 case reports, a typical comment being:

> *'In many islets showing early atrophic changes there is a small-cell infiltration of the capsule, and often of the connective tissue within the islet as well'. (Author's translation)***

However, he did not appear to find these infiltrates particularly remarkable, considering them to represent an early stage in the fibrosis often accompanying 'atrophic' changes in the islets. Furthermore, such infiltrates were rare in the experience of most pathologists, and their importance was overlooked for many years.

Nevertheless, when Warren & Root (1925) noted this appearance in three of 26 cases, they were struck by the fact that all three were children who had died soon after onset of the disease. Two years and ten childhood patients later, Warren (1927) emphasized that 'this lymphocytic infiltration is found in young persons, if at all, and fairly early in

---

*'eine fast isolierte, akute, interstitielle Entzündung der Langerhans'schen Inseln bei einem 10 jahrigen Kind mit 6.8% Zucker...'

**'In vielen Inseln mit beginnender Atrophie ist die Kapsel oder auch das intrainsulare Bindegewebe kleinzellig infiltriert'

the course of the disease'. In this study most of the children had been diabetic for more than a year, but soon afterwards Stansfield & Warren (1928) described a similar picture in two children who had died at presentation with fulminating diabetes.

The term 'insulitis' was first used by von Meyenburg (1940) in his case report on a girl aged 18 years, who, refusing insulin, had died in ketoacidosis five months after diagnosis of diabetes and nine months after onset of symptoms. In the pancreas, although over half of the islets appeared normal the remainder showed 'degenerative' changes (wrinkling, pyknotic nuclei) or a rather profuse round cell infiltration, or both. Commenting on the rarity of such reports, this author endorsed the opinion of an earlier pathologist, Heiberg, 'who held that these cellular infiltrates made only fleeting appearances, ("hours, never days")'. According to von Meyenburg's interpretation insulitis was probably a reaction to islet cell damage induced by functional overloading, rather than representing a process important in pathogenesis (see 'Implications...', 3-C below).

Perhaps the real credit for awakening interest in insulitis should go to LeCompte (Gepts, personal communication) and to Gepts himself. In 1958 the former reported insulitis in four children dying in ketoacidosis within hours of diagnosis of diabetes. He suspected that in such cases it might not be so rare as previously believed, as it was easily overlooked. Since insulitis seemed confined to recent-onset juvenile diabetes, "one may perhaps be justified in regarding [it] as a potential clue to the nature of the injurious agent in human diabetes, at least of the 'growth onset' type". Gepts (1965) reported insulitis in 15 of the 22 juvenile cases he examined within six months of onset of symptoms. In half of these only a few islets were affected, so that the phenomenon was easily overlooked. In contrast, no insulitis was seen in the longer-standing cases.

The frequency of insulitis is still not clear. Maclean & Ogilvie (1959) mentioned in passing: 'Slight lymphocytic infiltration of some of the islets was noted in three of the [18] acute cases'. Doniach & Morgan

(1973) found none at all in their 13 cases of acute Type 1 diabetes. This difference from the Belgian and North American experience can hardly be explained by geographical variation. Junker et al (1977) suggested that it may reflect heterogeneity within Type 1 diabetes between young children and adolescents or adults. On the other hand insulitis may occur in Type 1 diabetes even in old age, as reported in two cases by Lecompte and Legg (1972). This important question is examined in more detail below (Section C).

3-B-ii. Morphological characteristics of insulitis

Insulitis is described as an infiltration of inflammatory cells within and around some islets of Langerhans. Characteristically the intervening exocrine tissue is not involved, although exceptionally a co-existing polymorphonuclear infiltration amongst the acini has been observed, for example in one of the fulminating childhood cases described by Stansfield & Warren (1928). All authors agree that the great majority of cells taking part in the insulitis proper are small round cells having the appearance of lymphocytes. In addition, smaller numbers of other cells are described in some patients: histiocytes (Gepts, 1965) or large mononuclear cells (LeCompte, 1958), and polymorphonuclear leukocytes (von Meyenburg, 1940; LeCompte, 1958). Plasma cells are sparse (von Meyenburg, 1940) or absent (Gepts, 1965), 'lymphocytes with plasmacytoid features' may be present in small numbers (Gepts, personal communication, 1981).

Gepts & De Mey (1978) and Gepts and LeCompte (1981) report a predilection of insulitis for areas of pancreas with residual B cells. Gepts and Lecompte also state that 'lymphocytes have never been observed in the ... islets composed exclusively of PP cells'. They suggest that 'in recent-onset juvenile onset diabetes, lymphocytes are primarily attracted by an antigen present in the B cells' but not in other endocrine cells. This interesting idea may help to explain why Doniach & Morgan (1973) saw no evidence of insulitis in their single acute case aged under 10y, since no B cells were detected in the specimen from that child.

3-C. IMPLICATIONS CONCERNING THE DISEASE PROCESS

3-C-i. Evidence for heterogeneity of diabetes

The concept of aetiopathogenetic heterogeneity between the two principal clinical categories of primary diabetes is strongly supported by histological evidence. To summarize, in juvenile onset diabetes the pancreas is typically (although not initially) small, with a reduction in islet number and total islet mass which is much more marked than in maturity-onset diabetes. Furthermore, in juvenile but not maturity-onset diabetes B cells are either absent or are sparse and often abnormal. (Gepts & Lecompte, 1981; Rahier et al,1983). Conversely, hyalinosis is seen only in maturity onset case, although often absent (Weichselbaum, 1910; Gepts & LeCompte, 1981). In support of this histological evidence is the finding that the total amount of insulin extractable from the pancreas is much lower in the juvenile disease (Wrenshall et al, 1952).

The finding that islet morphology is heterogeneous in relation to age at onset of clinical diabetes does not alone constitute conclusive evidence for aetiological heterogeneity between Type 1 and Type 2 diabetes as defined in Table 1; it might be argued, for example, that the same underlying process could produce different histological pictures at different ages because of age-related variations in tissue regeneration or in other aspects of response to injury.

Gepts and LeCompte (1981) have commented that islet morphology in truly insulin-dependent diabetes developing in adults is different from that in children, which perhaps supports this notion. Professor Gepts has stated (personal communication, 1981) that he has never observed in adult-onset cases the predominance (90% or more) of pseudoatrophic islets so characteristic of juvenile-onset diabetes. There seems to be no good study comparing islet appearances in the two types of diabetes at the same age of onset, but the gross differences described above could hardly be effects of age itself.

The occurrence of insulitis provides rather more concrete evidence in that it 'has never been observed in maturity-onset, non-insulin-dependent diabetic subjects' (Gepts & LeCompte, 1981). LeCompte & Legg

(1972) described it in two acute cases of true, insulin-dependent, ketosis-prone diabetes developing at ages as late as 64 and 74y, supporting the view that such cases belong to the 'Type 1' category.

The possibility of a relationship between insulitis and age at onset of clinical disease (Junker et al, 1977) is important here. The following new analysis of data from the literature includes only reports where there is sufficient information to allow reasonable categorization of cases as Type 1 diabetes, and where it is clear that cases were not selected with insulitis in mind.

Among the 183 cases reported by Weichselbaum (1910), forty-four had died before age 30. Of these, we exclude four on the grounds that they probably did not have Type 1 diabetes: two had evidence of chronic or relapsing pancreatitis, including one with a cirrhotic liver; a third was acromegalic; and another aged 27 y, with diabetes of 8 years' standing, had no ketonuria before death from 'chronic nephritis'. A further four are excluded because there is no record of duration of diabetes. The remaining 36 cases, and especially those dying within six months of onset, all had one or usually several features of Type 1 diabetes: typical clinical descriptions, ketonuria, ketosis ('Acetonämie') as the certified cause of death, or characteristic islet morphology.

These 36 Type 1 diabetic cases are shown in Table III-2-A, grouped according to age in decades at onset of diabetes and duration of disease at death, measured from onset of symptoms. It is interesting that, of the 18 with onset in the third decade, 14 had survived more than six months (though rarely beyond 2 years), a much greater proportion than in younger age-groups. Thus only four could be classed as 'acute' Type 1 diabetics in the 20 - 29y age group, and little significance can be attached to the fact that none of these had insulitis, whereas the frequency among acute cases dying in the first two decades of life was 50%.

On closer examination, however, we see that among the 36 patients insulitis was almost confined to those aged under 17 years (Table III-2-B), among whom the prevalence in acute cases was 71%. This difference from acute cases aged 17 - 30 years (0% insulitis) may be significant (Fisher's Exact Test: p = 0.021), but because the age of 17 was deliberately chosen to maximize the difference, statistical analysis is not strictly valid and further confirmation is necessary.

Furthermore, although four of the nine cases of insulitis noted among 90 diabetics by Cecil (1909) were aged over 30 at onset, and a fifth case had no age recorded, all these older subjects as well as one

TABLE III-2. INSULITIS IN TYPE 1 DIABETES:
OCCURRENCE IN RELATION TO AGE AT DIAGNOSIS
FROM DATA OF WEICHSELBAUM (1910)

A. SUBJECTS GROUPED BY AGE AT DIAGNOSIS - IN DECADES

Age at onset (years)	0 - 9		10 - 19		20 - 29		0 - 29	
Duration of diabetes (months)	<6	>6	<6	>6	<6	>6	<6	>6
Number of subjects	2	0	8[a]	8	4	14	14	22
Number with insulitis	1	0	4	1	0	1	5[b]	1

Notes:
a. One of these had transient diabetic symptoms three years earlier.
b. A seventh case of insulitis was in a man of 35 with definite Type 1 diabetes beginning three weeks before death.

B. SUBJECTS GROUPED BY AGE AT DIAGNOSIS - ABOVE OR BELOW 17 YEARS

Age at onset (years)	0 - 16		17 - 29		0 - 29	
Duration of diabetes (months)	<6	>6	<6	>6	<6	>6
Number of subjects with insulitis:	7	6	7	16	14	22
number	5	1	0	1	5	2
(%)	(71)	(17)	(0)	(6)	(36)	(9)

Notes:
Frequency of insulitis in younger compared with older group (duration of disease less than six months): P = 0.0210 using Fisher's Exact Test.

aged 28y had changes of chronic or suppurative pancreatitis. If all such atypical cases are excluded, there remain only six acute diabetics with onset before age 30y, only one being under 20y at diagnosis. Of these, insulitis was present in three.

Table III-3 gives combined data from five series of reports, showing that there is indeed a decrease in frequency of insulitis in acute (less than six months) Type 1 diabetes with increasing age over the first three decades of life. This trend is statistically significant.

TABLE III-3. INSULITIS IN ACUTE TYPE 1 DIABETES

Age at onset (years)	0 - 9		10 - 19		20 - 29		0 - 29	
Number of cases: total = N, with insulitis = I	N	I	N	I	N	I	N	I
Cecil (1909)	0	0	1	0	5	3	6	3
Weichselbaum(1910)	2	1	8	4	4	0	14	5
Gepts (1965)	10	9	9	6	3	0	22	15
Doniach,Morgan(1973)	1	0	6	0	2	0	9	0
Junker et al (1977)	4	4	3	1	3	1	10	6
All sources (frequency, %)	17	14 (82)	27	11 (41)	17	4 (24)	61	29 (47)

Difference from equal frequency of insulitis in all three decades (χ^2 test on combined data): $\chi^2 = 17.72$, $p < 0.001$ for 2 deg. freedom.
Analysis for linear trend: $\chi^2 = 11.78$, $p < 0.001$ for 1 deg. freedom.

Thus it is possible that Type 1 diabetes with onset in childhood may have a pathogenesis somewhat different from that in young adults. An alternative possibility is that insulitis occurs early on in the pre-diabetic period at all ages, but persists long enough to be detected at the time of diagnosis in only a proportion of cases. It seems quite likely that the pre-diabetic period is longer in older than in younger patients, just as in untreated cases the progression from diagnosis to eventual demise tended to be slower in the former (Weichselbaum, 1910). If so, then insulitis would be less likely to be *observed* in the older than in the childhood cases even though it might in fact be equally common in both groups.

On balance there is good histopathological evidence for a difference in aetiopathogenesis between Type 1 and Type 2 diabetes, as well as some much less convincing evidence compatible with further heterogeneity within Type 1 diabetes.

3-C-ii. Specificity of the endocrine lesion

Whatever the nature of the pathological process leading to Type 1 diabetes, it predominantly depletes B cells, leaving the numbers of other islet cells relatively unaltered. It does not necessarily follow, however, that the destructive process itself is truly B-cell-specific. An alternative possibility might be that all islet cells are susceptible to damage, and that there is sufficient regenerative capacity to restore the populations of A, D and PP cells but not of B cells. Such a specific inadequacy of B cell regeneration could in theory be either constitutional (perhaps as an inherited abnormality) or acquired (Hellerström et al, 1976).

Although histopathological evidence does not exclude this second hypothesis, it rather favours the idea of B-cell-specific cell destruction in the first place. Thus, as emphasized by Gepts (1965 and later papers), in acute juvenile Type 1 diabetes surviving B cells often appear by histological criteria to be hyperactive in a functional sense. Furthermore, in some cases a few islets consisting largely of hyperactive B cells appear to be regenerating from exocrine ductules. Morphological appearances suggesting cellular degeneration are well recognized in B cells of acute cases, whereas no such changes are reported in A, D or PP cells (Gepts & De Mey, 1978); and there is the intriguing suggestion already mentioned, that lymphocytes involved in insulitis may be attracted specifically to B cells or to an antigen associated with them (Gepts and De Mey, 1978). On the other hand, such evidence applies only to a minority of cases, and it is quite feasible that different mechanisms could lead to the same result of specific B cell depletion in different patients.

3-C-iii. Timing

A recurrent question when considering pathogenesis of Type 1 diabetes is its timing in relation to onset of symptoms. For example, are the B cells destroyed rapidly, over a period of only a few days or weeks, as might be suggested on the basis of the typical rapid progression from earliest symptoms to ketoacidosis? Alternatively, is there perhaps a congenital inadequacy of the B cells, leading to a gradual diminution in number throughout the pre-diabetic life of the patient? Or does the truth lie somewhere between these two extremes? Histology can make only a limited contribution in this situation since it is necessarily retrospective, being based on material obtained post mortem.

However, there are some clues. The small size of the pancreas in childhood cases appeared at first compatible with the view that the whole organ, including the islets, was congenitally inadequate; and although Warren (1927) was not convinced that this was a reliable observation, his finding of a reduction in numbers of islets without evidence of scarred islet remnants led him to support the idea.

Evidence that the pancreas becomes smaller than normal only after diabetes has been established for some time (3-A-ii) is rather against the 'congenital' hypothesis. Likewise, B cell function manifested by C peptide secretion may persist for a while after diagnosis (Ludvigsson & Heding, 1978), just as histological work has shown that the loss of B cells and the characteristic cytoarchitectural abnormalities of the islets may become fully established only after a number of months or a few years (see above). In addition there are the observations of insulitis, degenerating B cells and 'regenerating' islets. These findings all suggest that the pathological process is active at the time of onset of clinical disease.

On the other hand, in even the most acute cases islets showing such evidence of active changes are in the minority, most of the remainder having a well established pseudoatrophic appearance with no detectable B cells (see above). Thus, despite the limitations of histopathological work, it is reasonable to accept as a working hypothesis the opinion of

Gepts (1965): 'It seems probable that in the pancreas of acute juvenile diabetics we ... catch the final stages of a process which has been going for an indefinite time, perhaps from birth on.' This concept forms part of the rationale of the present study.

3-C-iv. Mechanism of B cell destruction

The evidence outlined above indicates that, although minor impairment of B cell regenerative capacity might still be a contributary factor, the immediate cause of Type 1 diabetes is an active and specific destruction of B cells, by a pathological process beginning some time before eventual onset of clinical diabetes and generally still active at diagnosis. What, then, is the nature of the active destructive process? Here, insulitis is likely to be a major clue.

Lecompte (1958) discussed several possible explanations for insulitis. The first possibility was that it represented a reaction to viral or bacterial invasion of the islets. This he considered quite likely in the few reported cases showing an intense cellular infiltration of many islets, but less so in the majority where only a few islets were affected since it seemed to him unlikely that an infective agent should be so selective. However, in recently diagnosed Type 1 diabetes only a minority of islets contain any B cells, so an infective agent exhibiting specific tropism for B cells might well induce a similarly patchy insulitis. In the single case reported by Gepts and De Mey (1978) insulitis was indeed associated with B-cell-containing islets.

It might be argued that islets containing B cells but free of insulitis do occur in acute cases of the disease. However, not all B cells in a given pancreas need express viral antigen at the same time. Some might have escaped involvement in a light infection, or be newly formed by regeneration since resolution of viraemia; or possibly viral replication within them is not yet sufficiently advanced. Furthermore, the infiltrate is very often sparse and confined to small sections of the affected islets, so that unless serial sections of each islet are examined it cannot definitely be stated that insulitis is absent.

In certain animals viral infections are known to cause diabetes, either naturally or experimentally. The insulitis described in such animals is typically very similar to that seen in human diabetes, both in its cellular make-up and in its patchy distribution (Craighead, 1976; see also Chapter 8). Thus, in the light of current understanding, viral or possibly bacterial infection of B cells seems an acceptable explanation for insulitis.

The second possible explanation discussed by Lecompte (1958) was that of von Meyenburg (1940), who considered it a reaction to cell damage induced by functional overloading. Lecompte criticized this on the grounds that it was too vague, and pointed out that in some types of experimental diabetes (such as that due to pituitary or adrenal hormones, partial pancreatectomy or low doses of alloxan) the remaining B cells were undoubtedly overloaded and yet insulitis was never seen.

Lecompte (1958) particularly favoured a third idea, which was that insulitis represented a non-specific response to B cell injury by some unknown pathogenic agent. In this context it is interesting that insulitis often consists of no more than a scattering of lymphocytes in the connective tissue around the periphery of an islet (the 'Kapsel' of Weichselbaum, 1910), rather than spreading intimately among the endocrine cells themselves. Such examples, it has been suggested, might well represent a reaction to antigenic substances released by damaged B cells into the vascular outflow from the islet (Gepts et al, 1976).

A fourth possibility not mentioned by Lecompte (1958) is that the infiltrating immunocytes might actually be autoaggressive, and thereby responsible for B cell destruction. By analogy, lymphocytic infiltration is invariably present in endocrine glands affected by three conditions widely accepted to be autoimmune diseases: autoimmune thyroiditis, Addison's disease, and hypergonadotrophic hypogonadism (Doniach et al, 1979; Irvine & Barnes, 1975). The analogy is not perfect, however, since plasma cells, although so conspicuously absent from islets affected by insulitis, are almost invariably present in the three diseases mentioned; indeed, in Hashimoto's disease of many years' standing they predominate

over lymphocytes. Germinal centres also feature in autoimmune thyroiditis and sometimes in Addison's disease, but not in insulitis. This latter difference could perhaps reflect the transience of insulitis and the small size of the islets, but the lack of plasma cells sugests that if insulitis does represent an autoimmune process, then the immune mechanisms involved may not be quite the same as those in other autoimmune endocrinopathies.

In several species of animals, autoimmune insulitis similar to that seen in man has been induced by immunization with homologous or heterologous 'insulin' (that is, recrystallized insulin preparations usually including small amounts of impurities) in Freund's adjuvant (Toreson et al, 1964; Toreson et al, 1968; Renold et al, 1966; Renold et al, 1969; Klōppel, 1976). Such insulitis is typically lymphocytic and affects only a proportion of islets. Although there may be evidence of B cell destruction and compensatory hyperactivity, diabetes is an uncommon result. In rabbits, development of insulitis depends on the strain of animal used (Kloppel, 1976).

When small doses of streptozotocin, insufficient to cause direct destruction of islet B cells, are administered repeatedly over a few days to certain strains of laboratory rodents, insulitis involving mononuclear cells develops and may be followed after about a week by hyperclycaemia (Like & Rossini, 1976 and 1978). This effect probably involves an autoimmune process (see 7-D-ii). It is particularly interesting that some strains of mice, not susceptible to either low-dose streptozotocin or infection and betacytolysis by diabetogenic virus, may become diabetic when exposed in sequence to both agents (Toniolo et al, 1980). Thus, different environmental agents may interact to cause insulitis (see 8-D)

Another interesting animal model is provided by four strains of New Zealand Black mice which exhibit spontaneous generalized autoimmunity, manifested as a lupus-like syndrome. Kolb et al (1980) have reported that insulitis is common in these animals, and that under the electron microscope the infiltrating mononuclear cells appear specifically cytotoxic to B cells.

Insulitis is also found during the development of spontaneous ketotic diabetes in an outbred strain of Wistar rats, known as 'BB' rats. At onset of overt diabetes the BB rat has severe mononuclear-cell insulitis, few if any islets being spared, with destruction of virtually all islet B cells within a few days (Nakhooda et al, 1977). Such a fulminating course is rare in human diabetes, but it is now recognized that many animals of the same strain develop milder insulitis, patchily distributed within the pancreas, which is apparently identical to that in man. Animals with insulitis may be metabolically normal at the time of study, or may have mild hyperglycaemia (Marliss et al, 1982). There is mounting evidence to implicate autoimmunity in pathogenesis of the diabetes of BB rats (see 7-D-iii).

An interesting feature of these animal models is that plasma cells are not seen in experimental insulitis. Klöppel has related that when experimental insulitis and thyroiditis were induced simultaneously in a rabbit, plasma cells were present in the thyroid but not in the islets (Gepts et al, 1976; p.140) If this anecdotal report is confirmed, then the lack of plasma cells in human insulitis does not necessarily argue against the view that this, too, has an autoimmune basis.

In conclusion, insulitis may well be a clue to the mechanism of B cell damage in human Type 1 diabetes. If so, it is compatible with both viral and autoimmune hypotheses, or indeed a combination of both. Further evidence relating to autoimmunity in Type 1 diabetes is described in Chapter 7, and possible viral involvement in aetiology is covered in Chapter 8.

PART III — CHAPTER 4

EPIDEMIOLOGY OF TYPE 1 DIABETES

CONTENTS

A. Estimates of prevalence and incidence 32
i. Prevalence 32
a. All ages 32
b. Childhood and adolescence 33
ii. Incidence 34
iii. Sex distribution 36
iv. Geographical and other sources of variation 37
v. Age-specific prevalence and potential prevalence . . . 38
B. Other aspects of epidemiology of Type 1 diabetes 45
i. Seasonal variation in incidence 45
ii. Clustering 47
iii. Secular trends 47
iv. Climate, race and socioeconomic factors 49

The aim of this chapter is twofold: (A) to construct from published data a curve relating cumulative frequency of Type 1 diabetes to age, in order to facilitate interpretation of some of our own findings (Part VIII, Chapter 4); and (B) to examine briefly certain other epidemiological aspects of the disease.

4-A. ESTIMATES OF PREVALENCE AND INCIDENCE

4-A-i. Crude prevalence

a. All ages

Despite many attempts to measure overall prevalence of diabetes in various communities (West, 1978, p. 132), until very recently no reliable information seems to have been published concerning Type 1 as opposed to Type 2 diabetes, except insofar as most childhood cases are Type 1.

Of the features listed in Table III-1 (page 6), the only criterion of Type 1 diabetes suitable for epidemiological work, apart from age, is probably insulin dependence. Even this is not a reliable guide (see Chapter III-1). In particular, insulin usage in Type 2 diabetes has fluctuated widely over the past few decades, particularly in the United States, since it depends not only on factors such as availability of oral hypoglycaemic agents but also on the ebb and flow of medical opinion.

Nevertheless, the frequency of insulin-treated diabetes in a population should give at least an approximate indication of prevalence of Type 1 diabetes. West (1978, pp 29-30) cites estimates of 21% to 83% for the proportion of known diabetics taking insulin in various places, with a figure of 28% in USA in 1964-5, a time when both sulphonylureas and biguanides were widely used. Thus at that time about one fifth of known diabetics were probably insulin dependent. If it is assumed that recognized diabetics constituted 0.6 - 1.6% of the population (Harkness, 1962: 0.6% in Halstead, England; Working Party of the College of General Practitioners, 1962: 0.6% in Birmingham, England; Palumbo et al, 1976: 1.6% in Rochester, Minnesota, USA), then as a very rough approximation some 0.1 - 0.3% would have had Type 1 diabetes.

Direct determinations of prevalence of insulin-treated diabetes in the whole population have been performed only recently. Hedley et al (1982; presented to British Diabetic Association in 1980) reported that 0.3% of the population of 871,000 in Nottingham, England were receiving insulin treatment. One third of these patients were aged under 40 years, and there was a male excess of 12%. Cases were ascertained through insulin prescriptions and hospital clinic records, and validated by

comparison with general practitioners' records. A similar method gave a crude prevalence of 0.36% in males and 0.33% in females in Funen County, Denmark, where the total population numbers 450,000 (Green et al, 1980).

4-A-i-b. Children and young people

A number of studies have reported crude prevalence in children of 'paediatric' age (e.g. 0-14 years), of school age, or sometimes aged 0 - 20 years. Most earlier reports which did not take insulin treatment as a criterion need not be excluded from consideration, because only a minority of overtly diabetic white Caucasoid subjects aged under 20 are not insulin treated (e.g. 2% - Silwer, 1958). However, some surveys also include milder glucose intolerance discovered during screening for military service (Proust and Smithurst, 1968) or during pregnancy.

In 1964, according to Beardmore and Reid (1966), 39 of 45,500 schoolchildren aged 5-16 were known to be diabetic in Northampton, England, giving a crude prevalence of 0.086%. These cases had been accumulated over several years by the School Health Service from schools and physicians within the community. At that time the population of Northampton was predominantly white and of European origin (author's personal experience, 1955-1966). Outside the United Kingdom, the prevalence in white schoolchildren (aged 4-17) in Michigan, USA was reported as 0.16% (Gorwitz, Howen and Thompson, 1976). In both these school studies there are several possible sources of error. Thus, cases were often reported by non-medical personnel, particularly in Michigan, and there was no validation of ascertainment rate by an alternative approach. All the Northampton cases were insulin-treated, but as many as 15.5% of those in Michigan were said not to be, which casts some doubt on relevance of this study to Type 1 diabetes. Furthermore it was acknowledged that some elementary schools were not included in the Michigan study. This would tend to inflate overall prevalence figures.

In Rochester, Minnesota the mean prevalence of diabetes in school age children between 1945 and 1970 was 0.12% (Palumbo et al, 1976). In the same state, Kyllo and Nuttall (1978) found a prevalence of 0.19%

among children attending state schools in 1975. Here, 95% of the diabetics were on insulin.

'Paediatric' surveys include children below school age, and would be expected to give lower overall prevalence figures than school surveys. Thus in children under 15y, Palumbo et al, (1976) obtained a figure of 0.057% in Rochester, Minnesota; and in France Lestradet & Besse (1977) found only 0.024%, in a study based on several sources of information which seems unlikely to have missed many cases.

4-A-ii. Incidence studies

Incidence studies have the advantage that, at least in Europe and North America, nearly all cases of Type 1 diabetes are referred to a hospital or specialist on diagnosis (Christau et al, 1977; Åkerblom, Reunen & Käär, 1980). A high level of ascertainment is therefore easier to achieve than in cross-sectional prevalence studies. Furthermore, loss due to associated mortality is not a problem, and any fluctuations in incidence are obvious. Both these factors probably contribute to the 'frequency reduction' noted by Falconer, Duncan and Smith (1971) - that is, a decrease in the apparent frequency of diabetes in a specific age-at-onset class with increasing duration of diabetes.

The authors of that outstandingly careful and balanced paper were able to correct for frequency reduction and to calculate age-specific incidence (morbidity risk) and potential prevalence (i.e. cumulative frequency of diabetes assuming no loss with increasing duration) from prevalence data obtained in Edinburgh. From one of their Figures it appears that annual incidence in the group aged <5y was about 3.10-5 at risk, rising to 6 at age 5-9 and 11 at 10-14y. The ascertainment rate was high - around 90% for all ages and probably higher in children - but insulin treatment was not a criterion, so only the younger onset-age groups are relevant here. Unfortunately, the numbers were rather small when these groups were further subdivided by duration of disease in order to correct for the frequency reduction. The estimates of incidence in childhood are therefore only approximate.

Direct measurements of incidence are mostly rather recent. From the voluntary Paediatric Register of the British Diabetic Association, Bloom, Hayes & Gamble (1975) obtained an overall figure of 7.7 $.10^{-5}.y^{-1}$ (age 0-16). There was however a wide variation between different areas of UK and Ireland, mainly attributable to patchy ascertainment, and the true figure is probably closer to the highest estimate of 13 $.10^{-5}.y^{-1}$ in the South of England.

Lestradet & Besse (1977) found a lower annual incidence for insulin-treated diabetes among French children of 0-14y in 1975, rising from 2.0 in the group aged <4y to 5.5 aged 10-14, with an overall mean of $3.7 \cdot 10^{-5}.y^{-1}$. The ascertainment method seems reasonably complete.

Perhaps the most reliable data come from Denmark. Christau et al (1977) achieved nearly complete ascertainment of patients aged <30y who presented with insulin-dependent diabetes over a period of 5 years in the Copenhagen area, and reported a mean overall age-corrected incidence of 13.2 $.10^{-5}.y^{-1}$ (14.5 in males and 11.9 in females) - very similar to that in the South of England. Elsewhere in Scandinavia incidence generally appears to be higher (all in children of 0-14y): $20.10^{-5}.y^{-1}$ in Sweden, 1970-78 (Sterky et al, 1978), and $29.10^{-5}.y^{-1}$ in Finland, 1970-79 (Akerblom, Käär & Reunanen, 1981) although in Northern Norway, 1965-74, Bratlid et al (1976) reported only $7.6.10^{-5}.y^{-1}$. This variation is discussed further below (A-iv; B-iv).

In the New World, incidence figures reported in schoolchildren include 14 $.10^{-5}.y^{-1}$ in Rochester, Minnesota (Palumbo et al, 1976), and 20 $.10^{-5}.y^{-1}$ in Michigan (North, Gorwitz & Sultz, 1977), although the latter figure was based on data of Gorwitz et al (1976) and may not be accurate (see above). For insulin-dependent diabetes, West et al (1979) found a mean annual incidence of 8.8 $.10^{-5}.y^{-1}$ in Montreal children aged <17y, while in Pittsburgh the rate was 16 $.10^{-5}.y^{-1}$ in white boys and 15 $.10^{-5}.y^{-1}$ in white girls aged <20y between 1965 and 1976 (Laporte et al, 1981). In the same age group the incidence of diabetes in New Zealand was 10 $.10^{-5}.y^{-1}$ (Crossley & Upsdell 1980).

4-A-iii. Sex distribution

Estimates of male:female incidence ratios vary from one report to another, according to the geographical situation, age distribution and ascertainment method of the sample studied, and depending on whether incidence or prevalence is investigated.

Vinke et al (1959) examined prevalence of diabetes in a population aged 15y or more in Groningen, The Netherlands (at a time when Type 2 diabetes was rare under age 40y because of the depredations of the Second World War). They found a male:female ratio of 1.4 in the group aged 15-

TABLE III-4. MALE:FEMALE RATIOS FOR NEW CASES OF TYPE 1 DIABETES

Reference		Number aged 0-14y	M:F ratio in age-groups 0-4y	5-9y	10-14y	15-19y	0-14y	0-19y
INCIDENCE RATIOS[a]								
Christau et al	1977	255	1.23	0.87	1.17	1.47	1.07	1.16
Akerblom et al	1980	2641	1.20	0.98	1.11	-	1.07	-
Weighted mean		2896	1.20	0.97	1.12	-	1.07	-
PROPORTIONS OF CASES[b]								
Bloom et al	1975	2330	1.28	0.94	1.14	-	1.09	-
Lestradet & Besse	1977	467	1.76	0.88	1.11	-	1.11	-
Christau et al	1977[c]	255	(1.27)	(0.93)	(1.26)	(1.56)[c]	(1.16)	(1.25)
West et al	1979	483	1.10	0.85	0.97	-	0.95	-
Akerblom et al	1980	2641	1.25	1.02	1.15	-	1.12	-
Laporte et al	1981	682	1.18	1.22	0.89	1.61	1.03	1.13
Weighted mean[d]		6603	1.28	0.99	1.10	-	1.09	-

Notes: a Corrected for age and sex distribution in general population
b Uncorrected for general population characteristics
c Calculated from age- and sex-corrected incidence figures and total numbers of males and females, and therefore approximate. Not included in mean.
d Excluding Christau et al (1977)

29y, the male excess disappearing after age 40y. Cohen et al (1970) likewise found a ratio of 1.4 in Israel, ages 2-16y. Falconer et al (1971) found approximately equal prevalence of diabetes in the two sexes in Edinburgh for ages up to 40y, but the frequency reduction effect (see above) was greater in the males and after allowing for this the 'annual morbidity risk' in males exceeded that in females, except in the very young. Thus incidence data are preferable to prevalence studies in this respect.

Most incidence studies show that the major peak on the incidence-versus-age-at-onset curve is a year or two earlier in girls than in boys (see below). As a result, for onset-ages between about 7 and 11y the sex ratio is often close to 1.0, whereas before and after this most reports show a male excess (Table III-4).

4-A-iv. Geographical and other sources of variation

There are several possible explanations for the wide variations in prevalence and incidence reported by different authors. Differences in age distribution between the populations examined would obviously be important in prevalence studies, but certainly cannot account for the very striking variation in incidence. Incomplete ascertainment may have led to the low overall figure obtained by Bloom et al (1975) in UK, but again is unlikely to be so marked as to explain the substantial difference in incidence between, for example, Finland and France.

Careful examination of the data suggests a consistent geographical trend. Over fifty years ago Adams (1926) pointed out a much higher recorded mortality from 'acute diabetes' in northern than in southern states of the USA. It is not clear how far these cases represented Type 1 diabetes. In Europe the annual incidence of Type 1 diabetes in children increases progressively from around $4.10^{-5}.y^{-1}$ in the south (3.7 in France and 4.65 in Venice; Lestradet & Besse, 1977) through $13.10^{-5}.y^{-1}$ in Denmark and $20.10^{-5}.y^{-1}$ in Sweden to $29.10^{-5}.y^{-1}$ in Finland, the latter being the highest reported anywhere in the world (see above for references). Within New Zealand, South Island has 1.4 times the incidence of North Island.

These geographical trends have interesting implications for aetiology which are discussed in more detail below (3-B-iv), but they also mean that great care must be taken when data from one country are extrapolated to another. Other sources of variation between different studies of incidence include short and long-term fluctuations and maintained secular trends (see 3-B-iii).

4-A-v. Age-specific prevalence in UK

Later in this thesis (VIII-3) a comparison is made between the frequency of Type 1 diabetes in siblings of affected children and the age-corrected frequency in the general population, over the age range 0-33y. Unfortunately the age-specific potential prevalence curve derived by Falconer et al (1971) is insufficiently detailed for our purpose, and above age 20y or so it would include a significant proportion of Type 2 diabetes. Indeed, there is as yet no UK study giving the necessary information. We have therefore derived a curve relating cumulative frequency of Type 1 diabetes to age, by combining age-specific data from other countries in such a way as to compensate for differences in overall incidence.

Ten published reports present sufficiently detailed information on distribution of age at onset in a sample of diabetics, incidence of new cases within a population, or age-specific prevalence. Of these, however, the two reports on Michigan school-children already mentioned (Gorwitz, Howen & Thompson, 1976; North, Gorwitz & Sultz, 1977) are excluded on the following grounds: ascertainment was incomplete, with a probable age bias; there was an obvious deficit of adolescent cases; 15.5% of the diabetics were said not to be on insulin; and there was no accurate information on the age distribution of the population actually sampled. The remaining eight papers are summarized in Table III-5.

Because of the problems of geographical variation and differing presentation of information between these eight studies, the data were first transformed as follows. Age-specific cumulative frequency was calculated from incidence data (Q in Table III-5); or, from data on age

TABLE III-5. SOURCES OF DATA USED IN FIGURE III-1

	Authors	Year published	Place	Type of study [a]	Subjects in this analysis				
					Age range	No. aged <15			All ages
						M	F	Total	
1	Bloom et al	1975	UK	DNR	<16	1211	1115	2326	2424
2	Christau et al	1977	Denmark	QNRI	<29	137	118	255[b]	474
3	Lestradet & Besse	1977	France	DNRI	<15 [c]	246	221	467	467
4	Kyllo & Nuttall	1978 [d]	Minnesota	PKS	5-16	-	-	994	1533
5	Durruty et al	1979	Chile	DNHI	<30	155	173	328[e]	525
6	Fleegler et al	1979[d]	Pittsb'gh, Texas, Florida, Melbourne	DI N+K	?	1682	1778	3460	3538[f]
7	West et al	1979	Montreal	DKHI	<17	235	241	476	516
8	Laporte et al	1981	Pittsb'gh	QNHI	<20	347	336	683[g]	865

NOTES
(a) KEY: D = distribution of age at onset in a population of diabetics; H = hospital-based survey; I = insulin-treated diabetics only; K = known cases; N = newly-diagnosed cases; P = prevalence study; Q = incidence study; R = registry of cases used.
(b) Numbers calculated from age-specific incidence data given.
(c) Data on subjects aged 15-19y excluded because stated to be un-reliable.
(d) NOT used in combined analysis (Figure III-2) - see text.
(e) Corrected for 0.5 year age shift (ages originally expressed to nearest birthday).
(f) Total of male and female subjects in Results = 3538; total stated in Summary = 3241.
(g) Calculated from age-specific incidence data given (white subjects only).

distribution within a sample of diabetics (D in the table), the cumulative number at each year of age was determined. From the study of Kyllo & Nuttall (1978) age-specific prevalence data were taken direct. The result for each age class in a given study was first expressed as a fraction % of the value in the same study at a reference age of 14y, which was the greatest age included in all the reports. Where possible only white subjects were included and the sexes were analysed separately. Figure III-1 shows the graphs obtained when both sexes were combined. Graphs for males and females are not shown separately here, but are broadly similar.

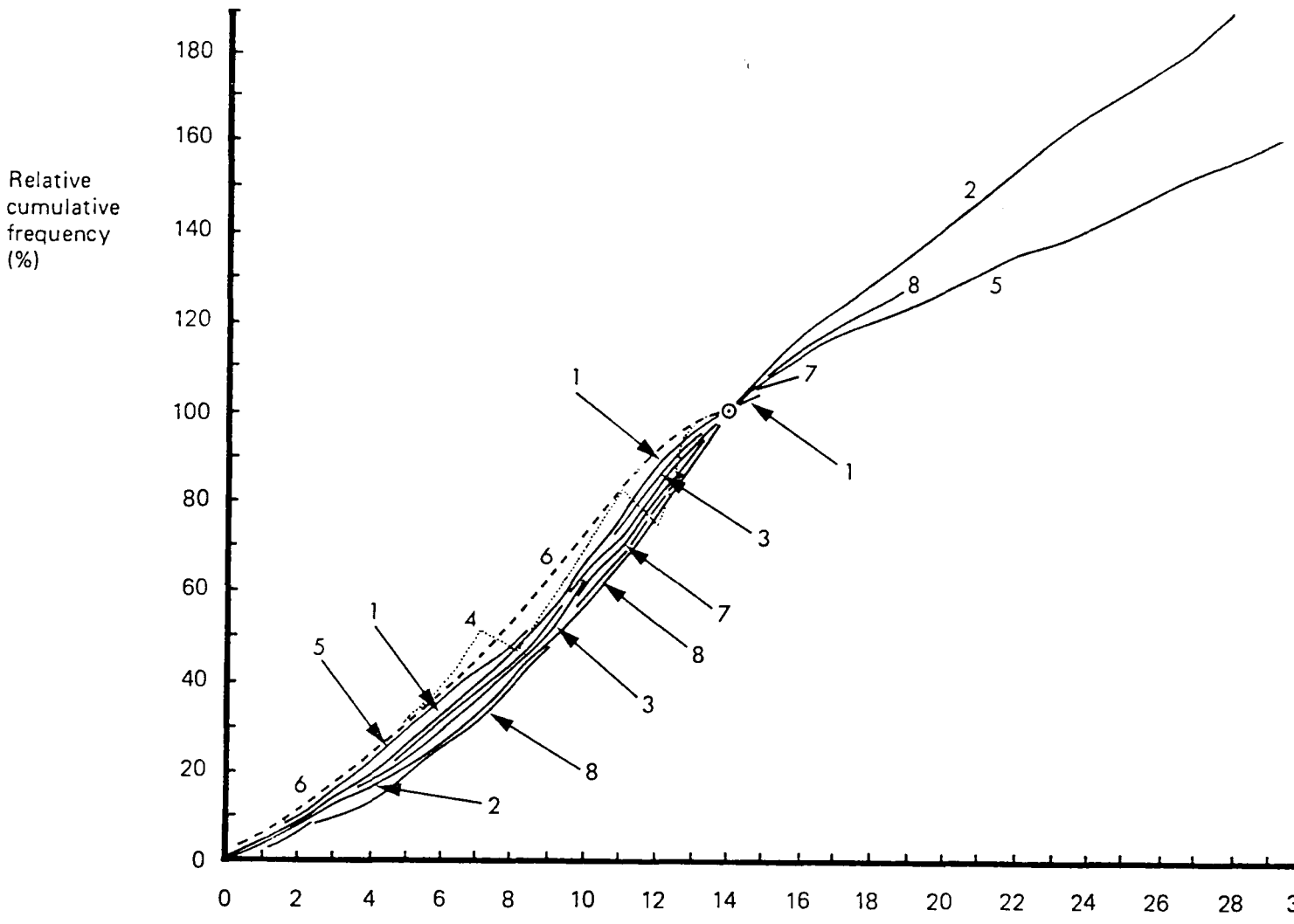


FIGURE III-1. RELATIVE CUMULATIVE FREQUENCY OF TYPE 1 DIABETES, RELATED TO AGE. From the eight studies detailed in Table VIII-5. For each study, the cumulative frequency at each year of age was expressed relative to that at age 14y and plotted as a percentage.

In Figure III-1, curve 6 (Fleegler et al, 1978) is obviously out of line; this is probably due to age-biased sampling in their very heterogeneous group of children, who were ascertained by several different methods. Furthermore, the upper age limit is not given, and the total number of patients stated in the Summary is different from that in the Results. This study is therefore excluded from the final analysis. Curve 4 (Kyllo & Nuttall) is also atypical, partly because of the frequency reduction effect, partly because only a relatively small number of cases is informative for each point on this non-cumulative curve, and probably also due to uneven ascertainment. This, too, is excluded.

The remaining six curves are more comparable, although above age 20y only two studies are informative, and that from Chile (5) gives consistently lower values than Christau et al (1977) (2). The principal reason for this discrepancy is probably that the Chilean data were not corrected for the age distribution of the population at risk, an important consideration because the proportion of Chilean inhabitants aged <20y is 45% - approximately double that in most developed countries (Jackson, 1981). Another possible reason for the difference between curves 2 and 5 is that the latter was based on known cases identified through clinical records, and thus would show a 'frequency reduction effect', whereas the Danish work is a true incidence study.

At lower ages, Curve 1 (UK) shows a marked falling-off in gradient after age 12 when compared with others. This can probably be explained by relative under-ascertainment of the older children; below age 12 this curve appears to run about 10% above the reliable Danish curve (2), suggesting a 10% under-ascertainment at the reference age of 14y.

With the exception of curves 4 and 6, the data summarized in Figure III-1 have been combined by calculating the weighted mean of the cumulative frequencies at each age. The UK figures (1) were first reduced by 10% and data above age 12 were excluded (see above). The weighting factor was in most cases equal to the number of diabetic subjects aged <15y, but because of the reservations expressed above the Chilean data were arbitrarily given only half this weighting, and the reliable Danish study was given 50% extra weighting. Linear extrapolation extends the resulting mean curves to age 33. Weighted means are given in Appendix 6-2 (column 3) and plotted graphically in Figure III-2.

A number of assumptions are involved in these calculations, quite apart from the possibility of biased ascertainment already mentioned. First, most of the reports examined (Table III-5) either express age-specific results in terms of age at last birthday, or do not state exactly what they mean by 'age'. In one case (Durruty et al, 1979),

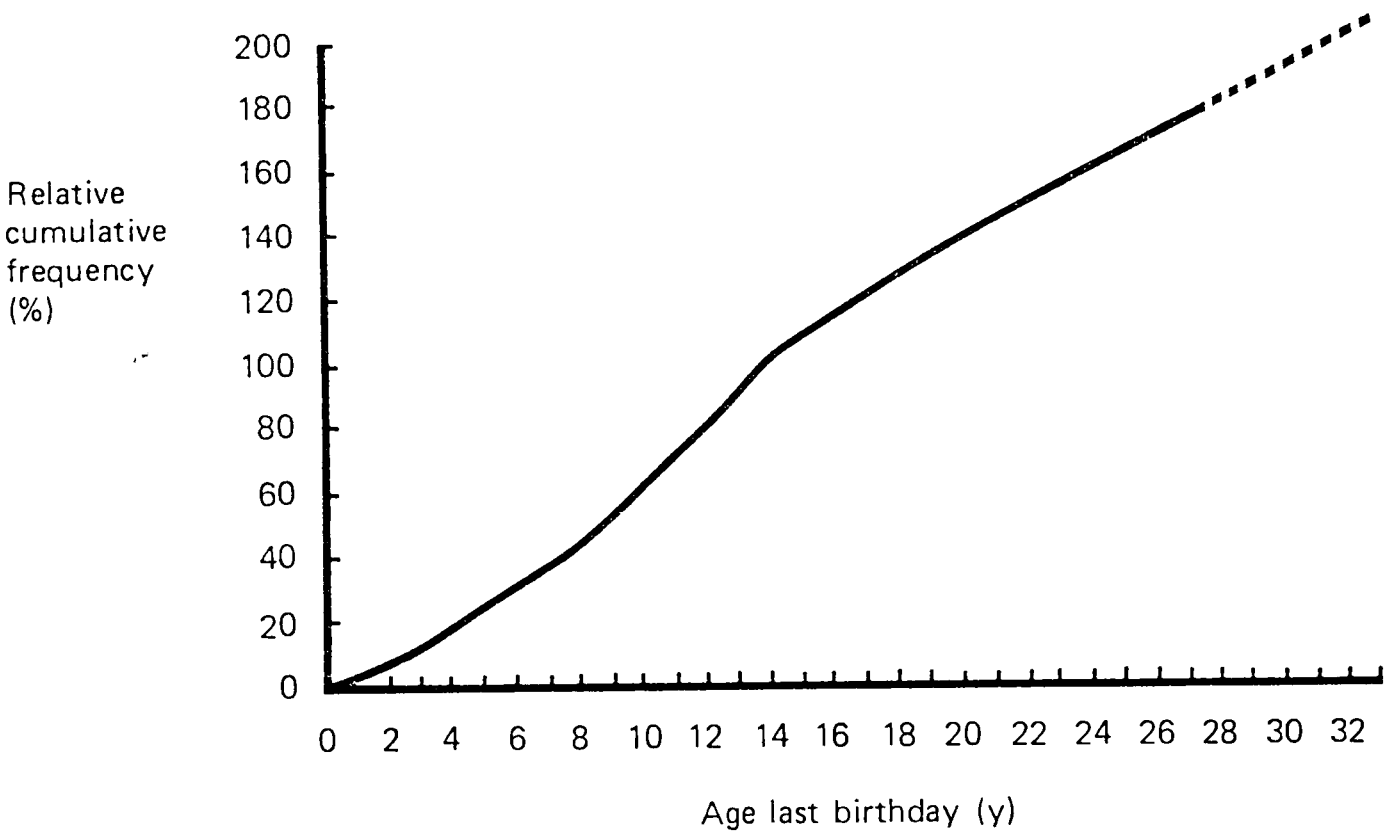


FIGURE III-2. AGE-SPECIFIC CUMULATIVE FREQUENCY OF DIABETES, BOTH SEXES. Based on age at onset in 4539 cases aged <15y and in 810 insulin dependent cases aged 15-29y. Calculated as weighted means of data summarized in Figure III-1, excluding Curves 4 and 6.

where age is apparently expressed to nearest birthday, we have corrected the data by interpolation. Any error from this source is likely to be small.

Secondly, there may be genuine differences in relative age distribution at onset of Type 1 diabetes between UK and other countries. For example, while the major peak is usually in early adolescence, a subsidiary peak earlier in childhood is often described. This appears at age 4-5y in UK (Bloom, Hayes & Gamble, 1975) and perhaps in France (Lestradet & Besse, 1977), 6y in Chile (Durruty et al, 1979) and 7-9 in Scandinavia (Christau et al, 1977; Sterky et al, 1978; Åkerblom, Käär & Reunen, 1981), whereas in New Zealand it was not apparent at all to Crossley & Upsdell (1980). (Bloom et al (1975) suggested that such variation may be related to age at school entry.) There is no reliable comparative information on incidence beyond age 20y.

Finally, there may be secular trends or cyclic variations of incidence within the same geographical area. This possibility is discussed below; if it is true, then old data may not apply.

Nevertheless, these curves are probably reasonably accurate, and as yet no better alternatives are available. The next question concerns the level at which to fix them so as best to describe the population in the area of our own study. The potential prevalence at age 20y, according to Falconer et al (1971; Fig.11), was about 0.18% in Edinburgh; if the curve for both sexes in Figure III-2 is set at this level, the value at our reference age (14y) is about 0.14%. The cumulative frequency at 14y in Denmark was higher at 0.21% (from data of Christau et al, 1977). However, it cannot be assumed that these estimates apply to south-east England.

The population of schoolchildren studied by Beardmore & Reid (1966) was in Northampton, only about 80km from the centre of the area of interest here. It seems reasonable to use the crude prevalence in that population (0.86%) to set the absolute level of our relative potential prevalence curves, provided allowance is made for age distribution and the 'frequency reduction' effect.

In general, the mean potential prevalence over an inclusive range of one-year age classes a to b, $P_{a,b}$, is given by:

$$P_{a,b} = P_c \cdot \frac{\sum_{i=a}^{b} r_i}{100(a - b + 1)} \qquad (4\text{-}1)$$

where P_i is absolute potential prevalence at age i and r_i is (%) potential prevalence relative to that at age c. This assumes an even age distribution in the population over the range a-b, and is thus an approximation. From Figure III-2-c,

$$P_{5,15} = P_{14} \times 0.62 \qquad (4\text{-}2)$$

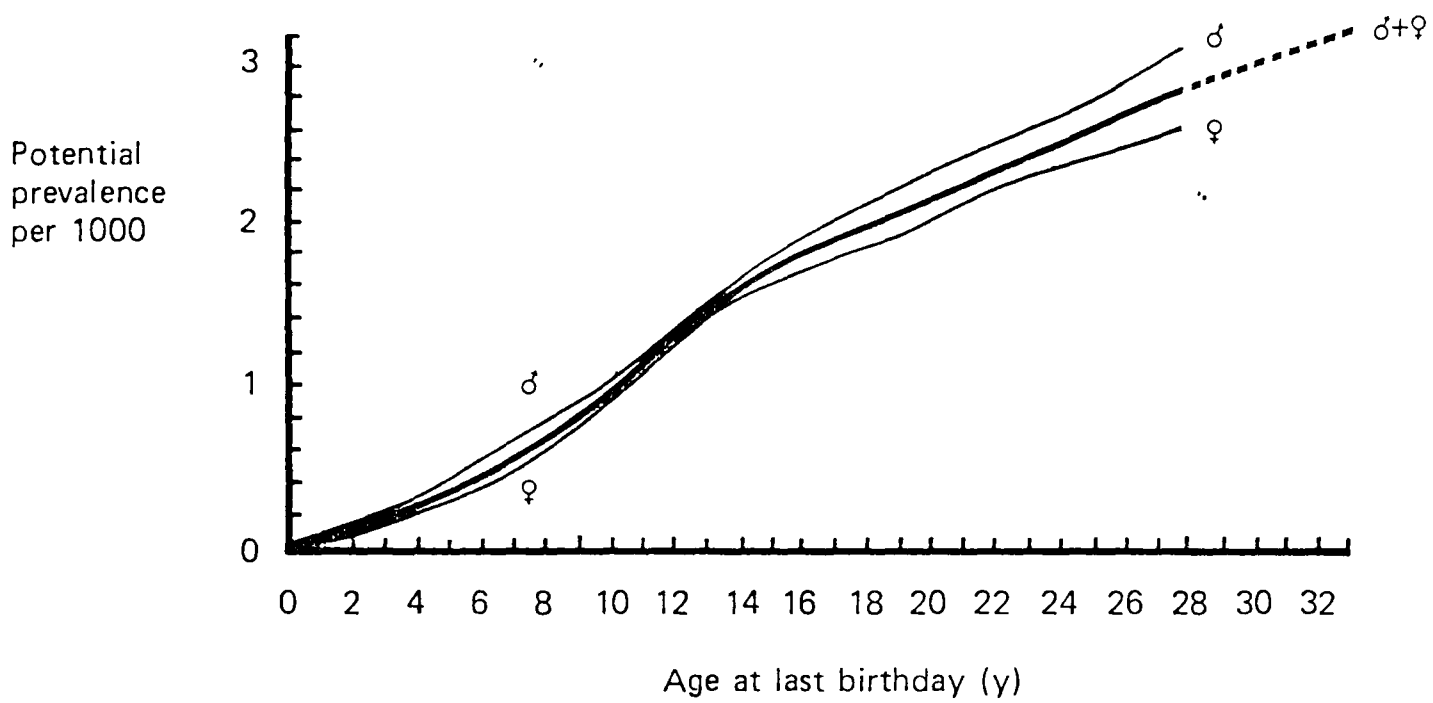


FIGURE III-3. POTENTIAL PREVALENCE OF TYPE 1 DIABETES. Calculated from data of Fig III-2, taking P_{14} = 0.16 (see text)

In order to compare potential prevalence with actual prevalence a correction is needed for the 'frequency reduction' effect, which varies with age (Falconer et al, 1971). In Denmark, the mean prevalence of Type 1 diabetes is 0.082% in the age group 5-9 and 0.185 at ages 10-14 (Green et al 1980). The potential prevalence figures in the same country and the same age groups, calculated directly from the incidence data of Christau et al (1977), are 0.087 and 0.193%, exceeding the observed values by a factor of 1.1 (to one decimal place). Thus the mean potential prevalence in Northampton schoolchildren, $P_{a,b}$, would probably exceed the observed prevalence by about 10%. Hence, from Equation 4-2 and the Northampton figure,

$$P_{5,15} == 0.62 \times P_{14} == 0.086 \times 1.1 ,$$

and so

$$P_{14} == 0.16\% .$$

Figure III-3 shows age-specific potential prevalence in absolute terms calculated on this basis. For the separate curves relating to the two sexes, P_{14} in males was assumed to exceed that in females by 7%, as

in the data of Christau et al (1977) and Akerblom et al (1980) (see above).

It might be argued that the recent Danish figures are more appropriate than the older, less reliable Northampton estimate of prevalence, despite geographical separation from the area of interest. In that case the values of Figure III-3 should all be increased by a factor of 0.21/0.16 = 1.31. Even better would be to obtain a new figure by direct observation (see VII-B-i and VIII-4).

4-B. OTHER ASPECTS OF EPIDEMIOLOGY OF TYPE 1 DIABETES

4-B-i. Seasonal variation

In Minnesota, USA, Adams (1926) observed a significant excess of acute onset diabetes in autumn and winter months. Although some of his cases may not have had Type 1 diabetes, others have reported similar observations in school-age children in various parts of the world: UK (Gamble & Taylor, 1969; Patterson et al, 1981b), Denmark (Christau et al, 1977), Finland (Akerblom et al, 1981) and Iceland (Helgason & Jonasson, 1981) in Europe; Chile (Durruty et al, 1979) and Australia (Fleegler et al, 1979) in the Southern Hemisphere; and several states in both northern and southern USA (MacMillan et al, 1977; Kyllo & Nuttall, 1978; Fleegler et al, 1979). Several other series showed a similar but non-significant pattern. There is a single exception: even using a sophisticated spectral analysis technique, Crossley & Upsdell (1980) found no consistent seasonal variation in onset of 608 cases in New Zealand.

In some of these studies, even though school-age children did show seasonal variation, pre-school children did not (Bloom et al, 1975: 308 cases in UK*; Patterson et al, 1981 in Scotland; Fleegler et al: 300 cases in Australia). This is not simply a failure to reach statistical

---

* DR Gamble has since found a minor trend to seasonal variation in pre-school children in UK (personal communication, 1982).

significance because of small numbers, since the distribution diagrams show no evidence of any seasonal trend at all. The observation is all the more puzzling in that a clear seasonal pattern in such very young children is apparent in studies from France (Lestradet & Besse, 1977: n = 379), Finland (Akerblom et al, 1980: n = 436), and in three US cities (Fleegler et al, 1979: n = 621). The negative findings of Macmillan et al (1977) in Minnesota can probably be disregarded because so few children (34) were under 6y old.**

Beyond school-leaving age, the few informative studies agree that the seasonal variation fades away or disappears, at least in Chile (Durruty et al, 1977), Scotland (Gray, Duncan & Clarke, 1979) and Iceland (Helgason & Jonasson, 1981)

The most popular theory to explain seasonal variation is that it may reflect prevalence in the community of unidentified viruses important in precipitating metabolic decompensation, if not actually causative. This is discussed in Chapter 8. The lack of seasonality beyond school age could mean that different processes operate (i.e. that there is aetiological heterogeneity), or may simply be a result of slower progression of the disease process (cf. insulitis, Chapter 3-C-i).

An alternative theory, suggested by Adams (1926), is that a harsh winter climate enforces physical inactivity and that inappropriately high food intake places increased load on the failing B cells at that time. Adams supported this by citing national statistics which showed much higher mortality from diabetes in northern than in southern states of USA, the latter having mild winters. However, this argument breaks down in the face of the new data showing a seasonal incidence pattern even in southern states -Florida and Texas - as well as in Australia (Fleegler et al, 1979).

** This paper is misleading, and has been quoted as showing seasonal variation in pre-school, but not in school-age children (West, 1978). In fact, the distribution diagrams for the two age groups have obviously been interchanged in printing!

4-B-ii. Clustering

There is very little convincing evidence of clustering of this rather uncommon disease within communities. However, Gamble (1980a) reports a curious finding from the Paediatric Register of the British Diabetic Association. Where a diabetic child (onset age <16y) has an affected sibling, the interval between the dates of diagnosis of the two cases was less than one year in 42 pairs, a highly significant excess over the 18 expected. In 32 of these pairs the discordance interval was less than six months (nearly four times the expected number), and in three it was less than a week. Although the author acknowledged that an ascertainment bias was likely, it could not have been sufficient to explain these observations. Gamble concludes that some environmental event affecting both members of each of these pairs influenced the timing of onset of symptoms. Furthermore, the interval between that event and progression to clinical diabetes appears unlikely to be more than a few months. This is relevant to the discussion in Chapter 9 on the latent period of Type 1 diabetes.

4-B-iii. Secular trends and fluctuations

The 'frequency reduction' mentioned above was defined by Falconer et al (1971) as 'the reduction of frequency [of diabetes] with increasing current age [and therefore increasing duration] in any onset-age group'. Of various possible explanations they favoured two: differential mortality, and a secular increase in detection rate. They were unable to distinguish between these, considering that both were likely to be involved. So far as childhood-onset cases are concerned, if there were no excess mortality an increase in detection rate of 1-2% per annum would account for the frequency reduction observed.

North, Gorwitz & Sultz (1977) reported an apparent doubling of incidence of childhood diabetes in Erie County, New York over the period 1947-72. Up to 1961 the evidence was obtained retrospectively, but thereafter cases were recorded as they occurred. Changes in ascertainment seemed unlikely to explain the trend. A more rapid increase over a shorter period was deduced from prevalence figures from Michigan schools, but these are rather less reliable (see A-iv above). However,

North et al also cited data from Norway in 1900-1920, which suggested a much lower but rising incidence, the values appearing to lie on the same trend line as their own incidence rates. This somewhat tenuous evidence is compatible with the view that Type 1 diabetes is becoming more common during the present century.

Some other authors have reached a similar conclusion, but again rarely from convincing evidence. In retrospective, hospital-based studies (eg Patterson et al, 1981) there is always doubt about adequacy of old records, and as yet few of the more recent, careful incidence studies have been going long enough. Laporte et al (1981) found no consistent trend in incidence of 937 cases over 12y. Likewise, Palumbo et al (1976) reported no significant trend in incidence of juvenile cases (age <30) in Rochester, Minnesota, between 1945 and 1970 - which is perhaps not surprising as the 62 cases must have been rather thinly spread over the 25y. In contrast, in the Finnish incidence study the incidence in the years 1977-79 exceeded that in 1970-76 among children aged less than 10y. However, there was no obvious rising trend during the earlier period, and although the authors claimed the difference to be statistically significant, it is clear that the two periods they compared were selected a posteriori, after inspection of the results. This is not statistically valid. Nevertheless, the mean incidence of $29.10^{-5}.y^{-1}$ over the whole decade was higher than the $13.10^{-5}.y^{-1}$ reported for 1953 (Åkerblom et al 1980, 1981).

West (1978: p.297) has suggested that, rather than showing a consistent trend, incidence might fluctuate over periods of a few years, perhaps independently in different countries. Apparently in support of this, among a national cohort of 5362 children born in UK in a week of March, 1946, only one case of diabetes was identified by the age of 16 in 1962 (Wadsworth & Jarrett, 1974). In contrast not less than 22 cases of insulin-dependent diabetes occurred by the same age in another UK cohort of about 15,500 children born in 1958 ($p = 0.0198$ by our calculation using Fisher's exact test) (Calnan & Peckham, 1977). During this latter period (1958-74) a further 15 cases were diagnosed in the older cohort, 7 being insulin-dependent. This might suggest that incidence of Type 1

diabetes in UK was lower in the period 1946-62 than between 1958 and 1974. However, the single case identified by age 16 in the earlier cohort is significantly less than the 9.6 expected from our Figure III-3 above ($p<0.01$ by either χ^2 or binomial test), even though the absolute level of the curves in Figure III-3 is based on prevalence in schoolchildren in 1964, soon after the end of the same period, in a UK town. This discrepancy is difficult to explain, but calls into question the witness of the earlier cohort. Presumably it was a freak sample.

There is other evidence to suggest periodic fluctuations in incidence. Laporte et al (1981) noted quite marked differences between successive years in Pittsburgh, USA. Graphs of annual incidence figures in individual areas of Finland, 1970-79 (Åkerblom et al, 1981) show a striking pattern, with two maxima approximately 6y apart, separated by a clear minimum; the east appears to lead the north by a year or so, thus obscuring periodicity in the graph for the whole country.

Sultz et al (1975) reported that the annual incidence of diabetes in boys varied widely between 3 and 11.10^{-5} population from 1946 to 1971 (the Erie County data again), and showed a correlation with the incidence of mumps in the community four years before. Whether or not this implies a causative role for mumps virus is debatable, but there is no doubt that if such wide fluctuations are commonplace, then long-term trends will be very difficult to identify. Similarly, Laporte (1981) noted quite marked differences between successive years in Pittsburgh.

4-B-iv. Climate, race and socioeconomic factors as bases for geographical variation.

On the global scale there are marked differences in incidence of this disease between different areas. True Type 1 diabetes is distinctly rare in the Far East, in black Africans, American Indians, Eskimoes and Melanesians (West, 1978: 290-6); in general it seems to be predominantly a disease of white Caucasoids. How far this depends on racial differences in genetic constitution rather than on environmental factors is uncertain. Cohen (1970) reported that in Israel childhood diabetes

was nearly three times more common in those whose fathers had been born in Israel or Europe than in those with Asian or African fathers. Although this supports a genetically-determined basis for geographical variation, there are still substantial cultural differences between these groups in Israel, so environmental factors may yet prove to be important.

On a somewhat smaller scale there is the south-to-north incidence gradient within Europe, where quite marked differences appear over distances of several hundred to a few thousand kilometres (see A-iii above). This, too, could reflect the genetic clines which are known to occur within Europe (Menozzi, Piazza & Cavalli-Sforza, 1978), but it is often difficult to distinguish epidemiologically between genetic and environmental effects. The difference between North Island and South Island in New Zealand cannot be explained by different populations of non-white peoples, but Crossley & Upsdell (1980) do not mention the proportions originating in northern as opposed to southern Europe. However, some environmental influence may be operative.

In Scandinavia, too, the incidence is greater in the north and east than in the south and west, although in the extreme north of both Norway and Finland it is lower than elsewhere in these countries (Bratlid et al, 1976; Koivisto et al, 1976; Sterky et al, 1978; Akerblom et al, 1981). Again these differences could involve genetic, climatological or microbiological factors. Akerblom et al (1981) emphasize the first possibility, the low incidence in the north perhaps being due to admixture of Lappish genes; Type 1 diabetes was present in only one of the 1288 Lapps (of all ages) tested by Kirjarinta & Eriksson (1976).*

Socioeconomic factors such as population density could be important. They might, for example, modify patterns of infection. Christau et al (1977) reported a twofold excess incidence of Type 1 diabetes in the north when compared with the south of Copenhagen, the north having more

*It is interesting (vide Chapter 6) that HLA-B8 is relatively rare in Lapps, as low a frequency as 1% being found in the most pure-bred Skolt Lapps, although B15 is very common (58% in Skolt Lapps) (Kirjarinta & Eriksson, 1976).

low-income families. West et al (1979) found the opposite in Montreal. In Iceland, Helgason & Jonasson (1981) noted a significant excess of Type 1 diabetes in the two cities compared with rural areas. In contrast, no such difference was reported in Denmark, the Netherlands or Finland by Christy et al (1979a), Vaandrager et al (1984) and Akerblom et al (1981), respectively, although the latter did suggest that the very low population density in the far north might be one factor in the low incidence there.

The evidence reviewed in this chapter suggests that the incidence of diabetes in a population is determined by a complex interaction of factors including genetic/racial determinants and environmental agents. The next two chapters deal with genetics, while the possible aetiological role of viruses is discussed in Chapter 8.

PART III — CHAPTER 5

GENETICS: EARLY WORK ON DIABETES

"[Diabetes] is apparently inherited quite irregularly; some even consider that it is not inherited according to the same rules as other characters, or alternatively that there are several distinct forms of the disease each of which is inherited in its own fashion, or perhaps in some cases not at all." (Hansen, 1925; author's translation)*

CONTENTS

A. Evidence for existence of genetic factors 52
B. Mendelian theories 53
C. Multifactorial inheritance 56
D. Limitations of early work. Heterogeneity 57

5-A. EVIDENCE FOR EXISTENCE OF GENETIC FACTORS

That diabetes may have an hereditary basis was recognized by the ancient Hindu physicians, Charaka and Susruta, but in Europe it was not until the 16th and 17th centuries AD that Rondoletius and Morton reported multiple cases within families. This suggestion of a familial tendency was supported in 1906 when Naunyn stated that 18% of his diabetic patients had affected relatives (Simpson, 1976). Pincus & White (1933) were the first of many to confirm a statistically significant increase in prevalence in parents and siblings of their patients.

---

*"vererbt sich in anscheinend ganz unregelmässiger Weise, so unregelmässig, dass viele sogar meinen, das er überhaupt nicht nach den selben Gesetzen wie andere Eigenschaften vererbt wird, oder dass es mehrere verschiedene Krankheits-formen gibt, die jede auf ihre Weise oder überhaupt nicht wererbt werden."

Then Bergh (1938) and others reported that monozygotic twins were more often concordant for diabetes than dizygotic twins, implying that the familial occurrence of the disease was due at least in part to genetic factors (Langenbeck & Jörgensen, 1976). This is now generally accepted to be true, but for Type 1 diabetes even monozygotic twins are concordant in less than 50% of cases, so environmental factors must also be involved (Tattersall & Pyke, 1972).

5-B. MENDELIAN THEORIES

After the delayed penetration of Mendelian genetic theory into the body of scientific knowledge, there appeared between 1900 and 1930 a number of conflicting reports purporting to show simple dominant, recessive or even sex-linked inheritance on the basis of individual diabetic pedigrees. It soon became apparent that this condition was more complex (Hansen, 1925, quoted at the head of this chapter).

Notwithstanding these doubts, single-gene (Mendelian) hypotheses held sway for many years, apparently supported by new quantitative evidence from pooled family data. Dominant inheritance was suggested by Levit & Pessikova (1934) because the prevalences in siblings and parents of their patients were equal, but they did not correct for age. Allan (1933) preferred a recessive hypothesis, and although he had insufficient data to give much support he appears to have initiated the technique of using a hypothesis to construct a model predicting incidence of diabetes in parents, siblings and children of patients, and then comparing the results with the frequencies actually observed.

This type of approach was used by Pincus & White (1933, 1934a) Like Allan, they considered simple dominant inheritance unlikely because "the disease may skip one or more generations", and "the percentage of diabetic parents is much too small", so they, too, used the recessive model. The main importance of this work is their recognition of the need to correct for inability to identify the diabetic genotype in many subjects in whom clinical diabetes had not yet appeared at the time of the study. They estimated genotype prevalence in parents and siblings of 521 patients by assuming that 100% of such potential diabetes would

become manifest by the 9th decade of life, and applying an age correction factor based on distribution of age at onset in other reported series. The frequencies of diabetes predicted in families resulting from different matings were close to what was observed, and it was concluded that the data were compatible with a Mendelian recessive determining "the potential for diabetes".

Some of the assumptions in these papers have been challenged. Thus it now seems unlikely that 100% of genetically susceptible individuals would become diabetic, even if they did live to age 90y. Further, the age-specific data used for the age corrections were derived from different populations, and although the authors acknowledged a relative age bias in their own series their attempt to correct for this was inappropriate (Neel, 1976). In fact they did not claim the recessive hypothesis to be proven, merely that it was the simplest explanation compatible with the evidence; an alternative hypothesis they suggested, "a dominant gene the expression of which is suppressed or not possible in a proportion of cases", is not far removed from the multifactorial genetic determinants and environmental factors now believed by many to underlie Type 2 diabetes.

Soon afterwards Pincus & White (1934b) attempted to detect the 'diabetic' genotype more directly, by measuring glucose tolerance in the relatives of diabetic patients. They showed that, in families with neither, one or both parents diabetic, the respective proportions of children (other than the probands) with 'abnormal' tolerance were in the ratio 1 : 2.6 : 3.7. By the recessive hypothesis the proportions of genotypically susceptible children to be expected would be in the ratio 1 : 2 : 4, and the proportions of phenotypically abnormal children should be in the same ratio even if penetrance is reduced. Thus once again the observations were compatible with recessive but not with dominant inheritance.

A similar approach was used by other proponents of recessive inheritance, including Steinberg & Wilder (1952), who obtained a ratio of 1 : 2.4 : 3.4. In addition, they found some evidence for slightly

increased consanguinity between the parents of diabetic patients, particularly those diagnosed after age 30y, thus supporting one of the findings of Harris (1950). A third test used by Steinberg & Wilder involved calculating the relative proportions to be expected for the three possible types of mating between parents heterozygous or homozygous for a recessive diabetes gene, and comparing the result with the proportion of probands in their sample who had both, one or neither parent diabetic. To do this they assumed random mating, equal fertility and equal ascertainment of the three mating types, and used a maximum-likelihood estimate of gene frequency. The predicted proportions were very close to those observed (1%, 19% and 80% respectively), which the authors took as good evidence in favour of recessive inheritance. In fact, as later acknowledged by Steinberg (1961, footnote), this method is a test not of recessivity but simply of random mating: if p is the proportion diabetic in this population of parents of diabetic probands, the random probabilities of a pair of parents including two, one or no diabetics are p^2, $2p(1-p)$ and $(1-p)^2$. From Steinberg's data $p = 0.105$, so the proportions expected are 1.1%, 18.8% and 80.1%.

In the summary of their paper, Steinberg & Wilder (1952) stated categorically that "only the hypothesis of a single autosomal recessive gene is consistent with the data...", giving a gene frequency of 20-25% and a 'penetrance' of about 20%. This view was strongly reasserted (eg Steinberg, 1961) and by the early 1960's was firmly established in the literature. In fact, of course, they had not really proved the recessive hypothesis. Close examination of the 1952 paper shows that they were aware of two possible problems which they considered unlikely to be important, but which later proved to be so. The first was the question of genetic heterogeneity of diabetes (see below). The second was that genetic determination of susceptibility might depend on several interacting genes. They did acknowledge the possibility of "genetic modifiers affecting the time of onset of the disease", but nowhere did they give serious consideration to multifactorial hypotheses.

5-C. MULTIFACTORIAL INHERITANCE

As early as 1925 Hansen proposed that diabetes was inherited as a multifactorial character. This could explain its irregular occurrence within families, as well as sporadic cases. According to his theory,

> "Apparently isolated cases arise when the latent predisposing factors from the two parents accumulate in the fertilized ovum..." "Moreover, further investigation commonly shows that such cases are after all familial." (Author's translation)*

He believed that the inherited predisposition resulted from abnormalities affecting a variable proportion of the numerous genes he considered likely to be necessary for normal pancreatic islet development; the greater the proportion abnormal, the more severe and the earlier the onset of the diabetes.

By such a theory it would not be a discrete disease entity which was inherited. Instead, the character determined by the interacting genes could be considered as a quasi-continuous variable, which may or may not be measurable. Presence or absence of disease would depend on whether this variable exceeded or failed to exceed an arbitrary threshold level.

In a paper entitled 'The simulation of Mendelism', Edwards (1960) analysed mathematically the properties of such an "arbitrarily dichotomized variable" on the assumption that it was normally distributed. He showed that for diseases with population prevalence in the range 0.1-1.0%, the prevalence in relatives of propositi would be very similar to that expected if they had been determined by incompletely penetrant recessive genes. He also pointed out that, if impaired or variable prevalence is due to the effects of modifying genes, then the two mechanisms become more or less equivalent in fact as well as in the familial patterns of disease produced.

---

*"Anscheinend isolierte Fälle entstehen, wenn sich die latenten Anlagen der Eltern im befruchteten Ei kumulieren..." "Nahere Untersuchung zeigt übrigens sehr haüfig, dass solche Fälle in der Tät familiare sind."

Edwards (1960) did however indicate some properties of multifactorial inheritance which might be useful in distinguishing it from single-gene mechanisms. First, only in the former case would the risk to siblings increase with the number of relatives already affected. Secondly, where the arbitrarily dichotomized variable could be measured (eg height, blood pressure, or some aspect of glucose tolerance), its frequency distribution in the whole population should be unimodal, unless there were self-reinforcement above a certain level (as, for example, in malignant hypertension). Thirdly, the levels of this variable in apparently undiseased carrier parents should be higher than normal. Penrose (1953) had already predicted a fourth theoretical distinction: if incidence in monozygotic twins of patients is more than four times that in dizygotic twins or other siblings, multifactorial inheritance is likely.

Evidence of this nature suggests that genetic susceptibility to 'diabetes' is multifactorial. Thus Keen & Track (1968) among others showed that blood glucose levels after a glucose load are unimodally distributed, tending to be higher than the mean in relatives of diabetics. Further support comes from the finding that concordance for clinical maturity-onset diabetes between identical twins is much higher than that between non-identical twins (Tattersall & Pyke, 1972). That this (and indeed the other evidence for multifactorial inheritance) may not apply to juvenile-onset cases points to one of the major limitations of the work described so far: genetic heterogeneity of diabetes.

5-D. LIMITATIONS OF EARLY WORK. GENETIC HETEROGENEITY OF DIABETES

Simpson (1962) has summarized the problems which hinder interpretation of genetic research in diabetes as follows:

> "... there ... are three main difficulties. Those individuals who are of the diabetic genotype but who show no abnormal tolerance of glucose are indistinguishable from those who are of the genotype for normal; the proportion of individuals of the diabetic genotype manifesting the disease increases with age; and it is not known if the predisposition for all diabetes is a result of the same genotype."

Pincus & White (1933, 1934a,b) and Steinberg & Wilder (1952) attempted to allow for the second of these problems (age), although only partly successfully. More important is the first difficulty, particularly if the 'penetrance' of the diabetic genotype is as low as proposed by Steinberg and Wilder, since the distinction between recessive and multifactorial hypotheses is then virtually impossible by quantitative genetics (Penrose, 1953; Edwards, 1960).

But it is the third of Simpson's problems, the recurring question of heterogeneity of diabetes, which creates a fundamental problem affecting all the evidence so far discussed. Some aspects of this topic have been described in Chapters 1 and 3 above. The concept of genetic heterogeneity was first mooted by Hansen (1925; see above), and also by Cammidge (1928,1934): of several hundred pedigrees, some suggested to him a 'dominant' and some a 'recessive' pattern, the onset of the latter tending to be earlier in life.

Harris (1950) cited other evidence of this type, and himself demonstrated in a much more sophisticated analysis that there was a significant correlation between age at onset of diabetes in probands and that in their diabetic siblings, even after allowing for spurious correlation due to age stratification effects in the sample. This was confirmed by Steinberg & Wilder (1952), but they were sceptical of Harris's deduction that there was genetic heterogeneity.

More recent work has established the existence of genetic heterogeneity beyond any doubt. The newer evidence includes a tendency for Type 1 and Type 2 diabetes to 'breed true to type' (see below), a markedly lower concordance rate for juvenile-onset than for maturity-onset diabetes in identical twins (Tattersall & Pyke, 1972), an excess of organ-specific autoimmunity in non-diabetic relatives of Type 1 diabetics alone (Fialkow et al, 1975; see also Chapter 7), and, most importantly, different associations of the two types of diabetes with genetic markers. Thus, there are HLA associations with Type 1 but not with Type 2 diabetes (see Chapter 6), whereas Type 2 diabetes is associated with chlorpropamide-alcohol flushing (Pyke, 1980); and the incidence ratio

between blood group A and group O is significantly raised for onset ages above 20y, but not for juvenile-onset cases (Mourant, Kopec & Domaniewska-Sobczak, 1978).

The tendency to breed true with respect to age at onset is manifested by Harris's (1950) observations(see above), augmented by those of Degnbol & Green (1978), by an increase of juvenile- onset diabetes in siblings of patients with juvenile-onset but not of those with maturity-onset diabetes (Simpson, 1962), by lack of excess maturity-onset cases among ancestors of diabetic children when compared with controls (Lestradet, Batiselli & Ledoux, 1972; MacDonald, 1974), and by a significant difference in relative age-corrected frequencies of maturity- and juvenile-onset diabetes between families of patients with the two types of disease (Köbberling, 1969).

Other statistical evidence has been used to support heterogeneity with respect to age: the 'K ratio' for diabetes is higher in young than in older subjects, and the genetic correlation between the disease in young people and that in their parents is low. However this appears to some to be rather less convincing (Köbberling, 1976).

Penrose's (1953) K ratio is the ratio between frequency of a disease in a class of relatives of a patient and that in relatives of control subjects. It has been shown to be higher for juvenile-onset than for maturity-onset diabetes (Working Party of the College of General Practitioners, 1965; Simpson, 1968). Although others have suggested that this difference could be due to the lower frequency in young than in older control subjects (Keen & Track, 1968; Köbberling, 1976), their argument seems to us partly fallacious, in that it depends on a constant (arithmetic) excess in frequency of diabetes among the relatives of patients above that in controls, rather than a geometric increase in risk.

Another approach is to assess heritability of liability. Liability is defined as a hypothetical, continuously distributed variable to which both environmental and genetic factors may contribute. Disease results

when liability exceeds a threshold level, and an increase in disease incidence may result from an increase in mean liability or an increase in variance. Heritability of liability is that proportion which is due to additive genetic variation; i.e. the genetic contribution is itself counted as a continuous variable, so this method assumes multifactorial inheritance. Even though for human diseases liability cannot in general be measured directly, if it is assumed to be normally distributed an estimate of heritability can be made from frequency of the disease in relatives of patients and in control subjects (Falconer, 1965).

Falconer (1967) estimated that genetic correlation for diabetes between juvenile-onset diabetics (aged <20) and their parents (mean age 48y) was less than 0.5, and in females <0.25. His assumption that there was no excess mortality in the diabetic parents may not be valid, but it is difficult to see how a bias from this source could have been so large as to explain the low figures obtained. Simpson (1969) later examined a larger sample which allowed more precision. She, too, found low genetic correlation (0.31) between early-onset (age <20y) and late-onset (40y or more) diabetes; the intermediate group (age 20-39y) gave genetic correlations of 0.74 and 0.67 with the other two, respectively. suggesting overlap between two fairly distinct genetic forms of the disease, with tendencies to early and late onset respectively.

This work on heritability of liability to diabetes has been criticized in general terms on the grounds that continuous multifactorial inheritance cannot necessarily be assumed for all variants of a heterogeneous condition (Köbberling, 1976). However, the conclusion that there is heterogeneity still stands.

Although most of the evidence so far discussed demonstrates heterogeneity with respect to age at onset of diabetes, it is now clear that the primary distinction is between Type 1 and Type 2 diabetes (see definitions in Chapter 1). Thus, in identical twins, the concordance rate for non insulin-dependent diabetes is >90%, much greater than for insulin-dependent diabetes, irrespective of age at onset (Barnett et al, 1981). Further evidence comes from the use of genetic markers. In an

early study of this type, Lister & Tanner (1955) examined somatotype, which is genetically determined (and distinct from adiposity). Acute-onset diabetics, who reported symptoms present for less than three months before diagnosis and can probably be considered as broadly equivalent to Type 1 diabetics, had a significantly higher ectomorphy rating than insidious (Type 2) cases, even though there was considerable overlap between the two groups in respect of age at onset. More convincing evidence was provided by Cudworth & Woodrow (1976) who reported that HLA associations with adult-onset diabetes were detectable only when it was insulin-dependent.

More recently genetic heterogeneity has been demonstrated in juvenile-onset diabetes, between insulin-dependent cases and 'MODY' (see chapter 1). Thus, the latter shows a much stronger family history, at least some pedigrees suggesting dominant inheritance (Tattersall, 1976). In these and some other families chlorpropamide-alcohol flushing (CPAF), a dominantly-inherited trait, appears to act as a genetic marker for Type 2 diabetes. Although there has been some controversy because of differences in method and interpretation of the test, it seems that the 20% or so of Type 2 diabetics who are CPAF positive tend to have a milder form of the disease and a stronger family history than the others, whatever the age at onset (Pyke & Leslie, 1978; Leslie & Pyke, 1978). No association has been reported between CPAF and Type 1 diabetes; and conversely, MODY families, unlike Type 1 diabetic families, do not show HLA linkage (Nelson & Pyke, 1976).

The appreciation that diabetes was genetically heterogeneous was one major landmark underlying the great advances in knowledge over the past decade. Equally important was the discovery of a useful genetic marker for Type 1 diabetes, as described in the next chapter.

PART III — CHAPTER 6

GENETICS (2)

SUSCEPTIBILITY TO TYPE 1 DIABETES

CONTENTS

6-A	The concept of susceptibility. Penetrance	62
B	The HLA system	64
	i General description	64
	ii Phenotypes and genotypes	67
	iii Linkage disequilibrium	68
	iv Segregation and zygotic assortment	70
	v Disease associations: some statistical points	72
C	HLA and Type 1 diabetes	73
	i Association: evidence from population studies	73
	ii Family studies: linkage	78
	iii Family studies: other aspects	80
D	Other genetic markers	82
E	Mode of inheritance of susceptibility	86
	i Genetic heterogeneity within Type 1 diabetes	86
	ii Single-gene hypotheses	88
	iii Other hypotheses	89
	iv A protective gene?	90
	v Epistasis between HLA-DR3 and DR4?	90
	vi Susceptibility genes not linked to HLA	91
F	Risk of Type 1 diabetes in siblings	92

6-A. THE CONCEPT OF SUSCEPTIBILITY. PENETRANCE.

Some of the earlier writers suggested that the inherited character was not diabetes itself but an unidentified 'condition predisposing to its development' (Cammidge, 1928), also termed 'potentiality' (Pincus & White, 1953), 'predisposition' and 'susceptibility' (Penrose, 1953), and 'liability', or rather its heritable component (5-C). In keeping with current preference we use 'susceptibility'.

Several questions arise. First, what proportion of genetically susceptible children actually develop diabetes; in other words, what is the penetrance? Secondly, what is the mode of inheritance? Thirdly, can susceptible but (as yet) unaffected individuals be identified; and finally, what is the actual nature of susceptibility, or what is (are) the character(s) determined by the susceptibility gene(s)?

Concerning the first question, it follows from the relationship between prevalence of Type 1 diabetes and age (see Chapter III-4, Fig 3) that penetrance must be low in young children and must rise with age, as Pincus & White (1933) pointed out for diabetes as a whole. However, their assumption of 100% prevalence by age 90y is almost certainly not valid for Type 1 diabetes. The strongest evidence against it comes from studies on identical twins, such as the UK study most recently reported by Barnett et al (1981).

In that series, among 80 pairs of monozygotic twins concordant for Type 1 diabetes the discordance interval was usually short (<5y in 70%, and >10y in only one pair), whereas all but three of the 67 initially discordant pairs remained discordant throughout observation periods which exceeded 10y in 40% of cases. The three initially unaffected twins who later became diabetic did so within thirteen months of their co-twins. Barnett et al concluded that most pairs discordant for more than a few years were likely to remain so, even though the unaffected members must be genetically susceptible.

These authors pointed out that the penetrance of the susceptibility genes was undoubtedly much lower than the sample concordance rate of 55%, for two reasons: ascertainment favoured concordant pairs; and in general identical twins share not only genes but also environment, especially when young, so even if the concordance rate were accurately known it would still overestimate the penetrance in the general population. Another consequence of the high discordance rate in identical twins is that non-genetic (and therefore, presumably, environmental) factors must also be important, as suggested by epidemiological evidence (Chapters III-4 & III-8).

Because penetrance is low, the second of the above questions, concerning mode of inheritance, is unlikely to be answered satisfactorily by conventional analyses of familial patterns of manifest Type 1 diabetes. What is needed is some other way of detecting susceptibility. Thus it was an important advance when a genetic marker for the disease was discovered less than a decade ago.

6-B. THE HLA SYSTEM

6-B-i. General description.

So far as is known, histocompatibility antigens are present on the surface of nucleated cells in all but the most lowly of animals. In their principal biological role they are believed to facilitate T-lymphocyte responses to most antigens, but they were first detected by virtue of rapid rejection of tissue grafts (Munro & Waldmann, 1978), and were shown to be genetically determined at a complex of closely linked loci known as the Major Histocompatibility Complex (MHC).

The MHC of man is known as the HLA system (a term derived from the older nomenclature 'HL-A', for 'Human Leukocyte system A'). Early family studies showed that HLA phenotypes formed two segregant series; that is, they behaved as if determined by sets of allelomorphic genes at two loci, which were later found to be closely linked (Ceppellini et al, 1967).

These two series and their loci, at first called 'LA' and '4', were later re-named 'A' and 'B'. A third series, HLA-C, was defined in 1975. HLA-A, B and C antigens are glycopeptides (Mr 43,000) and are present on most nucleated cells, where they are non-covalently bound to ß-2 microglobulin (Mr 12,000). In the laboratory they are identified on peripheral blood lymphocytes in a complement-mediated micro-lymphocytotoxicity test using sera obtained from multiparous women or, recently, monoclonal antibodies (Bodmer, 1978, 1980).

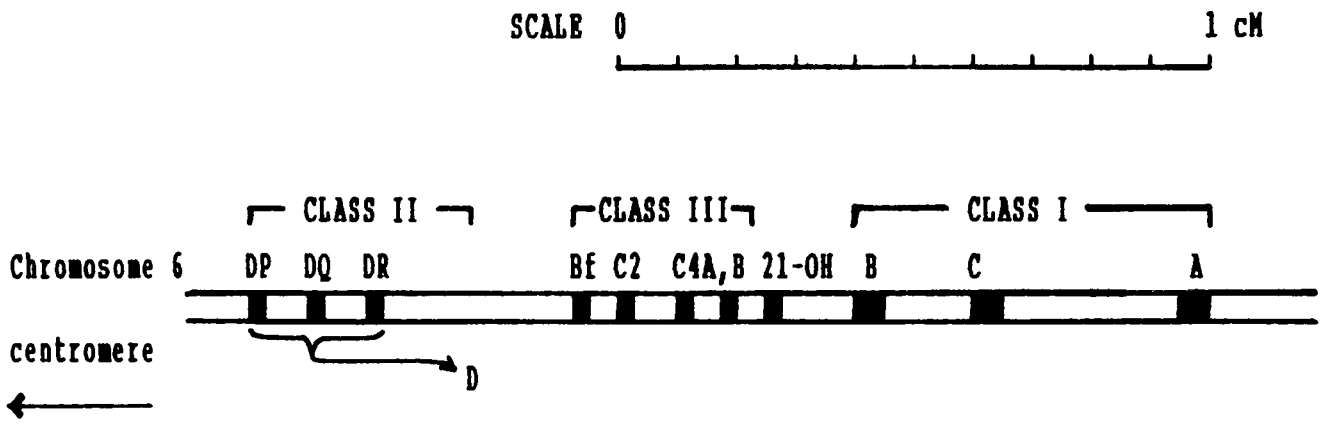
HLA-D specificities were first characterized in the mid-1970's by mixed lymphocyte culture (MLC). The genetic structure of the Class II region is complex (Figure III-4). By 1979, DR (D-related) antigens were

being typed serologically on B-lymphocytes. They also occur on monocytes, activated T-lymphocytes, and on some other cells under certain conditions (Bodmer, 1980). Specialized typing methods indicated two further polymorphic loci nearby, named DP and DQ in 1984 (formerly SB and DC respectively). The HLA-D specificities typed by MLC probably reflect mainly the DR genes, with lesser contributions from DP, DQ and perhaps other Class II loci (Bodmer & Bodmer, 1984; Bach 1985).

The HLA loci sparsely occupy a region of about 2000kilobases on the short arm of Chromosome 6, representing about one-thousandth of the total human genome. Other loci in this region include those for three complement components (Factor B (Bf), C2 and C4, the latter having two loci), two red-cell blood groups (Chido and Rodgers, believed to be the same as C4B and C4A) (Figure III-4), and steroid 21-hydroxylase. Nearby are loci for C3d complement receptors on lymphocytes, effector cell stimulating genes (ECS), and more distantly (to the left, towards the centromere) phosphoglyceromutase 3 (PGM_3) and glyoxalase (GLO) (Festenstein & Demant, 1978).

The HLA system is extremely polymorphic. For most of the HLA loci proper at least 8 serological specificities have been defined (Figure III-4). Many HLA antigens have proved to be 'supertypic', consisting of two or more specificities which cross-react serologically, as indicated in the figure. It does not necessarily follow that each of these cross-reacting components or 'splits' is determined by a single, distinct DNA sequence. For example, the molecular configuration of the antigenic sites could be influenced by the structure of the backbone of the HLA molecule, which may itself be determined by different genes in the HLA complex, as in the case of the DR-α chain. This may explain the cross-reactivity of some of the earlier HLA-B supertypes with both Bw4 and Bw6 (two of the first specificities defined in the 'Four' system). Furthermore, different antigenic sites on the same gene-product may react with different sera; this does appear to be the case for the cross-reactions of B7 and B27 with Bw6 and Bw4 respectively (Joysey & Wolf, 1978).

FIGURE III-4. THE HLA SYSTEM



DR	D	Bf	C2	B	C	A
1	w1	F	A	5*a	w1	1
2	w2	F1	B	7	w2	2
3	w3	S	C	8	w3	3
4	w4	S1		12*b	w4	9*i
5*p	w5			13	w5	10*j
w6*q	w6*n			14*m	w6	11
7	w7*o			15*c	w7	w19*k
w8	w8			16*d	w8	23(i)
w9	w9			17*e		24(i)
w10	w10			18		25(j)
w11(p)	w11(o)			21*f		26(j)
w12(p)	w12			w22*g		28*l
w13(q)	w13			27		29(k)
w14(q)	w14			35		30(k)
	w15			37		31(k)
w52	w16			38(d)		32(k)
w53	w17(o)			39(d)		w33(k)
	w18(n)			40*h		w34
	w19(n)			w41		w36
				w42		w43
				w44(b)		w66(j)
				w45(b)		w68(l)
				w46		w69(l)
				w47		
				49(f)		
				w50(f)		
				51(a)		
				w52(a)		
				w53		
				w54(g)		
				w55(g)		
				w56(g)		
				w57(e)		
				w58(e)		
				w59		
				w60(h)		
				w61(h)		
				w62(c)		
				w63(c)		
				w64(m)		
				w65(m)		
				w67		
				w70*n		
				w71(n)		
				w72(n)		
				w73		

The map of the HLA region is drawn approximately to scale; 1cM (centimorgan) represents a 1% recombination fraction, or approximately 1000 kilobase (Bodmer, 1978). A, B, C, DR, DP and DQ are HLA loci, C2 and C4 are the second and fourth components of complement, Bf is Factor B of the alternative complement pathway, and 21-OH is the double locus for steroid 21-hydroxylase. D is a series of specificities characterized by mixed lymphocyte culture, and probably determined genetically by the combination of Class II alleles present. Below the map are the HLA specificities recognized by the Ninth International Histocompatibility Workshop (Bodmer & Bodmer, 1984). The prefix w indicates that the specificity is provisionally recognized but has not been assigned a permanent number. At the C locus, however, w is always used, to distinguish these specificities from C2 and C4. The specificities marked *a, *b ... are supertypic or 'broad', their 'splits' being marked (a), (b) respectively.

Nevertheless, whether gene or supergene, the genetic determinant of each HLA specificity behaves during inheritance as an allele at the appropriate locus, and is therefore termed an allele here as in other works (Festenstein & Demant, 1978).

6-B-ii. HLA phenotypes and genotypes

Where only one HLA specificity is detected for a particular locus, this may be due to homozygosity of the corresponding allele, or to the presence of another allele whose gene product is not detectable because

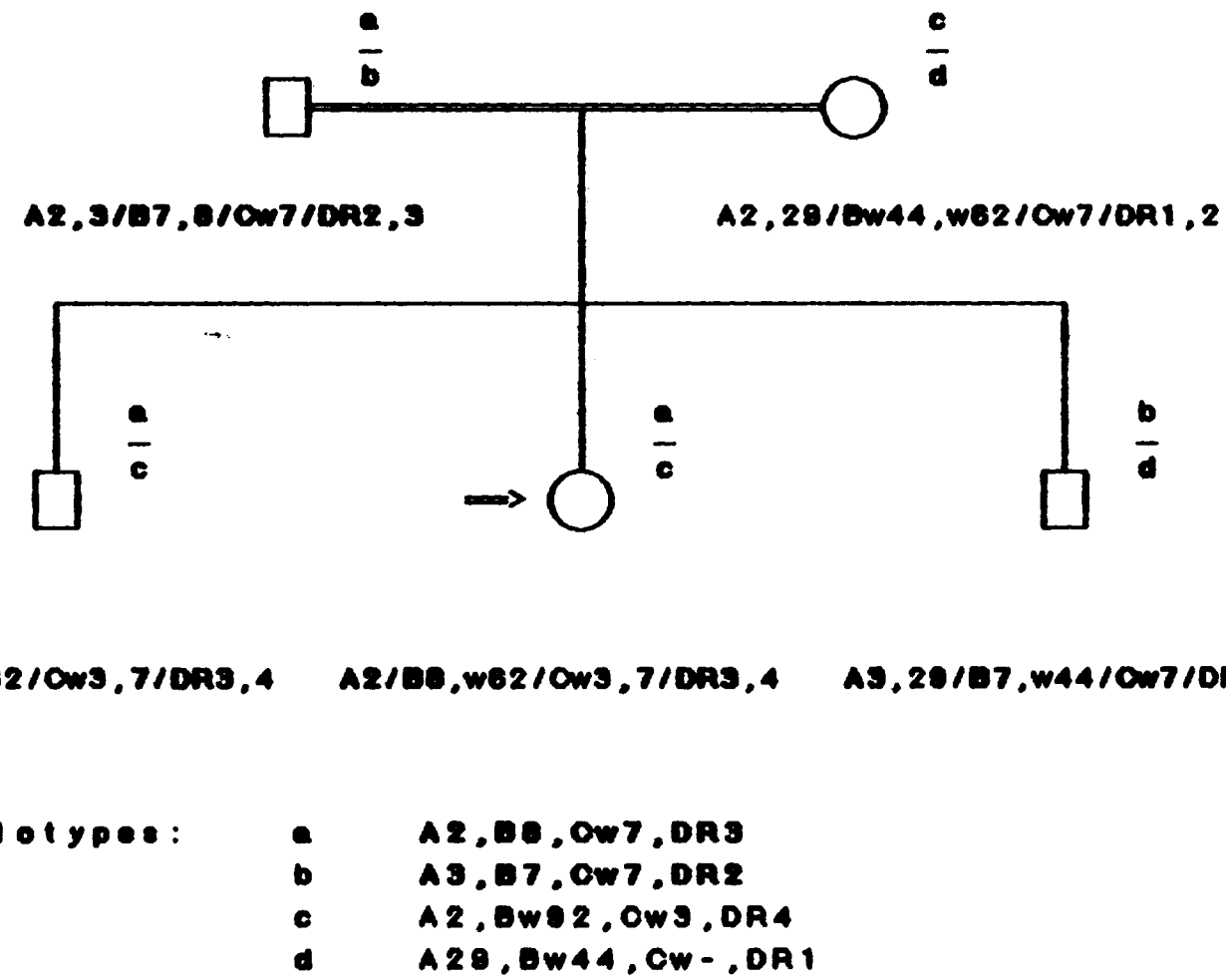


FIGURE III-5. DETERMINATION OF HLA HAPLOTYPES. In this example the proband (arrowed) is to be genotyped. First the HLA phenotypes of the proband and first degree relatives are determined (shown in full). Since each antigen behaves as if determined by a single allele, those in bold type in the children must have been inherited from the father, and the remainder from the mother. Thus the four parental haplotypes can be worked out, and hence the genotypes of all family members. ('Haplotype', from 'haploid genotype', is the particular set of alleles on one chromosome in a given individual; two haplotypes make up the genotype.) In this example the father is homozygous for Cw7 and the proband for A2. Although only one Cw antigen is detected in the mother, this is due to a 'blank' rather than to homozygosity, since there is no Cw3 in the third child. There is no evidence of recombination.

the appropriate specificity is not represented in the battery of sera used - a serological 'blank'. These two possibilities can be distinguished only by testing the first degree relatives of the individual concerned and thereby determining his HLA genotype (Figure III-5). A third, theoretical, possibility, that the gene product is not expressed at all, cannot be distinguished from a serological blank.

The gene frequency of a particular allele is its frequency among the population of haplotypes (defined in Figure III-5), rather than among the population of diploid individuals. It can be determined accurately only in an HLA-genotyped population.

However, gene frequency p_a can be calculated indirectly from phenotype frequency f_a on the assumption of Hardy-Weinberg genetic equilibrium (an assumption which is not necessarily valid: 6-B-iv). By the Hardy-Weinberg rule,

$$p_a^2 + 2p_a.p_{a-} + p_{a-}^2 = 1 \qquad (6\text{-}1)$$

where p_{a-} is the frequency of 'allele not-a', i.e. the sum of the gene frequencies of all other alleles (including blanks). Now the phenotype frequency is given by the sum of the first two terms in Equation 6-1, representing the population frequencies of homozygotes and heterozygotes respectively. Therefore,

$$p = 1 - \sqrt{1 - f_a} \qquad (6\text{-}2)$$

6-B-iii. Linkage disequilibrium

It is well established that within the HLA system certain alleles at different loci tend to occur together in the same haplotype more often than expected on the basis of random assortment (that is, assortment by chromosomal recombination over many generations) This phenomenon of linkage disequilibrium, or gametic association, is quantified by the standard parameter:

$$\Delta = p_{a,b} - p_a.p_b \qquad (6\text{-}3)$$

where p_a and p are the gene frequencies of alleles a at one locus and b at another, and $p_{a,b}$ is the haplotype frequency (H.F. in Table 5), expressed as a proportion of the total number of haplotypes in the

population concerned (Bodmer et al, 1966).

When genotypes are not available Δ cannot be measured directly because $p_{a,b}$ is not known. However, if Hardy-Weinberg equilibrium is assumed, the following expression is equivalent (derived from Ceppellini et al, 1967):

$$D_{a,b} = \sqrt{1 + f_{a,b} - f_a - f_b} - \sqrt{(1 - f_a)(1 - f_{ab})} \qquad (6\text{-}4)$$

or

$$D_{a,b} = \sqrt{f_{\langle a,b\rangle -}} - \sqrt{f_{a-} \cdot f_{b-}} \qquad (6\text{-}5)$$

where f_a, f_b, and $f_{a,b}$ are the phenotype frequencies of specificities a, b and the combination a + b respectively, while f_{a-}, f_{b-} and $f_{\langle a,b\rangle -}$ are the frequencies of individuals negative for these specifities.

When Δ values between different pairs of alleles are to be compared this cannot be done directly because they depend not only on the strength of the gametic association but also on the gene frequencies and haplotype frequencies themselves. Thus the 'relative Δ value' is used. Bodmer et al (1966) defined this as follows:

$$r\Delta_{a,b} = \frac{\Delta_{a,b}}{p_{a,b}} \qquad (6\text{-}6)$$

More recently, $r\Delta$ has been defined in terms of the maximum value possible, given the gene frequencies of the two alleles concerned (Baur & Danilovs 1980):

$$r\Delta_{a,b} = \frac{\Delta_{a,b}}{\max\Delta_{a,b}} \qquad (6\text{-}7)$$

where $\max\Delta_{a,b} = p_a(1-p_b)$

and p_a is the smaller of the two frequencies concerned.

The statistical significance of Δ is generally assessed using χ^2. For linkage disequilibrium between alleles at three loci the following expression applies (Yasuda et al, 1978)

$$\Delta_{a,b,c} = p_{a,b,c} - p_a.p_b.p_c - (p_a.\Delta_{b,c} + p_b.\Delta_{a,c} + p_c.\Delta_{a,b}) \qquad (6\text{-}8)$$

Table III-6 shows some significant second-order linkage disequilibria.

TABLE 6. LINKAGE DISEQUILIBRIA IN THE HLA SYSTEM Significant second-order linkage disequilibria between B-locus antigens and those of other loci in healthy Caucasoids. Delta is the linkage disequilibrium parameter. H.F. is the haplotype frequency (see text). * = $p < 0.01$; otherwise $0.01 < p < 0.05$. From Bodmer & Bodmer (1978)

Between A and B loci				Between C and B loci				Between B and DR loci			
Haplotp.		Delta	H.F.	Haplotp.		Delta	H.F.	Haplotp.		Delta	H.F.
A	B	(x10³)	(x10³)	Cw	B	(x10³)	(x10³)	B	DR	(x10³)	(x10³)
1	5	13.7	20.6	1	5	9.3	12.1	7	1	16.8	22.9
1	8	57.2*	64.1	1	w22	11.4*	12.2	7	2	37.6*	46.2
1	17	16.0	22.4	1	27	14.6*	15.2	8	3	62.3*	70.1
2	12	27.2	64.5	2	27	22.6*	22.7	12	7	26.7*	41.3
3	7	18.5	28.3	2	40	13.6*	14.7	13	7	16.8*	18.5
w23	12	17.6*	19.3	3	15	28.3*	29.9	14	1	12.5*	15.3
26	w38	5.3	5.5	3	w22	18.9*	20.7	15	w6	13.0	17.7
29	12	27.3*	33.1	3	40	24.5*	26.4	17	7	22.8*	29.3
w30	18	16.6*	17.0	4	w35	84.2*	89.7	18	3	12.9*	18.2
w33	14	6.4*	6.6	5	12	40.2*	48.2	w35	1	15.1*	19.2
				5	18	19.4*	23.6	w38	4	8.6	11.7
				6	13	22.8*	25.0	w39	2	11.4*	14.2
				6	17	35.4*	38.7	40	4	14.5	22.8
				6	37	6.1	6.6	w47	4	5.8*	6.1

6-B-iv. Segregation and zygotic assortment.

> *'The existence of a polymorphism is a sure sign of evolutionary change if it be transient, and of the action of contending selective forces if it be the normal balanced form'* (Ford, 1965, p.88)

The HLA system is generally considered as a balanced rather than a transient polymorphism, and there is some evidence that resistance to

infective organisms may be an important factor in maintaining it (Bodmer & Bodmer, 1978). An important feature of genetic polymorphism is that selection often appears to favour the heterozygote. Thus Hardy-Weinberg equilibrium between HLA alleles cannot confidently be assumed to exist.

It is also possible that the 'selective forces' might act before birth, either at the zygotic or at the prezygotic (gametic) level. Clarke et al (1968) have delineated three possible mechanisms of prezygotic selection: meiotic drive, that is, unequal rates of gametogenesis with respect to the polymorphism concerned; sperm competition, the efficiency of the spermatozoa being affected by their genetic consitution; and sperm incompatibility. Clarke et al cite evidence that sperm incompatibility probably is a selective factor in the ABO polymorphism, involving secretion of anti-A and anti-B into the maternal genital tract. Spermatozoa express HLA-A,B,C and DR antigens, and have been used as stimulator cells for HLA-D typing in place of the more usual homozygous typing cells in the MLC reaction (Festenstein & Demant, 1978). Maternal anti-HLA antibodies could in theory lead to prezygotic selection in individual matings.

Zygotic selection is also well recognized, as for example in ABO incompatibility. Maternal immunization by fetal antigens occurs in many pregnancies, and indeed has been the principal source of sera for typing. If maternal reactivity to the set of HLA specificities encoded on one paternal haplotype were different from that to the other set, then differential zygotic elimination could result.

If zygotic or prezygotic selection is important in maintaining the HLA polymorphism then in some families there should be non-random distribution of HLA haplotypes among the children. Immunological mechanisms would be expected to affect second and subsequent rather than first pregnancies. Mattiuz et al (1971) investigated distribution of the four parental HLA haplotypes among the offspring in 93 families and found no evidence of non-random zygotic assortment, nor any deviation from the expected random distribution of members of each sibship between the four possible genotypes ac, ad, bc, bd (see Figure III-5). However, they did

not consider these results entirely to exclude such selective mechanisms, and were:

> *'aware of the possibility of more subtle effects which will require larger bodies of data and, perhaps, different methodological approaches, such as a study of the correlation between the HLA types of husband and wife and their fertility, as measured by the number of their surviving offspring'.*

Furthermore, they had insufficient data to investigate transmission of specific haplotypes. The present study suggests non-random segregation of HLA-B8 (RESULTS: VIII-1-B-iv; Cudworth et al 1979).

6-B-v. Disease associations: some statistical points

Certain diseases are associated with blood groups or other polymorphically determined phenotypes (Clarke et al, 1968). The strength of such an association may be expressed as the odds ratio. This is calculated as the cross-product ratio from a 2x2 contingency table constructed to show the numbers positive and negative for a particular phenotype in samples from unrelated subjects in the 'diseased' (h positive and k negative) and the control (H,K) populations. Now commonly known as the (approximate) relative risk (RR), it was first used by Woolf (1955). Haldane (1955) corrected for bias and discontinuity, and it is a modification of his formula which is used today:

$$X\ (=RR) = \frac{(h + 0.5)(K + 0.5)}{(H + 0.5)(k + 0.5)} \qquad (6\text{-}9)$$

This is a useful statistic in that it can be used to combine data from different populations even though the phenotype frequencies might be different. A further advantage is that it has a readily-appreciated 'direct medical meaning' in that, except for high frequencies, RR is a maximum-likelihood estimate of the ratio between the risk of disease in those with to that in those without the phenotypic specificity concerned (Woolf, 1955). In addition, $\log_e X$ is normally distributed with variance

$$V = \frac{1}{h + 0.5} + \frac{1}{k + 0.5} + \frac{1}{H + 0.5} + \frac{1}{K + 0.5} \qquad (6\text{-}10)$$

and thus the 95% confidence limits for RR are readily calculated.

The Woolf-Haldane method gives χ^2 for heterogeneity between

different populations and for significance of the combined results, but for data from a single population the standard χ^2 test (with Yates' correction) or an exact probability test is preferred.

Another measure of association used in this context is population attributable risk (PAR) or aetiological fraction, δ (Bengtsson & Thompson, 1981). This can be considered an estimate of the contribution to the total burden of the disease D in the population made by the excess in subjects positive for the HLA (or other) specificity concerned. It is calculated as:

$$\delta = \frac{f_D - f_C}{1 - f_C}, \qquad (6\text{-}11)$$

where f_D and f_C are the proportions positive for the specificity in the disease and control populations respectively.

An important point is that apparently significant associations are likely to occur by chance when many comparisons are made. For this reason, when new associations are sought the p values should normally be multiplied by the number of comparisons before any is deemed significant (see also under Heterogeneity, 10-B-iii-a).

6-C. HLA AND TYPE 1 DIABETES

6-i. Association: evidence from population studies

The first reported investigation of HLA phenotype frequencies in Type 1 diabetes failed to show any significant association, but this was in a small sample of patients with mixed racial origins including only 28 white Caucasoids (Finkelstein, Zeller & Walford, 1972). The following year Singal & Blajchman (1973), studying 50 Canadian patients, found a significant excess of the antigen later termed HLA-B15*. After some initial scepticism because of serological problems, Nerup et al (1974)

*Originally HL-Aw15, this specificity has now been 'split', the major component being Bw62 (see Figure III-4); in this discussion we use 'B15' as most of the work was done before the split was characterized.

and Cudworth & Woodrow (1974) confirmed the association with B15 and also showed one with B8. These first three reports agreed, however, that there was no evidence to suggest any association with Type 2 diabetes.

Over the next two years these findings in white Caucasoid Type I diabetics were confirmed, and further associations came to light: the frequencies of HLA-B18, A1 and A2 were significantly increased in affected patients, while those of A11 and B7 (and perhaps A3) were decreased (Nelson et al, 1975; Cathelineau et al, 1976; Mirouze et al, 1976; Cudworth & Woodrow, 1976; Ludwig, Schernthaner & Mayr, 1976).

In 1976 came the first of several reports based on the international HLA and Disease Registry (Ryder & Svejgaard, 1976; Nerup et al, 1977a; Christy et al, 1979b; Ryder, Andersen & Svejgaard, 1979) which summarized and combined data from various centres. These reports confirmed the associations already mentioned, while HLA-Cw3 was found to be increased and Cw4 to be decreased in frequency in Type 1 diabetes. There was some heterogeneity between different populations; for example, the association with HLA-B15 was more prominent in northern than in south-western countries of Europe, and conversely for B18. HLA-B40 (now split, Bw60 being the major component) also appeared to be increased in some populations (Cudworth & Festenstein, 1978). Statistically, these associations were in the main very highly significant. Nevertheless their strengths were low, in that all the RR values were below 3 or, if greater, had wide confidence limits.

Somewhat higher RR's were found when HLA-D and DR specificities were investigated. The first significant findings were from Copenhagen for MLC types (Thomsen et al, 1975; Nerup et al 1977b) and from Schernthaner et al (1977) and de Moerloose and co-workers (see de Moerloose et al, 1978) for DR antigens. In the HLA and Disease Registry in September 1978 there were data on around 200 patients giving significant RR's of 2.19 and 5.69 for Dw3 and DR3 respectively, and 4.01 and 2.84 for Dw4 and DR4, the last being based on only 76 patients (Christy et al, 1979b). HLA-Dw2, which occurs in about 25% of normal Caucasian populations, was not

found in any Type 1 diabetic patient until publication of our own work (Sachs et al, 1980).

More recently the Eighth International Histocompatibility Workshop combined HLA-DR data from several Caucasoid populations and found relative risks of 3.3 and 6.4 for DR3 and DR4 respectively, but there was significant heterogeneity between centres. The only single centre reporting more than 50 cases was in West Germany, where 74 patients gave respective values of 5.1 and 13.6 (Svejgaard, Platz & Ryder, 1980). HLA-DR2 was assigned a RR of 0.25 at this Workshop.

There is, then, a plethora of positive and negative associations between Type 1 diabetes and different HLA phenotypes, none of them very strong. The overall picture may be simplified by taking into account the known linkage disequilibria within the HLA system (Table III-6). Cudworth & Woodrow (1976) showed that, in subjects heterozygous for HLA-B8, HLA-A1 was no more frequent in patients than in controls; and likewise the association of B8 with Type 1 diabetes appears in its turn to be secondary to that of HLA-D/DR3 (Barbosa et al, 1979a; Svejgaard et al, 1980). The system of alleles in linkage disequilibrium with DR3 - including B18 - shows progressively increasing RR values as the HLA region of the chromosome is traversed from right to left (i.e. towards the centromere), and similarly for those in linkage disequilibrium with DR4 and DR2 (Figure III-6). The decrease in frequency of HLA-A11 has been attributed to serological cross-reactivity with A3, the latter being in linkage disequilibrium with B7 and D/DR2 (Cudworth & Festenstein, 1978).

An obvious implication of these complex HLA associations with Type 1 diabetes was apparent from the outset (Nerup et al, 1974): genetic susceptibility to this disease is probably determined by genes near the DR locus. It was not generally considered likely that the HLA genes per se are responsible because the RR's are so low (much lower than, for instance, that of >100 for HLA-B27 and ankylosing spondylitis; Bodmer, 1980), although this view is now challenged (6-E-v). Cudworth (1978) talked of 'susceptibility axes' (S1 and S2 in Figure III-6), each

representing HLA alleles in linkage disequilibrium with one or more susceptibility genes, as well as an 'R axis' associated with a putative gene suggested by Ludwig et al (1976) to confer some degree of resistance to the disease.

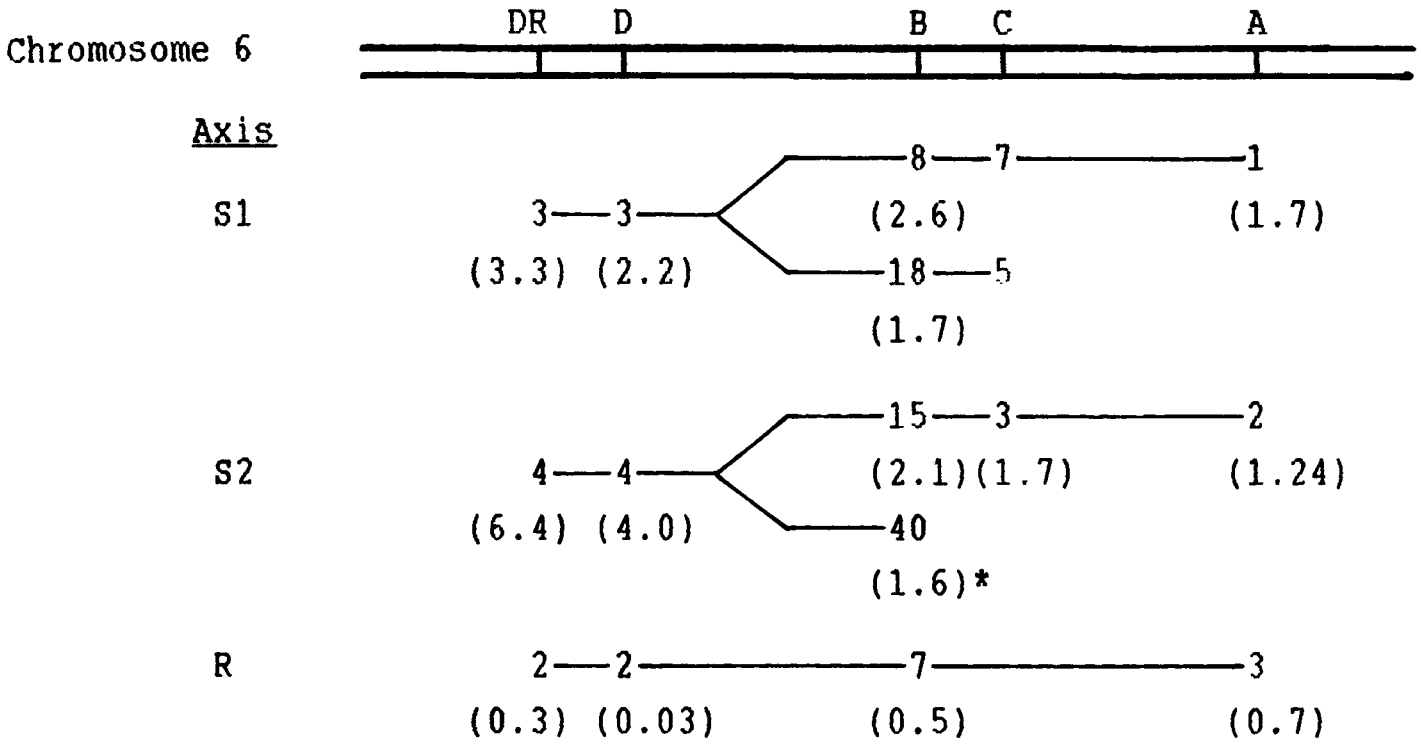


FIGURE III-6. HLA PHENOTYPE ASSOCIATIONS WITH TYPE 1 DIABETES

S1, S2 and R are the systems of HLA specificities in linkage disequilibrium with DR3, DR4 and DR2 respectively; the relative risk for the association of each with Type 1 diabetes is given in parenthesis (RR), illustrating the increasing deviation from 1.0 as the region is traversed from right to left (see text). The values used are those for white Caucasoids, from the Eighth International Histocompatibility Work-shop (Svejgaard et al, 1980) for DR; from Christy et al (1979b) for D, B (except B40) and C; and from Cudworth & Woodrow (1976) for A. *B40 from Ludvigsson et al (1977).

An important question here is whether a single susceptibility gene occurs in linkage disequilibrium with both the S1 and S2 axes, or whether on the other hand there are two different genes. This is discussed below (6-E-iii), but population studies provide the following evidence. A combination at the same locus of HLA specificities from both S1 and S2 axes carries a much higher relative risk than does either axis without the other. Thus Thomsen et al (1975) and Cudworth & Woodrow (1976) reported that the RR for the B8,15 phenotype approximately equalled the sum of the RR's for these antigens independently. Confidence limits were

wide in these early reports, but Barta & Simon (1977) in Hungary and Svejgaard & Ryder (1977) from the HLA & Disease Registry confirmed a significant difference. The Registry gave an RR of 9.8 (95% confidence limits 6.0-16.0) for the combination. There may also be enhancement of RR when B8 and B40 coexist (Cudworth & Festenstein, 1978). Farid et al (1978) found in Newfoundland a high RR for B8,18 as well as B8,15.

At the D/DR end of the HLA region, DR3,4 heterozygotes show an even more marked enhancement of RR, which exceeds 40 in some series (Platz et al, 1981; Deschamps et al, 1980). The figure for combined data from Eighth Workshop was 33 (confidence limits 21-51) compared with 3.3 (2.2-5.1) and 6.3 (4.3-9.2) for DR3,X and DR4,X respectively, where X is a DR antigen other than 3, 4 or blank (Svejgaard et al, 1980)

These findings suggest either that two (or more) different susceptibility genes are involved, one being in linkage disequilibrium with DR3 and the other with DR4, or that there is a single gene whose penetrance depends on gene dosage, i.e. subjects heterozygous for the gene are susceptible, but less so than homozygotes (intermediate inheritance). In the latter case homozygotes for 'high risk' HLA genes at a particular locus (e.g. DR3,3) should show relative risks for Type 1 diabetes just as high as heterozygotes having one HLA gene from each of the two S axes (e.g. DR3,4). Thus investigating RR's in these two situations is important.

Phenotype studies make only a limited contribution here. For the HLA-A,B,C loci, now that serological blanks are rare in Caucasoids most subjects typed as 'HLA-B8,-', for example, will in fact be B8 homozygous, but even so there is an element of uncertainty. Furthermore, in view of the low RR's for B-locus associations large numbers of subjects are needed. Thus data on this point are inconclusive (Cudworth & Festenstein, 1978). At the DR locus RR's are higher but blanks are more common, and it cannot be assumed that DR3,- and DR4,- individuals are homozygous. Nevertheless it appeared likely in the Eighth Workshop that DR3,4 heterozygotes do carry a higher risk of Type 1 diabetes than the corresponding homozygotes, although without statistical confirmation

(Svejgaard et al, 1980). Whether the homozygotes in turn have higher RR's than DR3,X and DR4,X is not yet clear.

Definitive answers to these questions will depend on large-scale studies of HLA genotypes, so that blank specificities can be distinguished from homozygosity in large numbers of patients. This forms one aspect of HLA family studies.

6-C-ii. Family studies: HLA linkage

If genes in the HLA region of Chromosome 6 do indeed determine susceptibility to Type 1 diabetes, then by HLA genotyping families with two or more diabetic children (Figure III-7) it should be possible to demonstrate linkage between HLA and the disease.

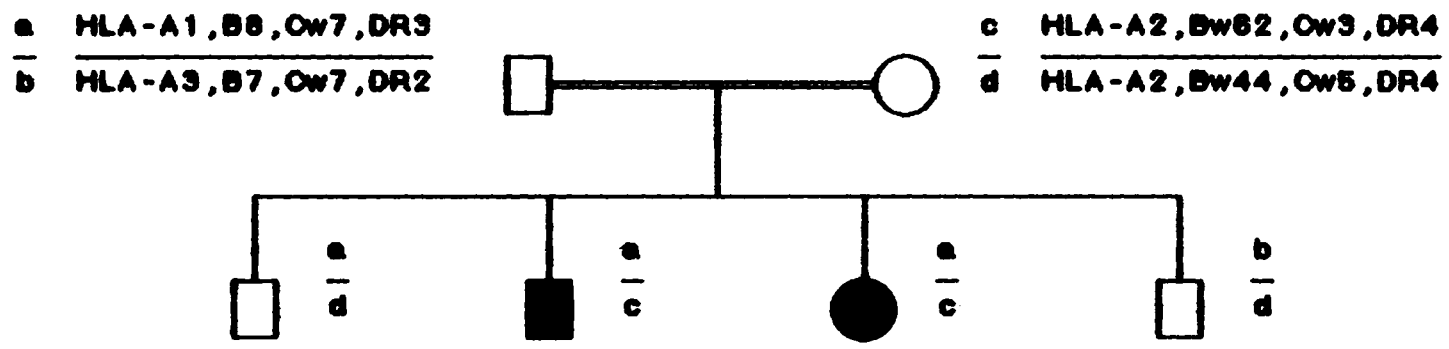


FIGURE III-7. HLA GENOTYPES IN A DIABETIC MULTIPLEX FAMILY. A hypothetical family with two diabetic children (solid symbols) who are HLA identical (see text).

Assuming random segregation of HLA haplotypes (see above), by the null hypothesis of no linkage 25% of pairs of diabetic siblings would be expected to be HLA-identical by descent (that is, identical for both paternal and maternal haplotypes), as, for example, the pair in Figure III-7; 50% would share only one haplotype, a state generally termed haplo-identity;* and 25% would be non-identical (Figure III-8).

---

*There has been some confusion in the literature over this terminology, and it is important to ascertain exactly what each author means.

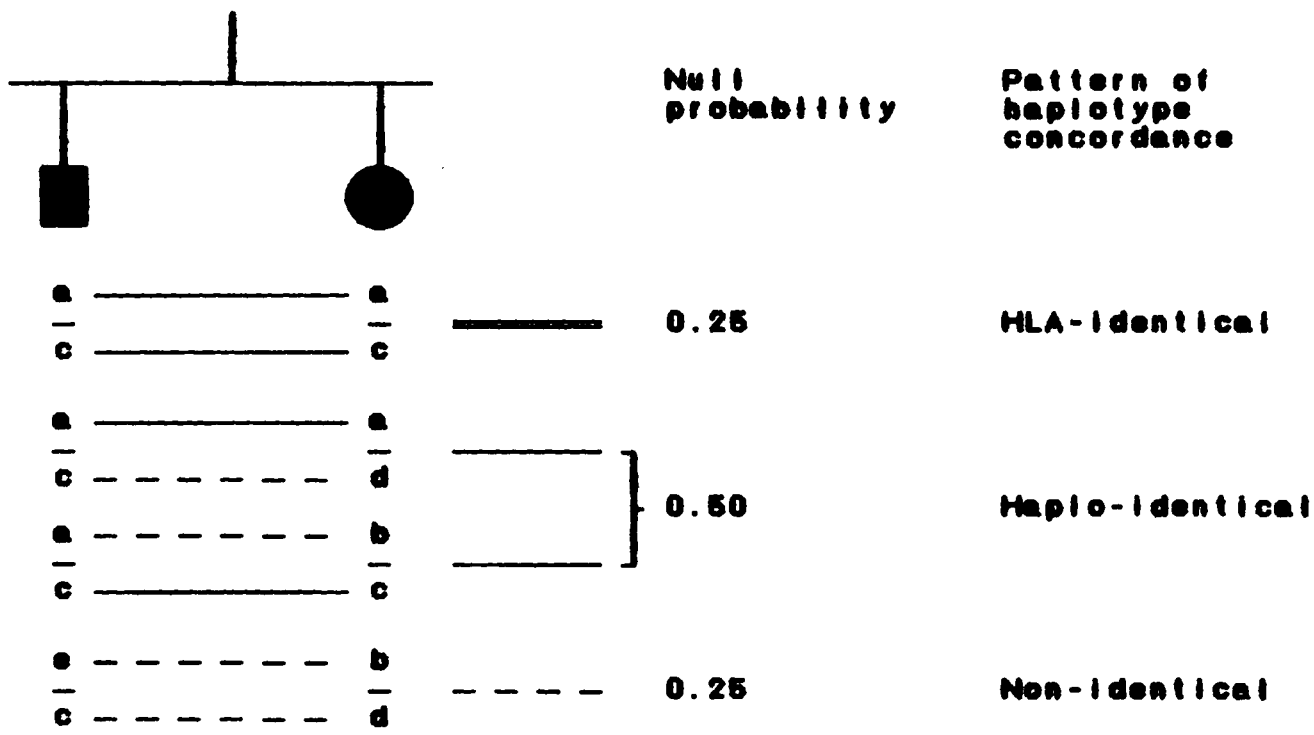


FIGURE III-8. EXPECTED HAPLOTYPE CONCORDANCE FOR DIABETIC SIB PAIRS, ASSUMING NO LINKAGE.

In the presence of linkage we would expect an excess of HLA-identical and perhaps of haplo-identical pairs. The exact proportions would depend on several variables: the closeness of linkage, the mode of inheritance of susceptibility, the population gene frequency, and related to this the penetrance of the susceptibility genes. Thus an accurate empirical determination of the proportions of sib pairs in the three HLA haplotype concordance categories would be important not only in demonstrating the existence of linkage (for which more powerful analyses are available - e.g. the shared-haplotype scoring method of Green & Woodrow, 1977), but perhaps also in helping to quantify these four variables.

The first study of this kind involved 17 families (Cudworth & Woodrow, 1975). Of 15 Type 1 diabetic sib pairs without further diabetic siblings, 10 (that is, 67%) were HLA identical, 7 were haplo-identical and only one was non-identical. This confirmed linkage between HLA and Type 1 diabetes ($p = 0.001$). In contrast, juvenile Type 2 diabetes (MODY) does not appear to be HLA-linked (Nelson & Pyke, 1976; Faber et

al, 1978; Platz et al, 1982). Barbosa et al (1976) did report in an abstract a positive lod score, but this was not significant, nor it seems was it upheld by investigation of additional families (Platz, P, personal communication, 1981).

Since 1977 there have been several further reports concerning Type 1 diabetes. The largest series is the UK study by Walker & Cudworth (1980), based on the Paediatric Register of the British Diabetic Association and thus concerning cases diagnosed before age 16y (see also RESULTS, VIII-2-B-i).

An important implication of the findings in Type 1 diabetic multiplex sibships is that a non-diabetic sibling of an affected child is himself likely to be genetically susceptible if HLA-identical to the diabetic child, whereas haplo- and non-identical siblings are progressively less likely to be at risk. This argument does not depend on the presence or absence of the 'high risk antigens' defined in population studies.

Research involving HLA-genotyped families, each including a diabetic proband with one or more sibling(s), promises to allow an empirical approach to the assessment of risk in siblings (discussed later), and may help to elucidate the nature of susceptibility. Furthermore, investigation of such families may permit identification of non-genetic factors in aetiology of any future diabetes in family members. These concepts form part of the rationale for the present work.

6-C-iii. Family studies: other aspects

Apart from establishing linkage and opening the way to prospective studies and assessment of risk in siblings, HLA investigations in Type 1 diabetic families are important for other reasons. First, they contribute to measurement of associations between HLA and diabetes. In particular, they enable HLA genotypes to be determined, thus making it possible to distinguish homozygotes (e.g. B8,8) from heterozygotes including 'blank' (e.g. B8,-); and furthermore they reduce the likelihood of mistyping since doubtful results can generally be corroborated in

relatives. Hitherto no HLA genotype study of Type 1 diabetes has been sufficiently large to show once and for all where the RR's in subjects homozygous for 'high risk antigens (e.g. B8,8, or DR3,3) lie in relation to those pertaining to a single such allele on the one hand, or to heterozygosity for alleles from both 'S axes' together (e.g. B8,15 or DR3,4) on the other hand. A large family study could make an important contribution here.

Another interesting point was the observation by Suciu-Foca & Rubinstein (1976) of a highly significant increase in frequency of intra-HLA crossovers, detected in six out of 48 informative meioses in 37 diabetic families. Others could not confirm this, finding recombination fractions close to the expected values of <1%, both between A and B and between B and D loci, in several hundred informative meioses (Barbosa et al, 1978b; Platz et al 1978). When the Columbia University team had genotyped more families their recombination rates of 7/352 for A-B and 5/411 for B-D crossovers were no longer significantly higher than in control families (Suciu-Foca et al, 1979). However, although the raised crossover rate was not confirmed, the fact that each of the four recombinant diabetic children was HLA D/DR identical to a diabetic sibling supports the view that the susceptibility gene(s) are near, or beyond, the DR locus (Rubinstein et al, 1976; Suciu-Foca et al, 1979).

A third point arising from HLA studies in diabetic families concerns segregation of the haplotypes possessed by the diabetic children among their non-diabetic siblings. Barbosa et al (1977) investigated haplotype concordance in three categories of sib pairs fom 24 diabetic multiplex families: 'both-diabetic', 'one-diabetic', and 'neither-diabetic' pairs (our terminology). Their aim was to confirm that the disturbance in distribution of 'both-diabetic' pairs among the three haplotype concordance categories shown by Cudworth & Woodrow (1975) did reflect HLA linkage of Type 1 diabetes. This they hoped to achieve by using the 'one-' and 'neither-diabetic' pairs for control purposes, the assumption presumably being that these pairs ought to show the distribution expected by Mendelian theory, unless the disturbance in distribution of 'both-diabetic' pairs was in fact due to non-random segregation of HLA

haplotypes in the first place. The use of such highly dependent data for control purposes is highly questionable, but it is reassuring that only the 'both-diabetic' pairs were over-represented in the HLA-identical group. The remaining pairings showed a minor deficit of 'one-diabetic' and excess of 'neither-diabetic' pairs in the HLA-identical category, but this was not significant.

Somewhat later Bengsch et al (1978) noted that, among 14 non-diabetic siblings of 15 diabetic sib-pairs, only one was HLA-identical to the index case - a non-significant deficit compared with the 'expected 25%' but sufficient to lead them to postulate 'negative selection against the HLA-associated diabetogenic factors independent of overt diabetes, for example by an increased number of early abortions'. In a further 15, simplex, diabetic families these authors found no deficit of HLA-identical non-diabetic siblings of the probands; but even if they had, this would not have been surprising, since ascertainment of families through diabetic probands is bound to introduce such a bias if Type 1 diabetes is HLA-linked (see RESULTS, VIII-1-B-v-a).

6-D OTHER GENETIC MARKERS

In addition to HLA itself, other polymorphic systems on Chromosome 6 have been investigated in the hope of finding stronger associations with Type 1 diabetes.

Factor B has four allotypes distinguishable by immunoelectrophoresis: BfF and BfS are common, and BfF1 and BfS1 are rare in most white Caucasoid populations. In Liverpool, UK, Cudworth, Usher & Woodrow (1977) found only a weak association of BfS with Type 1 diabetes, attributable to known linkage disequilibrium with HLA-B8. Two years later there were reports of stronger associations between Type 1 diabetes and BfF1, this allotype being present in 23% of 106 white Caucasoid patients and only 2% of controls in Massachusetts, USA, giving a RR of 15 (Raum et al, 1979). In some cases the finding of BfF1 was corroborated by genotyping relatives, and for this and other reasons it appeared unlikely that the high frequency in diabetics was artefactual (due, for example, to glycosylation of BfF).

In other populations some workers reported less striking but still significant associations, in West Germany (Bertrams et al, 1979, 1981), Australia (Kirk et al, 1979) and Minnesota, USA (Barbosa et al, 1979b), whereas others found moderate frequencies in diabetics which, while higher than in controls, were not significantly so, in upper midwest USA (Weitkamp et al, 1979) and Northumbria, UK (Bernal et al, 1979).

In each of the last two reports a possible reason for the apparent disagreement between different analyses was suggested: ethnic variation, and heterogeneity with respect to age at onset of Type 1 diabetes, respectively. In support of the first suggestion is the fact that BfF1 is in linkage disequilibrium with HLA-B18, a specificity known to be particularly frequent in south-western Europe, where it is typically found in the haplotype Aw30,Cw5,B18,BfF1,DR3 (Deschamps et al, 1979; Weitkamp et al, 1979). De Mouzon et al (1979) therefore examined French Basques and reported the highest known frequencies of BfF1, 78% and 24% respectively in 51 Type 1 diabetics and 87 controls, giving a highly significant RR of 11.4. Furthermore, the 'unexpressed' C4B allotype C4BQ0 (then known as $C4s^{-}$) was also excessively frequent in diabetics, and there was an enhanced RR for homozygosity of this allotype, although without any deviation from Hardy-Weinberg equilibrium. In 53% of patients compared with 13% controls the whole haplotype Aw30,Cw5,B18,BfF1,C4BQ0,DR3 was present (RR = 7.8). Relative Δ values for linkage disequilibria between these alleles were higher in the patients than in controls, confirming the findings of Bertrams et al (1979, 1981) for $r\Delta_{BfF1,B18}$. Dornan et al (1980) made similar observations in Newfoundland. Despite the impressive RR for BfF1, De Mouzon et al (1979) considered that the primary association of Type 1 diabetes was still with HLA-DR3 (RR = 23), as in other populations.

The question of age-at-onset-related differences in the BfF1-Type 1 diabetes association was suggested by Bernal et al (1979), who noted that although 44 of their 64 patients had been diagnosed after age 17y, four of the five who were BfF1-positive reported onset before this age ($p = 0.031$ by our calculation using Fisher's exact test). Others have

supported this idea (Barbosa et al, 1979b; de Mouzon et al, 1979; Bertrams et al, 1981) but it cannot be considered proven. This is discussed further in Chapter III-10.

Apart from the work in Basques, little was known at the time of the present work about possible associations between C2 and C4 allotypes and Type 1 diabetes. More recently, Dornan et al (1981) found no significant difference in C2 allotype frequencies between 55 patients and 800 controls in Newfoundland, but positive associations have been reported in West Germany for C4AQ0, C4BQ0, C4B4 (RR = 4.7) and possibly C4B2 (RR = 1.6) (Rittner & Bertrams, 1981; Bertrams et al, 1982).

Another polymorphic system on Chromosome 6 in man is that determining glyoxalase-I (GLO) (Kompff et al, 1975; Kompff et al, 1977; Mayr et al, 1978). The two allotypes recognized, 1 and 2, are in Hardy-Weinberg equilibrium in all most populations, an exception being European Lapps, who show a deficiency of heterozygotes (Busi et al, 1979; Parr et al, 1977). Kirk et al (1979) found no significant associations in 239 Australian Type 1 diabetic patients control subjects, but McCann, Davis and Welborn (1981) recently claimed in an abstract that there was a significant excess of GLO 1-1 in a sample of only 28 Australian patients. They did not correct the p value (<0.01) for the number of comparisons made, nor did they state that the patients were unrelated, so this claim should not be accepted without confirmation. Further studies are needed.

Although susceptibility to Type 1 diabetes appears to be principally HLA-linked, a contribution from genes on other chromosomes has not been excluded. Lucarelli et al (1978) investigated nine polymorphic systems in a series of 138 patients. Unfortunately once again the statistical analysis is suspect, in that these authors neither performed a multi-way analysis nor corrected their p values for the total number of comparisons made, a number which must have exceeded 40. Thus their claim of an association with the red-cell acid phosphatase CB phenotype in males only (χ^2 = 7.071, p <0.01) can probably be discounted. The same applies to the increased frequency of blood group B they found in female Type 1 diabetics. Somewhat more impressive, but still requiring confirmation,

is their reported excess of blood group M homozygous (MM) male diabetics (X^2 = 14.011, p <0.001). They observed no deviations from control phenotype frequencies for phosphoglucomutase (locus 1) (PGM_1), adenosine deaminase (ADA), adenylate kinase (AK), glucose-6-phosphate dehydrogenase (G-6-PD), haptoblobin (Hp) or rhesus (Rh). Farid et al (1978) found no association between Type 1 diabetes and IgG heavy chain allotypes (Gm) or C3 variants; only 36 patients were studied.

The Lewis (Le) blood group system has proved controversial. Vague et al (1978) reported increased frequency of the Le(-) blood group in both Type 1 and Type 2 diabetes, but this has been disputed (Leslie & Pyke, 1979). Mourant et al (1978) emphasized that for reliable Le typing the saliva rather than simply the red blood cells should be tested, pointing out that:

> *'because of the many interconnections with the [Secretor] system... and the weakness of most of the available reagents, the establishment of the Lewis genotype of an individual is a matter of some technical difficulty, so that many of the published data on frequencies in populations must be accepted with some reserve.'*

More recently Hodge et al (1981) reported weak but significant linkage between Type 1 diabetes and the Kidd (Jk) blood group system coded on chromosome 2. A larger study using virtually identical techniques in a different population failed to confirm this; the likely explanation, a chance sampling bias in the earlier study (Dunsworth et al, 1982), was confirmed when that study was enlarged and no significant association was found (Hodge et al, 1983).

However, recent evidence in favour of Type 1 diabetic susceptibility genes independent of HLA includes the recent demonstration by Bodansky et al (1981) of a weak but significant association with the fast acetylator phenotype, and by Bell et al (1984) of an association with insulin-gene restriction-fragment length classes.

6-E. MODE OF INHERITANCE OF SUSCEPTIBILITY TO TYPE 1 DIABETES

Much of the controversy over this topic can be resolved into two questions: is Type 1 diabetes genetically homogeneous; and if so, is susceptibility determined by a single gene, or by two or more?

6-E-i. Genetic heterogeneity within Type 1 diabetes

Genetic heterogeneity within Type 1 diabetes with respect to age at onset, already mentioned here in relation to BfF1, was first suggested by Cudworth & Woodrow (1975), who noted widely (but not significantly) varying frequencies of HLA-B15 in younger age groups, and by Cathelineau and colleagues in the same year (see Cathelineau et al, 1976) on the basis of an increased frequency of this antigen only in cases diagnosed before age 10y. Later, Ludwig, Schernthaner & Mayr (1977a) found the reverse, B15 appearing important only in later-onset cases, and on their expanded Liverpool series Cudworth & Woodrow (1976) found no significant heterogeneity in this respect, nor did Farid et al (1978) in Newfoundland.

Nevertheless, Ludvigsson, Safwenberg & Heding (1977) noted that the mean onset age of their cases positive for both HLA-B8 and B15 was somewhat lower (5.6 years) than that of the remainder (7.4 years), although this was not significant ($p < 0.1$). Kirk et al (1979) and Bertrams et al (1981) claimed the association with B15 to be particularly prominent in their childhood-onset cases. Christy et al (1979b) at first lent their support with the observation that among 43 newly-diagnosed Danish cases aged up to 15y, 49% were positive for HLA-Dw4 but not Dw3, compared with only 23% of 30 aged >15y. However this same group was recently unable to find any significant heterogeneity with respect to age in a much larger series of D- and DR-typed cases (Platz et al, 1981). Thus this whole concept remains highly controversial.

Type 1B, or 'primary autoimmune', diabetes is an uncommon clinical variant of Type 1 diabetes characterized by onset typically in adult life, a female preponderance, a particularly marked association with organ-specific autoimmunity, and persistence of islet cell antibody for

many years after diagnosis (see Chapter III-7). The principal HLA association of this variant appears to involve 'S1 axis' specificities rather than both S1 and S2, suggesting that Type 1B diabetes is, at least in part, genetically distinct from the classical 'Type 1A diabetes' typically of juvenile onset (Bottazzo & Doniach, 1976; Morris et al, 1976; Bottazzo et al, 1978a), even when biased ascertainment is allowed for (III-10-iii-c).

Another form of heterogeneity which has been suggested is between familial and sporadic Type 1 diabetes. By this hypothesis at its most extreme, some Type 1 diabetes would depend on genetically determined (and HLA-linked) susceptibility, whereas other cases would represent phenocopies, without such susceptibility. A less extreme hypothesis proposes that genetic factors are present in all cases, but are more important in some than in others where environmental factors take on a greater role. Alternatively, two or more distinct types of susceptibility may be transmitted by different genes with different modes of inheritance, and hence different patterns of recurrence within pedigrees.

Nelson et al (1975) investigated this question in identical twins. Among 31 pairs concordant for Type 1 diabetes 18 (58%) were HLA-B8-positive, compared with only 10 (32%) of 31 discordant pairs. HLA-B15 was equally common in both groups. The tentative conclusion, accepted by Rotter & Rimoin (1978), was that there might be heterogeneity of the less extreme type just mentioned, a susceptibility gene associated with B8 being more penetrant than one associated with B15. However, Nelson et al acknowledged that further data were needed, and even before correcting for the number of comparisons made the difference in frequency of B8 was not significant (by our calculation, $X^2_{(1)} = 3.19$ with Yates' correction). Bengsch et al (1978) found no significant differences in HLA phenotype frequencies between Type 1 diabetics with (n = 58) and without (n = 109) a positive family history of the disease; indeed, this time it was if anything B15 which was more frequent in the familial group.

More recently, combined data from several centres taking part in the 8th Histocompatibility Workshop suggested minor differences, with a slight excess of 'DR3 or 4 alone' (i.e. possible homozygotes) among the familial cases, but this was of debatable significance (Svejgaard et al, 1980). Thus what is needed is a large-scale study comparing familial with sporadic cases in the same population.

The claim of Barbosa et al (1980a) that there was heterogeneity between 'multiple generation' and 'multiple sib' Type 1 diabetes was not accepted (see below, 6-E-vi).

6-E-ii. Single-gene hypotheses

In the absence so far of convincing evidence for heterogeneity within Type 1 diabetes in childhood, what, then, is the mode of inheritance of susceptibility? Rubinstein, Suciu-Foca and Nicholson (1977) were convinced that it was autosomal recessive, with 50% penetrance. This was based on not more than 11 multiplex and few more than 16 simplex diabetic families (the exact numbers are not given); the assumption of 50% penetrance is nowhere adequately justified, and Barbosa et al (1978a) and Suarez & Hodge (1979) showed that these data were compatible with absence of any linkage at all.

In the same year, Thomson & Bodmer (1977) developed a method for distinguishing between the two Mendelian autosomal single-gene hypotheses: dominant (i.e. homozygote and heterozygote equally penetrant), and recessive (no risk for heterozygote). The method is based on the distribution of HLA haplotypes among affected siblings, and assumes close HLA-linkage. The recessive model fitted the data for gene frequencies between about 0.2 and 0.3, but no good fit could be obtained for the dominant model. Similar conclusions, with more definite exclusion of the dominant model, have been reached by others using the same method on larger bodies of data (Walker & Cudworth, 1980; Svejgaard et al, 1980; Platz et al, 1981).

However, the recessive theory is also difficult to accept. The model predicts that the lowest acceptable gene frequency for the HLA-

linked susceptibility gene is 0.23. If population prevalence at age 16 is taken as 0.002 (see Chapter 4), then penetrance at this age would be only $0.002/(0.23)^2 = 0.038$, or 3.8%. Thus 3.8% of susceptible siblings, and therefore only about 1% of all siblings of affected children, would be diabetic by age 16, a prediction so absurdly low (see 6-F) as to render the recessive model, like the dominant, untenable. Barbosa et al (1978a) viewed the problem from another angle. They showed that, for the data to fit the single-gene recessive hypothesis, the recombination rate between the gene and HLA would have to be 15-20%. This is much too high for the existence of the linkage disequilibrium supposed to underlie the population association between HLA and Type 1 diabetes.

Spielman, Baker & Zmijewski (1980), having earlier favoured dominant inheritance, recently showed that intermediate inheritance gives a significantly better explanation for family HLA data (in 164 families) than either the recessive or the dominant hypothesis. Using a population prevalence at age 30y of 0.004, their 'best fit' model gives a penetrance of 0.71 when the gene is homozygous and 0.065 when heterozygous, with a gene frequency of 0.03.

6-E-iii. Other hypotheses

With or without gene dosage effects, no single-gene hypothesis can explain the enhanced relative risk believed to exist in HLA-DR3,4 heterozygotes compared with homozygotes. Those who accept this enhanced RR therefore favour the idea that susceptibility is determined by interaction between two or more, different, HLA-linked genes (not necessarily at the same locus), one in linkage disequilibrium with DR3 and the other with DR4, as originally suggested by Cudworth & Woodrow (1975, 1976) on the basis of the high RR of the B8,15 phenotype. Thus accurate determination of the RR's of DR3,3 and DR4,4 homozygotes in relation to DR3,4, DR3,X and DR4,X heterozygotes (or similarly for the associated B-locus genotypes) is important. In this connection Curie-Cohen (1981) has commented that, in order to compare the relative risks of such compound genotypes, all RR's should be calculated relative to the same subpopulation, rather than to the population of all those lacking the genotype concerned as is generally done.

If the DR3,4 heterozygote is indeed at especially high risk of diabetes by comparison with the homozygotes, then a disturbance of Hardy-Weinberg equilibrium is expected in diabetics. This offers an alternative test of the hypothesis (see RESULTS, VIII-2-B-iii-a)

6-E-iv. A protective gene?

The proposed existence of a gene protecting against Type 1 diabetes, in linkage disequilibrium with the 'R axis' of Figure III-7, is another unresolved question. As already mentioned, HLA-B7, DR2 and DR2 carry significantly reduced RR's, but this could in theory be a secondary effect: DR3 and DR4 are so common in Type 1 diabetes that the other DR specificities are bound to be scarce, simply because there is less room for their alleles on the haplotypes. This is reflected in the Eighth Workshop data: all DR specificities analysed, except 3 and 4, showed RR's less than 1.0. DR2 would be expected to show the most marked such effect because it is the commonest antigen in most of the populations studied. However its frequency in Type 1 diabetes is very low indeed, and a protective gene has not been ruled out. Barbosa et al (1979a) estimated the expected gene frequencies of Dw2 in diabetic and control populations assumed free of Dw3, and still obtained a 'significant' reduction in frequency of Dw2 in the diabetics. However the method is not explained in sufficient detail to be convincing. For a definitive answer it will probably be necessary to examine segregation of DR2 to diabetic and non-diabetic individuals from parents positive for DR2 but not DR3 or DR4 (cf. RESULTS, VIII-2-B-iii-c).

6-E-v. Epistasis between HLA-DR3 and DR4?

Clerget-Darpoux & Bonaiti-Pellié (1980) have pointed out that HLA-linked susceptibility genes are not the only possible explanation for the HLA associations of Type 1 diabetes in the population and the haplotype distribution pattern among affected sib-pairs. A plausible alternative is an epistasis effect, whereby HLA genes themselves influence expression of a dominant susceptibility gene which could be located on any chromosome. Similar modification by MHC of expression of a (recessive)

gene is believed to underlie the spontaneous diabetes of the BB rat (Howard, 1983; III-7-D-iii).

There seems at present to be no compelling evidence in favour of this idea in human Type 1 diabetes, but it could be tested by investigating haplotype concordance between diabetic siblings for haplotypes from parents lacking both DR3 and DR4. By the epistasis hypothesis, such haplotypes should show a more random distribution to the diabetic sib-pairs than haplotypes from parents positive for DR3 or 4, (cf. RESULTS, VIII-2-B-iii-c).

Direct involvement of MHC gene products has implications for the mechanism of action of susceptibility genes (III-11-E).

6-E-vi. Susceptibility genes not linked to HLA

The possibility that genes not linked to HLA might contribute to susceptibility to Type 1 diabetes (6-D) has been considered by several authors. Barbosa et al (1980a) failed to demonstrate significant linkage between HLA and Type 1 diabetes in 28 pedigrees with multiple cases in different generations ('MG families'), unlike 'MS families' (with multiply-affected sibships). These two groups had been chosen as most likely to represented homogeneous autosomal dominant and autosomal recessive subgroups of the disease. One possible explanation for their findings was that susceptibility in MG families was determined by non-HLA-linked genes. However, these authors acknowledged that their findings were inconclusive.

Soon afterwards Suarez & van Eerdewegh (1981) pointed out the ascertainment bias inherent in this work, showing that the data were equally compatible with a genetically homogeneous disease. Furthermore, lod-score analysis depends on fitting models; exclusion of linkage according to a particular model (single-gene dominant) is irrelevant to other models, such as two-gene interactions. Thus there is no real evidence here for non-HLA-linked genes.

Nevertheless, Thomson (1980) has demonstrated that HLA haplotype concordance data in affected sib-pairs, and other population parameters for Type 1 diabetes, are compatible with a 'two-locus plus environment model', one locus being HLA-linked and behaving as a single-gene recessive with gene frequency 0.35, and the other not being HLA-linked at all.

A proven association with a marker known to be on another chromosome would be important evidence (6-D). Pending that, however, the question remains open.

6-F. RISK OF TYPE 1 DIABETES IN SIBLINGS OF AFFECTED CHILDREN

Several approaches to this problem are possible. The simplest is an empirical estimation of overall prevalence of diabetes in siblings of affected patients. To be of any practical value the applicability of the estimates should be defined in terms of age, sex, era, and geographical situation.

Among the earlier estimates (not necessarily excluding Type 2 diabetes), Harris (1950) recorded a prevalence of juvenile-onset diabetes (before age 30) of 4.1% among the siblings of such cases in London; Steinberg & Wilder (1952) found 6.0% by similar means in USA; and Simpson (1962) noted that 5.7% of the siblings of her Canadian cases, diagnosed by age 20y, were themselves affected. Darlow, Smith & Duncan (1973) give a figure of 8% in first-degree relatives in Edinburgh, but siblings are not treated separately. All these figures depend on the age distribution of the siblings. In a careful Danish study, Degnbol & Green (1978) estimated the risk of developing diabetes by age 30y, in siblings of such juvenile cases, to be about 5%. They pointed out that this was with 'large uncertainties'.

Recently Gamble (1980) used data from the Paediatric Register of the British Diabetic Association to calculate a potential prevalence of 5.7% by age 16y in siblings of patients diagnosed before that age. He

estimated that this was 26 times that in the general population at this age, using the data of Christau et al (1977) for comparison.

Another approach is to decide on the mode of inheritance and penetrance of susceptibility and thence to use Mendelian rules to calculate risks in siblings where both, one or neither parent is affected. This was done by Steinberg & Wilder (1952), treating diabetes as a single disease, and again by Rubinstein et al (1977) for Type 1 diabetes on the basis of their contentious '50% penetrant recessive' theory, taking into account HLA haplotype concordance with the proband. In view of the uncertainties over mode of inheritance, this approach is not justified.

Platz et al (1981) attempted to estimate indirectly the risk of Type 1 diabetes in HLA-identical, haplo-identical and non-identical siblings by age 20, as follows. They assumed a potential prevalence of 5.1% at this age in siblings overall (based on the data of Degnbol & Green - see above). They used the observation that 58.6% of affected sib-pairs were HLA-identical to each other to calculate that 3% of siblings would be both HLA-identical to their probands and diabetic at age 20y (0.586 x 0.051 = 0.030). Assuming that 25% of all siblings of the probands are HLA-identical to them (which is not necessarily so: 6-C-iii; RESULTS, VII-1-B-v), and that there is no heterogeneity between 'familial' and 'sporadic' diabetes (see above), they estimated that 0.03/0.25 = 0.12, or 12% of HLA-identical siblings would be diabetic by that age. Similarly they obtained figures of 4% and 1% for haplo- and non-identical siblings, respectively. So far there is no reliable empirical estimate of risk in siblings grouped according to HLA haplotype concordance with the diabetic probands.

For more accurate prediction of Type 1 diabetes in siblings it may in the future be possible to take into account not only HLA-linked but also perhaps other genes, as well as environmental factors. There is no doubt that a marker of the pre-diabetic state would be of great value. A possible candidate for this role, islet cell antibody, is discussed in the next chapter.

PART III — CHAPTER 7

TYPE 1 DIABETES AS AN AUTOIMMUNE DISEASE

CONTENTS

A. Introduction. Effects of diabetes on immunity. 95
B. Associated non-pancreatic autoimmunity 95
C. Pancreatic autoimmunity 98
i. Islet cell antibodies (ICA) 98
a. 'Conventional' ICA 98
b. Complement-fixing (CF-)ICA 102
c. Islet cell surface antibodies (ICSA) 102
d. A pathogenetic role for IC(S)A? 103
ii. Cell-mediated immunity (CMI) 104
a. Leucocyte migration inhibition 104
b. Circulating immunocyte populations 105
c. Target specificity of circulating immunocytes . . 107
d. Experimental work: passive transfer of diabetes? 108
D. Animal models 108
i. Experimental immunization 108
ii. Low-dose streptozotocin diabetes in mice 109
iii. The BB rat 109
E. Conclusion 110

7-A. INTRODUCTION

A plausible explanation for the phenomenon of insulitis is an autoimmune process directed against the islet B cells during the period leading up to diagnosis of Type 1 diabetes (Chapter 3). This notion is particularly attractive in the light of the HLA association (III-6): most of the genes recognized in the HLA region code for cell surface (and therefore immunologically important) components, or else for other molecules involved in immune reactions.

Most of the available evidence has to be interpreted with caution. One problem is that B cell destruction is already about 90% complete by the time of presentation, and investigation of diabetics themselves is largely retrospective. Furthermore the islets are highly inaccessible in life, newly-diagnosed Type 1 diabetics nowadays rarely come to necropsy, and most forms of human experimentation which might give direct evidence are unethical or impracticable. As a result most of the evidence in man is indirect.

Another problem is that diabetes itself may modify immune responses. It is well recognized that poorly-controlled patients are abnormally susceptible to certain infections (Wheat, 1980), but the mechanism is not always clear. There is no doubt that poor metabolic control can adversely affect function of some immunocytes. Thus, polymorphonuclear leukocytes show impaired chemotaxis, phagocytosis, intracellular killing and perhaps immune adherence in untreated diabetics (Bagdade, 1976). Similarly, lymphocytes from poorly controlled diabetics show impaired glucose metabolism and reduced phytohaemagglutinin (PHA) transformation compared with normals (Brody & Merlie, 1970). Further evidence that abnormal function and numbers of at least some T lymphocyte subpopulations may result from diabetes will be discussed below (7-C-ii).

7-B. ASSOCIATED NON-PANCREATIC AUTOIMMUNITY

For some time it has been recognized that Type 1 diabetes tends to associate in individuals and families with certain other conditions

believed to be autoimmune diseases (AID). Examples include primary AI myxoedema, Hashimoto's thyroiditis and Graves' disease (Hayles et al, 1959; Beaven et al, 1959; Perlman, 1961; Carpenter et al, 1964; Whittingham et al, 1971; Bastenie, 1975; Gray & Clarke, 1978); Addison's disease (Beaven et al, 1959; Carpenter et al, 1964; Bastenie, 1975; Riley, Maclaren & Neufeld, 1980); latent or overt pernicious anaemia (Ungar et al, 1968; Irvine et al, 1970; Bastenie, 1975; Gray & Clarke, 1978); and coeliac disease (Thain, Hamilton & Gholich, 1974). Like Type 1 diabetes, each of these conditions (except for isolated pernicious anaemia and perhaps Hashimoto's thyroiditis) is associated with HLA-DR3 or B8 in population studies (Ryder, Andersen & Svejgaard, 1979). As stated by Solomon & Blizzard (1963),

> *'The clinical associations of various autoimmune endocrine diseases suggests that investigation of a possible immunologic disturbance in this disease is also merited.'*

There is an even more striking association between Type 1 diabetes and serological organ-specific autoimmunity (OSAI) to thyroid and gastric parietal cell microsomal antigens. Pettitt, Landing & Guest (1961; see also Landing et al, 1963) found thyroid cytoplasmic antibody (Ab) in 13% of 83 juvenile diabetics compared with 1% of control children. Simkins (1968) reported similar findings in younger, but not in elderly, diabetic patients. Irvine et al (1970) examined sera of 871 control subjects and 1054 diabetics of various ages. Two thirds of the latter, including most of those aged < 40y, were insulin-dependent. In these both thyroid and gastric parietal cell Ab were more prevalent, especially in females, than in controls or in diabetics not dependent on insulin.

From this, and from the work of others (Goldstein et al, 1970; Nerup & Binder, 1973; Nissley et al, 1973; Bastenie et al, 1975; Fialkow, Zavola & Nielsen, 1975; Lendrum et al, 1976; Bottazzo et al, 1978b; Karmazsin et al, 1979; Nagaoka et al, 1979; Neufeld et al, 1980), it can be stated that in general approximately 20% of Caucasoid Type 1 diabetic children and young adults have thyroid and/or gastric parietal cell (thyrogastric) Ab, with a 2:1 female preponderance, whereas only about 2-5% of control sera are positive. Later in life the control prevalence

rises, but the excess in the diabetics remains. Dorchy et al (1981) found only a minor and non- significant excess in their juvenile diabetics, but their control population was inappropriate in that it consisted of blood donors who were much older. In contrast, Type 2 diabetics do not appear to show any increase in thyrogastric OSAI.

It might be held that this increase in humoral OSAI in Type 1 diabetes could be secondary to the diabetes itself (Bastenie, 1975). This seemed unlikely to Irvine et al (1970) and to Fialkow et al (1975), because the prevalence of thyroid and gastric parietal cell Ab was independent of duration of diabetes in their series. Perhaps a more powerful argument is based on the finding of an increased prevalence of these Ab in clinically unaffected relatives of Type 1 diabetics (Nissley et al, 1973; Fialkow et al, 1975; Bottazzo et al, 1978b). A few diabetics and relatives of diabetics also have adrenal steroidal cell cytoplasmic Ab, which is otherwise very rare except in Addison's disease (Carpenter, 1964; Nissley et al, 1973; Nerup & Binder, 1973; Irvine & Barnes, 1975; Bottazzo et al, 1978a; Neufeld et al, 1980; Riley et al, 1980).

Recently there has been some evidence of an association between Type 1 diabetes and non-organ-specific AI as manifested by Ab to nucleic acids (Huang and Maclaren, 1978; Huang et al, 1981) or to cell nuclei (Karmazsin et al, 1979; Dorchy et al, 1981).

Since most OSAID, like Type 1 diabetes, is associated with HLA-DR3, and there appears to be an inherited predisposition to OSAI in diabetic families, a reasonable proposition is that the HLA-DR3-associated diabetes susceptibility gene is also responsible for the OSAI diathesis. Until our own work, however, this hypothesis had not been tested in HLA-genotyped diabetic families.

7-C. PANCREATIC AUTOIMMUNITY

7-C-i. Islet cell antibody (ICA)

7-C-i-a. 'Conventional' ICA

Although Mancini, Constanzi & Zampa (1964) described patchy islet

fluorescence in a direct immunofluorescence test using fluorescein-labelled sera from 9 of 17 diabetics, this remained unconfirmed and has been attributed to insulin Ab (Bottazzo, personal communication). Not until 1974 were there convincing reports of a different antibody specific to the islets of Langerhans (Bottazzo, Florin-Christensen & Doniach, 1974; McCuish et al, 1974a). The standard technique is an indirect immunofluorescence test on sections of unfixed human pancreatic tissue taken from kidney donors and stored at -20°C, although islet cell antibody (ICA) is not usually species-specific.

As determined by this conventional technique, ICA binds to an unidentified cytoplasmic antigen common to all types of endocrine cells in the islets, which stand out as uniformly bright areas in the immunofluorescence test. Although this pattern is quite distinctive the fluorescence is much less bright than in the equivalent tests for thyroid or adrenal cytoplasmic antibodies, for example, and the initial identification of ICA in 1974 depended to some extent on technical advances in ultraviolet light microscopy (Bottazzo, Doniach & Pouplard, 1976; Bottazzo, personal communication). Background fluorescence due to presence of blood group antigens in exocrine pancreatic cells is obviated by using only pancreata from group O Rh+ donors (Lendrum & Walker, 1975). Occasional sera react only with A or D cells (Bottazzo, Doniach & Pouplard, 1976).

These first reports of ICA were in adult Type 1 diabetics with other AI endocrine diseases, but soon afterwards the new antibody was demonstrated in a proportion of diabetic children without other ditions (Lendrum, Walker & Gamble, 1975; Bottazzo & Doniach, 1976; Ludwig, Schernthaner & Mayr, 1977b). There seemed, at last, some justification for the classification of Type 1 diabetes as an OSAID analagous to AI thyroiditis and Addison's disease. However, it is now clear that the analogy is far from straightforward, and that ICA itself is unusual in a number of respects.

For example, although thyroid and other OS autoantibodies generally remain positive for many years once they have appeared in an individual,

and can still be found in most cases of clinical disease long after its onset, this does not apply to ICA. The prevalence of ICA may be as high as 85% at the time of diagnosis of Type 1 diabetes, but falls after a few months and declines further thereafter, leaving only a remnant of 10-20% of patients with ICA persisting over many years (Lendrum et al, 1976; Irvine et al, 1977; Del Prete et al, 1977; Neufeld et al, 1980).

Those patients with persistent ICA appear to form a sub-group (Type 1B diabetes) characterized in addition by female preponderance, onset typically in adult life (often in the fourth decade), a particularly marked tendency to OSAI, and an especially strong association with HLA-B8 rather than with B15 (Bottazzo & Doniach, 1976; Morris et al, 1976; Irvine et al, 1977; Bottazzo et al, 1978a,b). In no individual study is the association between persistent ICA and HLA-B8 in Type 1 diabetics significant beyond the 2.4% level, and unfortunately the results are presented in different ways so that a combined analysis is impossible. However, there is little doubt that this is a genuine association, since different reports agree. Nevertheless, further studies (preferably including HLA-DR) are needed to confirm this important evidence of genetic heterogeneity within Type 1 diabetes (see Chapter III-10).

According to Neufeld et al (1980), in Type 1 diabetes with onset in early childhood (age <5y) ICA is significantly less frequent than in later-onset cases, a finding which at first sight suggests further heterogeneity. However, the difference in ICA frequency between the younger- and older-onset cases in that report is no longer significant when data from the several duration-of-diabetes classes are combined by the method of Woolf and Haldane (see Chapter III-6) to compensate for differences in duration between the two groups: by our calculation, $\chi^2_{(1)} = 2.26$. Neufeld et al themselves noted that within 2y of onset there was no difference at all in ICA frequency between the two age-at-onset groups. It seems, therefore, that it is the rate of disappearance of ICA after diagnosis, rather than its initial prevalence, which varies with age at onset. Perhaps this is another manifestation of rapid progression of the disease process in childhood cases (cf. insulitis, Chapter 3-C-ii).

In non-diabetic populations without OSAI the prevalence of ICA is of the order of 1-2% in hospital controls (Lendrum et al, 1976; Del Prete et al, 1977) and even less in the general population (Del Prete et al, 1977; Irvine et al, 1977; Rodgers et al, 1980; Neufeld et al, 1980), confirming a strong association between the antibody and the disease. Furthermore, allowing for difficulty in categorizing some individual patients as Type 1 or Type 2 diabetics, ICA is not prominent in the latter: Lendrum et al (1976) found it in 6/112 insulin-independent diabetics, and then only in low titre; Del Prete et al (1977) reported it within a year of diagnosis in <10% of adult-onset diabetics not insulin-treated, compared with 19% of those on insulin, while Knowler et al (1979) found 2 ICA-positive subjects among 46 Type 2 diabetic Pima Indians. Irvine et al (1977) detected no ICA in 107 Type 2 diabetics treated by diet alone, but 8% of 217 on oral hypoglycaemic agents were positive. Such cases may represent a 'forme fruste' of Type 1 diabetes, since some later become insulin-dependent (Irvine et al, 1979).

Another unusual feature of ICA is its age and sex distribution. The other OSAb are twice as common in females as in males and show increasing prevalence with age in the general population. In contrast, ICA is about equally common in males and females, and in non-diabetics the prevalence in the second decade appears to equal that in the fifth, at least among relatives of diabetics (Lendrum et al, 1976; Irvine et al, 1977; Bottazzo et al, 1978b; Neufeld et al, 1980). Furthermore, our own work has shown differences in segregation in relation to HLA in diabetic families (see RESULTS, and also Gorsuch et al, 1980). Finally, even the immunoglobulin composition of ICA is unusual in that IgM and IgA have not been reported (Bottazzo et al, 1976) and there is often IgG subclass restriction to IgG2 (Dean, McNally & Doniach, 1980 - work based mainly on results in participants in our own study).

Because of the high prevalence of ICA in Type 1 diabetes at the time of diagnosis, together with its rarity in healthy subjects, it has been suggested that this antibody may be useful as an immunological marker of pre-diabetic damage to islet B-cells (Bottazzo, Doniach & Pouplard, 1976;

Irvine, Gray & McCallum, 1976a). This is supported by reports of increased prevalence of ICA in first degree relatives of Type 1 diabetics (Bottazzo et al, 1978b - 4%; Gorsuch et al, 1980 (see also RESULTS, VIII-1) - 4%; Dobersen et al, 1980 (using a different technique, as explained below) - 12.5%). In a few reported cases the finding of ICA has preceded development of Type 1 diabetes by periods ranging from months to a few years. In the main these were adults, generally with other OSAID (Bottazzo et al, 1974; Lendrum et al, 1976; Irvine et al, 1976a).

In view of the likelihood that these cases represented Type 1B diabetes further evidence was needed before acceptance of the 'marker' role for the more typical, childhood, Type 1 diabetes. Much of the evidence so far reported comes from the present work (Gorsuch et al, 1979, 1980, 1981; and RESULTS, VIII-3-B-vii). In addition, a single child described by Ilonen et al (1980) developed the disease 17 months after ICA (complement-fixing) had been detected when he had presented with transient symptoms; and in a small family study, one of six ICA-positive siblings of affected children became diabetic three months after the first test (Ginsberg-Fellner et al, 1979).

For some time it has been clear that, although ICA may have a 'marker' role, by no means always does its presence herald imminent diabetes. Several individuals (mostly adults) are known to have been ICA-positive for up ten or more years without developing the disease (Bottazzo & Doniach, 1976; Barnett et al, 1981; Irvine et al, 1976a; Rodger et al, 1980). More recently, Mutch, Bottazzo and Stowers (1981) reported 12 ICA-positive subjects out of 70 young patients with impaired glucose tolerance, none of whom had progressed to insulin dependence during a mean review period of 8y. Furthermore, Buschard et al (1977) found ICA in 6/18 unrelated spouses of Type 1 diabetics, a higher proportion than that of first-degree relatives. This curious observation requires confirmation. However, it is true to say that the true value of ICA as a predictive marker of Type 1 diabetes was not at all clear before the prospective study reported here.

7-C-i-b. Complement-fixing (CF-) ICA

In the conventional technique described above, a fluorescein-labelled rabbit antibody to human IgG is used to demonstrate ICA bound to the substrate tissue. Because of the weak fluorescence, interpretation of borderline resuts may vary between laboratories. In an attempt to improve reproducibility the test has been modified to detect complement-fixing (CF-) ICA. Here, after the test serum, a source of additional human complement may be used before the completion of the 'sandwich' by a fluorescein-labelled Ab to human complement component C3. Much of the early work on CF-ICA was performed at the Middlesex Hospital in collaboration with the present author. In general it appears that CF-ICA is more closely associated with the time of diagnosis of Type 1 diabetes than is 'conventional' ICA (Bottazzo et al, 1980a). Irvine (personal communication, 1981) has failed to find any difference between conventional and CF-ICA in this respect, possibly because of variation between laboratories in interpretation of borderline results; unfortunately there is as yet no generally recognized standard for ICA. The use of CF-ICA as a marker of prediabetes is investigated here (RESULTS, VIII-3-B-vii).

7-C-i-c. Islet cell surface antibody (ICSA)

The close association of ICA with Type 1 diabetes, and in at least some cases with the pre-diabetic period, raises the possibility that this antibody might be directly involved in the early stages of pathogenesis. There are, however, two problems to be overcome before this idea is acceptable. First, it is difficult to reconcile the specificity of islet B-cell destruction with the uniform islet fluorescence produced in the conventional test by ICA, which binds not only to the B cells but also to A, D and PP cells. Secondly, the antigen concerned is intracellular, and thus unlikely to be accessible to circulating ICA in vivo. The occurrence of ICA in spouses of patients (7-C-i-a) is another difficulty.

A more likely candidate is islet cell surface antibody (ICSA). This was probably first demonstrated by McLaren, Huang and Fogh (1975) using as substrate a culture of human insulinoma cells. Unfortunately this could not be confirmed because the cell line later lost its islet B-cell

characteristics. Later work using insulinoma cells has been inconclusive, perhaps because of variation in cell-surface antigens in these tumours (Pujol-Borrell et al, 1982).

A high frequency of ICSA at diagnosis of Type 1 diabetes has been confirmed by work using as antigen viable rat islet cells in suspension (Lernmark et al, 1978), monolayer cultures of neonatal rat islet cells (Dobersen et al, 1980), pure suspensions of rat islet B cells (Van de Winkel et al, 1981, and personal communication of work in press), and monolayer cultures of fetal human islet cells (Pujol-Borrell et al, 1982).

Lernmark et al (1981) have commented that prevalence of ICSA tends to parallel that of 'conventional' cytoplasmic ICA, but they agree with others working in this field that there is little if any quantitative correlation between the results of the two tests in the same individuals (op. cit; Dobersen et al, 1980). Thus it seems unlikely that the same antigen-antibody system is detected by both techniques.

Possible correlations between CF-ICA and ICSA have yet to be assessed. It is interesting that the former Ab often gives a patchy distribution of fluorescence within the substrate islet, perhaps reflecting different avidities for the various cell types present (Bottazzo et al, 1980a; Betterle et al, 1980). Furthermore, although the binding of CF-ICA is cytoplasmic in the dead, stored islets used, this does not preclude the possibility that the antigen concerned is the same as that binding ICSA, since the cytoplasmic antigen could also be expressed on the cell surface; this has been shown to be the case for thyroid cell antigens binding thyroid microsomal antibodies (Khoury et al, 1981).

7-C-i-d. A pathogenetic role for IC(S)A?

That ICA and ICSA commonly circulate in the blood of newly-diagnosed Type 1 diabetics is not in question. What is uncertain is whether these - or indeed other antibodies yet to be identified -play any part in pathogenesis, rather than being mere epiphenomena.

One approach is to study in vitro the effect of suitable sera on islet cells. Rittenhouse et al (1980) used Trypan Blue uptake to show substantial complement-mediated damage to isolated hamster islets in the presence of sera from some newly-diagnosed Type 1 diabetics, a cytotoxic effect which correlated with presence of ICA but which was considered unlikely to be specific for islet B cells.

Dobersen et al (1980, 1981) used ^{51}Cr-release and a double fluorescence technique to show lysis of neonatal rat islet cells by sera from ICSA-positive Type 1 diabetic patients. A more subtle effect on isolated mouse islets was demonstrated by Sai et al (1981). In presence of complement, the insulin-secretory response to glucose and theophylline was inhibited by sera from one of five newly-diagnosed Type 1 diabetics and from 10/16 Type 1 diabetics with additional OSAI, but not by control sera or those from Type 2 diabetic patients.

These recent developments are compatible with involvement of humoral autoimmunity in the islet B-cell destruction leading to Type 1 diabetes, but do not by any means prove the case. In vivo, at least in non-inflamed tissue, IgG is principally intravascular, so IC(S)A seems unlikely to initiate damage to previously unharmed islets. Furthermore the persistence of islet B-cell function after onset of clinical Type 1 diabetes, assessed by C-peptide studies, is unrelated to presence of cytoplasmic ICA (Tiengo et al, 1977; Madsbad et al, 1980). However, neither of these reservations yet applies to the cell-mediated immune mechanisms considered next, and the various branches of the immune system seldom act in isolation, so it is inherently likely that humoral mechanisms are involved.

7-C-ii. Cell-mediated immunity in Type 1 diabetes

7-C-ii-a. Leucocyte migration inhibition

Both organ-specific and non-specific abnormalities of cell-mediated immunity (CMI) have been described in Type 1 diabetes. The leukocyte migration inhibition test (LMIT) has given positive reactions with various pancreatic antigens in newly-diagnosed childhood cases (Nerup et

al, 1971; Nerup et al, 1974a; MacCuish et al, 1974b; Irvine et al, 1976b; Sensi & Pozzilli 1979). The positive reactions were unrelated to previous exogenous insulin therapy, and when insulin was used as antigen no inhibition of migration was seen. Furthermore, Sensi & Pozzilli (1979, plus personal communication of supplementary data) reported a positive pancreatic LMIT in 7 (16%) of 45 non-diabetic siblings HLA- or haplo-identical to Type 1 diabetic probands in the present study, compared with none of the 17 HLA-non-identical siblings tested. This supports the view that organ-specific cellular hypersensitivity to pancreas in Type 1 diabetes is not simply secondary to the metabolic changes.

It is generally accepted that a positive LMIT is the in-vitro equivalent of cell-mediated delayed (Type IV) hypersensitivity, and reflects release of migration inhibitory factor (MIF) in the presence of antigen by specifically sensitized T-lymphocytes, but Brostoff (1974) has cautioned that in some cases the presence of IgG antiglobulin may give positive results in the absence of specifically sensitized T-cells. Nevertheless, the LMIT has helped to focus attention on the possible role of these immunocytes in pathogenesis of Type 1 diabetes.

7-C-ii-b. Circulating immunocyte populations

Two rather crude and non-specific tests reflecting T-cell numbers and function are their total peripheral count, usually measured as spontaneous sheep-erythrocyte (E) rosette-forming cells (E-RFC), and phytohaemagglutinin transformation (PHAT) activity. Some workers using these tests have shown lower results in Type 1 diabetics than in controls (Cattaneo, Saibene & Pozza, 1976; Delespesse et al, 1974; Kadrnka-Lovrencic et al, 1978; Selam et al, 1979), probably because of the effect of poor metabolic control on T-cell function (Brodie & Merlie, 1970; MacCuish et al, 1974c; Delespesse et al, 1974; Di Mario et al, 1977; Selam et al, 1979; Cattaneo et al, 1980). In established diabetics with good control, PHAT and total E-RFC counts are near normal (Delespesse et al, 1974; MacCuish et al, 1974c; Hann, Kaye & Faulkner, 1976; Muller et al, 1980); although there may still be minor deviations, it is unlikely that in human Type 1 diabetes there exists the marked T-cell

lymphocytopenia found in BB rats before and after onset of diabetes (see below, 7-D-iii).

These rather unremarkable results do not imply that T-cells are unimportant in pathogenesis of diabetes. Thus total E-RFC counts and PHAT reflect numbers and activity of heterogeneous collections of lymphocyte subpopulations (e.g. helper, suppressor and natural killer T-cells; K-cells), which may vary independently and only some of which may be important in this context. Furthermore, even within T-lymphocyte subpopulations what really matters is not so much the total number of cells circulating in the blood as their specific reactivity in relation to islet B-cells in vivo. This is difficult to study in man, and not only for ethical and logistic reasons: specific T-lymphocyte cytotoxicity shows MHC-restriction, so ideally the target islet b-cells should have the same HLA type as the effector cells. Another relevant point is that most workers have concentrated on established diabetics, whereas the most important period is likely to be before diagnosis.

A simple way of distinguishing different T-lymphocyte sub-populations is by measuring spontaneous E-RFC under different experimental conditions. 'High-affinity' E-RFC constitute around 50% of peripheral blood mononuclear cells in healthy adults, and are thought to include the cytotoxic T-cells involved in target-specific CMI reactions. Muller et al (1980) investigated these cells and reported a small decrease (5% of total mononuclear cells) in 20 well-controlled Type 1 diabetics, relative to healthy controls. Many of the patients had long-standing disease, but Pozzilli and co-workers found the same in newly-diagnosed cases (see below).

'Low-affinity' E-RFC (normally up to 25% of circulating mononuclear cells) mainly represent killer lymphocytes or K-cells, which carry Fc receptors and mediate antibody-dependent cell-mediated cytotoxicity (ADCC) in vitro (West et al, 1977; West, Boozer & Herberman, 1978). The importance of K-cells in vivo is still somewhat controversial, but they may be involved in allograft rejection as well as the elimination of virus-infected and perhaps neoplastic cells. Furthermore, some findings

suggest a role in autoimmune disease, including thyroid OSAID (Calder et al, 1973, 1976; Feldman et al, 1976).

Pozzilli et al (1979, 1980) and Sensi et al (1981a, 1981b) measured peripheral blood E-RFC subpopulations in both longstanding and newly-diagnosed Type 1 diabetic children (mostly participants in the present study) as well as in healthy control subjects. The diabetics as a whole had slightly lower high-affinity E-RFC counts, but the more striking finding was in the newly-diagnosed children. These had significantly higher levels of low-affinity E-RFC, exceeding 2 standard deviations above the mean control level in over half. With increasing duration of diabetes the low-affinity E-RFC levels fell to normal. Raised levels of low-affinity E-RFC were associated with evidence of K-cell activity (ADCC) in vitro against two different targets.

These colleagues also investigated a small number of non-diabetic children from the present study who had been shown by HLA-genotyping to be potentially susceptible (see 6-C-ii; 6-F). Those who were ICA-positive had slightly but significantly higher levels of low-affinity E-RFC than controls; one has since become diabetic (Appendix 7-2, Case 2). These results suggest that K-cell numbers may be increased not only after but also perhaps before development of Type 1 diabetes.

Whether cell-mediated or complement-mediated, if antibody-dependent cytotoxicity directed against the pancreatic B-cells is indeed important in pathogenesis of Type 1 diabetes (rather than being merely a scavenging mechanism to remove defunct cells), then total peripheral suppressor T-lymphocyte counts in newly-diagnosed Type 1 diabetes (Buschard, Rygaard & Madsbad, 1980; Galluzzo et al, 1980) may perhaps reflect loss of pancreas-specific suppressor-cell clones, a possible mechanism by which organ-specific autoimmunity may arise.

7-C-ii-c. Target specificity of circulating immunocytes

Evidence from the LMIT suggests the existence of pancreas-specific CMI in diabetes (see above, 7-C-ii-a). Lymphocyte transformation by pancreatic antigens has also been described. MacCuish et al (1975)

reported positive results using insulin as antigen in five of 10 newly-diagnosed Type 1 diabetics (including 4 of 5 never treated), as well as in others already on treatment with exogenous insulin. Irvine et al (1976b) described comparable findings using insulinoma antigens. Similarly, Huang & Maclaren (1976) demonstrated immune adherence of lymphocytes from 18 of 23 Type 1 diabetic patients on cultured cells derived from a human insulinoma; cytotoxic effects, probably due to ADCC, were demonstrable.

More recently, Boitard et al (1981) showed that incubation of lymphocytes from newly-diagnosed diabetics with isolated mouse islet cells in vitro inhibited insulin release. It is not yet clear whether this represented a target-specific effect; in view of MHC restriction effects it seems unlikely that true T-cell cytotoxicity can have been responsible.

7-C-ii-d. Experimental work: passive transfer of diabetes?

In theory, if organ-specific CMI is responsible for Type 1 diabetes, it might be possible for donor immunocytes from a newly-diagnosed patient to induce diabetes in a recipient. However, without MHC identity, graft rejection and MHC-restriction of T-cell responses would normally prevent such an effect.

Buschard, Madsbad and Rygaard (1978) used as recipients nude mice, which are athymic and do not reject tissue grafts. They reported more hyperglycaemia in mice receiving large numbers of lymphocytes from human diabetic donors than in controls. They did not, however, examine the mouse islets histologically. Several other groups have attempted to reproduce these results and have failed (Lipsick et al, 1979; Serra et al, 1979; Thurneyssen et al, 1979; Sensi et al, 1982). On balance it appears that passive transfer of human Type 1 diabetes by immunocytes has not yet been achieved.

7-D. ANIMAL MODELS

7-D-i. Experimental immunization

Production of diabetes in animals by immunological procedures has a

surprisingly long history. In Brussels, de Meyer (1909) raised 'serum antipancreatique' by immunizing rabbits with an antigen prepared from dog pancreas, and used this to induce diabetes in dogs. The lesion was shown to affect the islets of Langerhans rather than the exocrine tissue (de Meyer, 1911). Insulitis is not apparent in the illustrations published.

Much more recent work, often involving active immunization, has been described in Chapter 3-C-iv; despite production of insulitis, rarely did anything like Type 1 diabetes result.

7-D-ii. Low-dose streptozotocin diabetes

The low-dose-streptozotocin-induced diabetes of mice and other rodents (Chapter 3-C-iv) appears to involve autoimmunity, in that it can be prevented by previous immunosuppression and depends on presence of thymic tissue, while passive transfer of the insulitis by spleen cells from affected mice to recipient mice has been demonstrated and is influenced by MHC genes (see Kiesel & Kolb, 1982).

Mouse islet cells affected by low-dose streptozotocin almost invariably contain large numbers of C-type virus particles, induced from viral DNA incorporated in the host genome and transmitted longitudinally from generation to generation (Like & Rossini, 1976). These authors suggested that modification of islet B-cell membrane antigen by viral antigen might underlie an autoimmune reaction.

7-D-iii. The BB rat

Of spontaneous animal diabetes, probably the closest model of human Type 1 diabetes is the ketotic form in BB rats. Nakhooda et al (1977) described many features in common with the human disease, including sudden onset of overt diabetes (often in adolescence), insulin dependence, a tendency to fatal ketoacidosis if not treated, and insulitis (see Chapter 3-C-iv).

Marliss et al (1982) have recently summarized evidence favouring an autoimmune pathogenesis. Perhaps most striking is the finding of a marked T-cell lymphopenia, probably affecting mainly helper T-cells,

which is present not only in rats already showing overt diabetes or milder hyperglycaemia, but also from the time of weaning in a proportion of apparently healthy littermates, some of which later develop carbohydrate intolerance (Bellgrau et al, 1982). Animals with lymphopenia also have insulitis.

Other evidence of autoimmunity includes passive transfer of insulitis (but not diabetes) by blood and spleen lymphocytes from BB rats to athymic mice; prevention or amelioration of the diabetes by neonatal thymectomy, by anti-rat-lymphocyte serum or by bone-marrow grafting from BB rats without lymphopenia (Marliss et al, 1982); detection of cell surface antibody, and also lymphocyte antibody, within the first few weeks after development of diabetes; frequent existence of lymphocytic thyroiditis or of B-cell lymphoproliferative conditions; occurrence of gastric parietal cell and thyroid colloid antibodies; and perhaps the susceptibility of these animals to infections, such that they are generally reared in a germ-free environment (which incidentally renders it unlikely that viruses are involved in pathogenesis, since vertical transmission of viral genome is not recognized in rats: Rossini et al, 1979).

Apart from the association with lymphoproliferative conditions and the severity of the T-lymphocyte defect, these features are very similar to those of human Type 1 diabetes. Further investigation of the BB rat, and establishment of an inbred strain in the next few years (Marliss et al, 1982), are awaited with interest.

7-E. CONCLUSION

The extensive literature reviewed here leaves little room for doubt that immune mechanisms are active during, and may even contribute to, pathogenesis of Type 1 diabetes. However, much is yet to be learned. Two unanswered questions seem fundamentally important.

First, what initiates the autoimmune reaction in the first place? Volpe (1978) has suggested in relation to Graves' disease that there

might be hereditary absence of a suppressor T-cell clone, or clones, specific for thyroid antigen(s) including the TSH receptor. Thyroid stimulating Ab would eventually appear in sufficient concentration to cause thyrotoxicosis only when a helper T-cell with TSH receptor specificity arose as a random event by somatic mutation. This intriguing idea parallels the theory that somatic mutation may contribute to the vast number of different clonal specificities of immunocytes expressed, for example, in the hypervariable region of immunoglobulins. However, it seems difficult to understand how a random event sufficiently rare to explain new appearance of thyroid stimulating Ab as late as middle age could at the same time be sufficiently common to provide a young baby with immunity against a host of different antigens. If autoimmunity in Type 1 diabetes does not arise at random, then presumably it is initiated by some environmental event, such as exposure to a virus or toxin.

The second question is related: is autoimmunity alone responsible for destruction of previously normal pancreatic B cells, or does it merely remove cells already affected by some other pathological process? If the latter then again an environmental factor may well be implicated.

PART III — CHAPTER 8

ENVIRONMENTAL FACTORS IN AETIOLOGY OF TYPE 1 DIABETES

CONTENTS

A Evidence that environmental factors are important · · · · · · · · 112
B Viruses and diabetes · 113
  i Animal work · 113
    a Encephalomyocarditis virus · · · · · · · · · · · · · · · · 114
    b Coxsackie B virus · 114
    c Other viruses · 115
  ii Viruses and human Type 1 diabetes · · · · · · · · · · · · · · 115
    a Mumps · 115
    b Coxsackie B · 117
    c Rubella · 118
    d Reo- and other viruses · · · · · · · · · · · · · · · · · · 118
C Toxic and nutritional factors · · · · · · · · · · · · · · · · · · 119
  i Energy and protein nutrition · · · · · · · · · · · · · · · · · 119
  ii Substances toxic to islet B-cells · · · · · · · · · · · · · · 119
    a Purine and pyrimidine derivatives · · · · · · · · · · · · · 120
    b Nitroso- compounds · 120
    c Other substances · 121
    d Possible role in human Type 1 diabetes · · · · · · · · · · 121
D Cumulative effects of different environmental agents · · · · · · · 122

8-A. EVIDENCE THAT ENVIRONMENTAL FACTORS ARE IMPORTANT

Since pairs of monozygotic twins having Type 1 diabetes are so often discordant for the disease, and generally remain so, it follows that genetic susceptibility alone is insufficient cause for diabetes, at least in the majority of cases (III-6-A). Thus an environmental factor (or, conceivably, a random event such as somatic mutation within the immune system; III-7-E) must also be involved.

Many of the epidemiological observations outlined in III-4 do appear to favour an environmental factor: geographical variation in incidence

(III-4-A-iv, III-4-B-iv), fluctuations and secular trends in incidence (III-4-B-ii), the distributions of age (III-4-A-v) and especially of season (III-4-B-i) at onset of clinical disease, and temporal clustering within sibships (III-4-B-ii). An obvious possibility is that infectious agents might be important in this respect.

Much of the evidence pertaining to human diabetes is circumstantial, and links environmental events not with the initiation of the disease process itself but with a later consequence, i.e. the onset of metabolic decompensation. The longer the latent period between these two stages of the disease, the less likely is it that any association between environmental agent and clinical diabetes is due to a true causal relationship, and the more plausible becomes the alternative interpretation: in a 'pre-diabetic' individual in whom the majority of islet B cells have been destroyed there might remain just sufficient insulin secretory capacity to maintain glucose homeostasis under normal conditions, but insufficient for the increased insulin requirement resulting from stresses such as incidental infections, emotional trauma or even dietary indiscretions. Thus an environmental event might indeed lead to onset of clinical diabetes without in any way causing the disease. The length of the pre-diabetic latent period is therefore of considerable importance (Chapter 9).

8-B. VIRUS INFECTIONS AND TYPE 1 DIABETES

A widely-held hypothesis suggests that viral infection of islet B cells occurs as a primary event and leads on to an immune response (cf. III-7) directed against virally-modified or unmodified B-cell surface antigens. Is there any evidence specifically implicating viruses in this role?

8-B-i. Animal models involving viruses

Viral diabetes in animals is now well recognized. Twenty years ago it was reported that a proportion of cattle surviving naturally-acquired or experimental foot-and-mouth disease develop insulitis and diabetes (Barboni & Manocchio, 1962; Barboni, Manocchio & Asdruballi, 1966).

8-B-i-a. Encephalomyocarditis virus

A more practical animal model was described by Craighead & McLane (1968). The D or M variants of wild encephalomyocarditis virus (EMCV), a picornavirus like that of foot-and-mouth disease, exhibit diabetogenic properties in some strains of mice and retain these properties after repeated passage in mice or mouse tissue cultures. After inoculation the virus multiplies in islet B cells, leading within a few days to degranulation and lysis associated with a rapid decline in pancreatic insulin content, resulting in hyperglycaemia of varying severity. Host susceptibility to EMC virus is influenced by more than one gene, probably recessive and not MHC-linked, and is enhanced by male sex hormones, by obesity, and by unidentified environmental factors. There is a transient mononuclear cell insulitis, and autoimmunity appears to be involved in that previously-irradiated mice, like athymic *nu/nu* mice, are resistant to the diabetogenic effects (Ross et al, 1976; Craighead, 1981).

8-B-i-b. Coxsackie B virus.

Another extensively studied animal model is diabetes induced by a common human picornavirus rather similar to EMCV, Coxsackie B virus. Robertson (1954) observed that mice infected with some strains commonly developed islet fibrosis and occasionally became hyperglycaemic. Later it was reported that when susceptible strains of mice were infected with Coxsackie B4 virus, blood glucose levels rose after an initial fall, reaching hyperglycaemic levels in 20-30% of cases. In only one mouse did frank diabetes persist beyond 10 weeks. Insulitis was present in those animals with impaired glucose tolerance but not in controls. The diabetogenic effect of the virus was markedly affected by both host factors (strain and sex of mouse) and strain of the virus; adaptation by passage in pancreatic tissue enhanced diabetogenesis, while heart passage weakened the effect (Coleman, Taylor & Gamble, 1974). Further work, much of it by A L Notkins and associates in Bethesda, has emphasized the importance of genetic strain of the host mouse, and has shown that all six Coxsackie B variants may be diabetogenic (Toniolo et al, 1982).

8-B-i-c. Other viruses

Several other human pathogens can produce islet cell damage in animals. In 1974, Rayfield and colleagues reported such an effect in hamsters and primates infected with the virus of Venezuelan encephalitis. The hyperglycaemia produced is usually mild and transient, and seems to involve a rather subtle abnormality of insulin release (Rayfield et al, 1981). In rabbits, congenital infection with rubella virus may lead to islet damage similar to that produced by EMC virus in mice (Menser et al, 1978). In immunosuppressed mice, murine strains of cytomegalovirus are said to infect islet cells (Craighead, 1981). Also in mice, some types of reovirus exhibit tropism for the islets of Langerhans, infecting not only B but also A and D cells, and leading to insulitis. Interestingly, islet cell antibody has been identified in serum from affected animals (Onodera et al, 1983).

Mention has already been made of the possible role of C-type virus in low-dose streptozotocin diabetes (III-3-C-iv; III-7-D-ii).

8-B-ii. Viruses and human Type 1 diabetes

Müller (1822) cited six earlier works implicating smallpox ('variolae') as a causative factor in diabetes, without giving any details. Since then anecdotal accounts have linked onset of diabetes with preceding illnesses now known to be viral, but in most cases it can be argued that there was only a chance association or a non-specific metabolic effect.

Several viruses have in fact been shown to infect human islet B-cells in vivo or in vitro, including Coxsackie B (Gladisch, Hoffmann & Waldherr, 1976; Yoon et al, 1979; Jenson et al, 1980), mumps (Prince et al, 1978), cytomegalovirus and rubella (Jenson et al, 1980), and reovirus (Yoon et al, 1981). As yet, however, none has been incontrovertibly implicated as a direct cause of Type 1 diabetes in man.

8-B-ii-a. Mumps.

Since 1899 there have been repeated anecdotal case reports of diabetes following mumps in childhood (Gamble, 1980b; Craighead, 1981).

Epidemiological evidence began in 1927 when Gundersen described increases in incidence of diabetes after outbreaks of mumps in isolated villages. The incidence data of Sultz et al (1975; see III-4-B-iii) were interpreted in terms of an aetiological role for mumps with a latency of about 4 years. Half of the 300 diabetics interviewed gave a history of exposure to mumps, but there are no control figures. More recently Laporte et al (1981) failed to show any change in incidence of Type 1 diabetes after the widespread adoption of mumps vaccination in the USA in the 1970's. It is of course possible that the vaccine and the wild virus might both be diabetogenic.

Using a different epidemiological approach in UK, Gamble (1980b) investigated nearly 4000 cases of childhood diabetes by questionnaire and, by comparison with a large control population, did demonstrate an excess of clinical mumps reported to have occurred in the year before onset of the diabetes. Although statistically significant and not explicable by biased reporting, the demonstrable excess of mumps represented only about 1% of all the diabetics. This does not exclude a role for mumps in other cases since infection is often subclinical, and the latent period before diagnosis may have exceeded 1 year (see III-9), but epidemiological evidence implicating this virus as a *major* factor in Type 1 diabetes is still lacking.

Evidence of other types is scanty. Exocrine pancreatitis is a well-recognized complication of mumps; but despite experimental infection of human islet B-cells in vitro (Prince et al, 1978), insulitis has not been reported (Craighead, 1981). An intriguing finding reported from Germany is that 14/30 children with recent complicated mumps had islet cell antibody in the serum, often complement-fixing. None developed diabetes while under prospective observation for 18 months (Helmke et al, 1980). The implication of this work is not clear; apparently, the high incidence of islet cell antibody after mumps was not seen again in later studies by the same group (reported by Kolb, Schernthaner & Gries, 1983)

8-B-ii-b. Coxsackie B virus.

Gamble et al (1969a, 1973) showed a greater prevalence of neutralizing antibody to Coxsackie B4 in children with recently-diagnosed acute diabetes (70%) than in controls (58%), although acknowledging that the latter were not matched geographically. Antibody titres were generally below the levels to be expected after recent infection. The excess frequency of antibody was confined to younger children (age <10y).

Interpretation of epidemiological studies based on Coxsackie B serology is difficult. Hierholzer & Farris (1974) failed to detect any increase in diabetes incidence after an outbreak of Coxsackie B3 and B4 infection, but this was in a population whose susceptibility to Type 1 diabetes was likely may from that in N. Europe. Sakurami et al (1982) reported higher frequency of Coxsackie B3 and B4 neutralizing antibody titres exceeding 80 in Type 1 diabetic children at summer camp in Japan than in controls; here, bias was possible in that control children were excluded if they had any history of recent infections. Palmer et al (1982) found *lower* titres to Coxsackie B3 and B4 in new cases of Type 1 diabetes during an outbreak of infection than in non-diabetics at the same time and place, suggesting a possible impairment of immune response in the former.

There have been several reports associating acute childhood Type 1 diabetes with individual cases of systemic Coxsackie B infection (Gibbs, 1974; Gladisch et al, 1976; Champsaur et al, 1980; Phillips & Pauli 1981). Perhaps the best-known and best-documented case was that described by Yoon et al (1979), a 10-year-old boy with fulminating Coxsackie B4 infection who developed diabetic ketoacidosis several days before dying some 10 days after onset of the illness. Not only was specific antibody present in the post-acute serum, but in addition B4 viral antigen was demonstrated by immunofluorescence in the pancreatic islets, and the isolated virus was shown to be diabetogenic in susceptible strains of mice (see III-8-B-i-b).

It is commonly held that such cases as these might represent a fulminating, 'viral' form of Type 1 diabetes, distinct from the majority.

However, sera from several of them have been tested for, and found to contain, CF-ICA (see III-7-C-i-b; Bottazzo, Minkowski Lecture 1982, published in revised form as Bottazzo, 1984) while the case of Yoon et al (1979), despite its extremely acute clinical course and the very widespread insulitis in the pancreas, did have a number of fully-developed 'pseudoatrophic' islets of Langerhans (personal observation by the author in the laboratory of Prof. W Gepts in 1982, confirmed by the latter). This, together with the finding of ICA (Bottazzo, op.cit.) suggests pre-existing 'Type 1 diabetic' islet cell damage. Thus it appears likely that such tragic cases, though atypical, do have much in common with the general run of Type 1 diabetes.

8-B-ii-c. Rubella.

In a long-term, prospective study of children affected by congenital rubella infection, Menser et al (1974, 1978) noted a very high incidence of glucose intolerance and frank diabetes during the first three decades of life. The relevance to Type 1 diabetes is not clear in that many of the patients were not insulin-dependent. Smithells et al (1978) found diabetes in only 2 of 482 children with congenital rubella, the disparity with Menser's figures perhaps being excplicable by the shorter followup period, a maximum of 3.5 years. Patterson, Chandra and Jenson (1981) described pancreatic histology in an infant who had died of diabetic ketoacidosis after congenital rubella: no insulin-positive B-cells were found, either within islets or as isolated nesidioblasts, and patchy insulitis was observed. That the virus could infect human islet B cells in vivo had already been shown (Jenson et al 1980), and it was concluded that in this child the virus had probably destroyed the B-cells in utero.

8-B-ii-d. Other viruses.

Reovirus is another common human pathogen capable of infecting human islet B cells (Yoon et al (1981). Animal work is mentioned above (III-8-B-i).

The extensive cytopathogenic spectrum of fatal human cytomegalovirus infection commonly includes pancreatic islets (Jenson et al, 1980). Both this virus and clinical infectious mononucleosis (sometimes due to cyto-

megalovirus, though usually attributable to the similar Epstein-Barr virus) have been anecdotally associated with diabetes mellitus (Craighead, 1981).

An epidemic of infectious hepatitis after civil war in Nigeria was associated with an outbreak of diabetes. This was not typical of Type 1 diabetes in Europeans since, even though initially insulin-dependent, most cases remitted spontaneously or followed a fluctuating course (Adi, 1974; Oli & Nwokolo, 1979).

8-C. TOXIC AND NUTRITIONAL FACTORS.

8-C-i. Energy and protein nutrition.

Obesity is described as the strongest risk factor for adult-onset diabetes in a community (WHO, 1980), and nutritional factors are almost certainly important in the high incidence of Type 2 diabetes in Pima Indians and in Nauru (Zimmet, 1982), as well as the low incidence in the Netherlands in the 1940's and in Paris, 1870 (Vinke et al, 1959).

Tropical malnutrition diabetes may be more relevant to Type 1 diabetes. McMillan & Geevarghese (1979) proposed a complex nutritional aetiology for this insulin-dependent condition, suggesting that inadequate intake of sulphur-containing amino acids in the diet impairs the ability to detoxify small amounts of cyanide released from cyanogens in the staple food of cassava, and that the cyanide causes pancreatic damage. The pancreatic lesion is generalized rather than B-cell specific, and this particular mechanism is clearly not important in Type 1 diabetes. Nevertheless it illustrates that nutritional status may modify toxic effects of environmental agents.

8-C-ii. Substances toxic to islet B-cells

Several chemical substances have been shown to impair glucose tolerance in man or animals. Some, such as cyproheptadine, thiazides and L-asparaginase, do not appear to act by impairing insulin release, and are not relevant here (Dulin & Soret, 1977). Other toxic substances are known to damage islet B cells.

8-C-ii-a. Purine and pyrimidine derivatives

Alloxan is a ring compound with oxidizing properties, derived from pyrimidine. When administered to experimental animals it leads to a brief initial hyperglycaemia followed by a few hours of hypoglycaemia as the islet B cells become degranulated, and finally by onset of diabetes with B-cell necrosis (Shaw Dunn et al, 1943a, 1943b (citing Jacobs' observation of hypoglycaemia due to alloxan in 1937); Shaw Dunn et al, 1943c). Insulitis is not mentioned in descriptions of alloxan diabetes.

Various authors have sought to implicate similar compounds in spontaneous human diabetes, so far without convincing success (Dulin et al, 1983). However on a number of occasions when men have been poisoned by a rodenticide, N-3-pyridylmethyl N-p-nitrophenyl (PNU or Vacor), insulin-dependent diabetes and autonomic neuropathy have resulted (Miller, Stokes & Silpipat, 1978). Histologically there is degeneration of islet B cells but no insulitis in the few cases studied post mortem; it is interesting that two of six surviving cases tested several months after exposure were positive for islet cell antibody (Karam et al, 1980).

8-C-ii-b. Nitroso- compounds.

Streptozotocin, an antibiotic from *Streptomyces achromogenes*, is the N-nitroso-N-methylurea derivative of 2-deoxyglucose. Rakieten et al (1963) administered single intravenous doses to animals and observed a triphasic glycaemic response similar to that after alloxan, the usual end result being permanent insulin-dependent diabetes with damage confined to islet B cells (Dulin & Soret, 1977). Such single doses of streptozotocin do not lead to insulitis, but repeated, individually sub-diabetogenic, doses may do so, leading to diabetes after about a week (III-3-C-iv; III-7-D-ii). Response is greatly affected by genetic factors.

There is some circumstantial evidence compatible with an aetiological role for nitroso- compounds in human Type 1 diabetes. Helgason & Jonasson (1981) noted an excess of early October birthdays among Type 1 diabetic boys in Iceland. Even when allowance is made for inappropriate statistical analysis this excess still appears significant.

The authors postulated that maternal consumption in January of smoke/cured mutton left over from Christmas might have affected the embryonic B cells early in gestation. High levels of nitrosamines were present in samples of the food, and Helgason et al (1982) later demonstrated diabetogenic properties of Icelandic smoked mutton in male (much more than female) mice exposed in utero (via the mother) and on weaning. There was a suggestion of a similar, though incomplete, effect when the meat was given to the parents only up to the moment of conception, implying an effect on the germ cell line.

This hypothesis is still tenuous, but it is not excluded by the lack of any demonstrable association between Type 1 diabetes and month of birth in the diabetic population which forms the basis of the present study (Spencer et al, 1982b).

8-C-ii-c. Other substances

Some quinoline derivatives (oxine, dithizone, 8'hydroxy-quinoline) have pancreatic B-cell toxic effects in experimental animals (Dulin et al, 1983), as does cyanide (McMillan & Geevarghese, 1979).

8-C-ii-d. Possible role of toxic substances in human Type 1 diabetes.

With the growth of the chemical industry, human exposure to synthetic chemical substances must have multiplied enormously during the present century, especially in industrialized countries. Had there been a proportionate rise in incidence of Type 1 diabetes over the same period there would be no room for the doubt expressed in III-4-B-iii over existence of any such change. Thus it seems highly unlikely that synthetic chemicals which are not widely distributed in nature should be important in aetiology.

It is, however, particularly interesting that two of the toxic substances implicated (if only rarely) in human insulin-deficient diabetes, viz. dietary cyanogens and nitrosamines, are respectively related to malnutrition and food spoilage, while streptozotocin (see above) is a naturally-occurring antibiotic. Presumably they are at least

as common in primitive or developing societies as in industrially advanced communities. Similar products of fungal or bacterial food spoilage might well be one factor in pathogenesis of Type 1 diabetes, either by contributing non-specifically to islet B-cell damage or, conceivably, by a more subtle and specific (immunological?) effect in genetically susceptible people. Retrospective identification of such effects will be especially difficult if the latency between exposure and clinical onset of disease is very long (III-9).

8-D. CUMULATIVE EFFECTS OF DIFFERENT ENVIRONMENTAL AGENTS.

It is obvious that, if two agents can each destroy a proportion of islet B-cells by different mechanisms, then their combined action will probably lead to more complete destruction. Alternatively, the mechanism by which one agent destroys individual B-cells might depend on another agent. The possible role of C-type virus in low-dose streptozotocin diabetes may be an example (III-3-C-iv; III-7-D-ii). In either case the two agents could be said to interact to produce diabetes.

In mice, induction of diabetes by EMC virus is affected by male sex hormones, obesity and unidentified environmental factors (III-8-B-i). Toniolo et al (1980) demonstrated that doses of streptozotocin which were not diabetogenic when given alone did lead to diabetes in mice pre-infected with strains of EMC and certain other viruses, even though the strain of mouse used was not otherwise susceptible to the diabetogenic effect of the virus. Similar interactions may occur in man, leading to one kind of heterogeneity within Type 1 diabetes and making the search for individual environmental factors even more difficult than it would otherwise be.

PART III — CHAPTER 9

THE PRE-DIABETIC PHASE OF TYPE 1 DIABETES

CONTENTS

A Introduction · 123
B Duration of the pre-diabetic phase · · · · · · · · · · · · · · · 124
i Duration of symptoms before diagnosis · · · · · · · · · · · 124
ii Duration of disease process before symptoms · · · · · · · 124
a Retrospective evidence · · · · · · · · · · · · · · · · 124
b Prospective evidence in man · · · · · · · · · · · · · 125
c Animal work on duration of latency · · · · · · · · · 126
C Markers of pre-diabetes · 126
D Implications · 126

9-A. INTRODUCTION

Historically, the term 'pre-diabetes' was used to describe cases where abnormal carbohydrate tolerance could be demonstrated under special circumstances (such as provocation by glucocorticoid administration, or pregnancy), while under normal conditions diagnostic criteria for diabetes were not met (Conn 1958; West 1960; Jackson 1960). In the present context, 'pre-diabetes' refers to the state in which the pathological changes characteristic of Type 1 diabetes exist and, as yet, the metabolic (or indeed clinical) features of diabetes have not developed. Until now very little was known about the duration of the pre-diabetic phase in childhood Type 1 diabetes or the reliability of islet-cell antibody as a marker of pre-diabetes in individual children (III-7-C-i).

9-B. DURATION OF THE PRE-DIABETIC PHASE.

Traditionally, Type 1 diabetes was said to begin rather dramatically and to progress rapidly from the first symptoms to severe metabolic decompensation. By backward extrapolation the time-course of pre-clinical islet-B-cell destruction seemed likely to be short, especially with onset of disease in infancy. There was, however, little firm evidence on the duration of symptoms before diagnosis, and even less on the pre-symptomatic course of the disease.

9-B-i. Duration of symptoms before diagnosis.

A few retrospective reports in fact suggested that symptoms began more than 1 month before diagnosis in about half of childhood cases, while occasionally the interval exceeded one year (Hamilton, Mundia & Lister, 1976; Cudworth et al 1977). There are also isolated reports of transient symptoms suggesting diabetes, or even of transient carbohydrate intolerance confirmed biochemically, followed by over a year without symptoms before apparently acute onset of Type 1 diabetes (Weichselbaum 1910 - see III-3-C-i, Table III-2A; Ilonen et al 1980).

9-B-ii. Duration of disease process before symptoms.

9-B-ii-a. Retrospective evidence

The theory that seasonality of viral infections explains seasonal variation in incidence of Type 1 diabetes (III-4-B-i), and that the same infections initiate pathogenesis, presupposes a short pre-diabetic phase. A similar argument applies to the excess of recent mumps reported in a minority of newly-diagnosed diabetic children (III-8-B-ii-a). On the other hand, pancreatic histology suggested to Gepts (1965) that the time-course was rather prolonged (III-3-C-iii). In that case viral infection might merely trigger the final stage in a much longer process, a good case in point being the child with apparently fulminating viral diabetes whose pancreas yet contained established pseudo-atrophic islets (III-8-B-ii-b).

The hint that a mumps epidemic might have been causally related to an increase in incidence of Type 1 diabetes four years later (III-4-B-iii; III-8-B-ii) is so circumstantial as to carry little weight here. A latent period of about 20 months in an infant is suggested by the diabetes diagnosed in a child affected by congenital rubella (III-8-B-ii-c), and much longer latency may be implied by the onset of diabetes in adult life in congenital rubella patients, if relevant to Type 1 diabetes (III-8-B-ii-c).

An intriguing observation of Steel et al (1979) also suggests a pre-diabetic phase of many years: among parous women diagnosed as having 'typically insulin-dependent' diabetes, the recalled birth-weights of more than 20% of babies born more than 3 years previously exceeded the 90th centile, an excess persisting even up to twenty or more years between delivery and clinical diabetes. Criteria for insulin dependence are not given, nor are other features of Type 1 diabetes recorded (family history, islet-cell antibody, HLA studies), so extrapolation to Type 1 diabetes in younger patients must be guarded.

It may be relevant here that, of ten Nigerian patients whose (presumably) post-hepatitis diabetes resolved and who were followed up, four again became diabetic after 1.5-4 years, three remained well after at least 4 years, and three were lost to follow-up after 1 year (III-8-B-ii-d).

9-B-ii-b. Prospective evidence in man.

If islet-cell antibody is a valid marker of the pre-diabetic phase of Type 1 diabetes (see below), then in adults with other autoimmune endocrine diseases the duration of this phase may reach 8 years or more before onset of (Type 1B) diabetes, according to Irvine et al (1976). In the more typical, childhood (Type 1A) diabetes, the single case of Ilonen et al (1980) has been mentioned. Otherwise, prospective information was lacking until reports based on the present study (Gorsuch et al, 1980, 1981b, Appendix 7-3). Subsequent work by others, confirming long latency, is mentioned in RESULTS, VIII-3-C.

9-B-ii-c. Animal work on duration of latency.

Laboratory animal models of Type 1 diabetes can have little relevance in this respect, since a very long latency would preclude experimental study, even supposing a sufficiently long-lived animal were chosen. The BB rat model (III-7-D-iii) shows abnormalities (lymphopenia) even at the time of weaning, while not developing diabetes until adolescence; in relation to a rat's lifespan this might be considered a long latency, but it only amounts to a few weeks. Barboni et & Manocchio (1962) mention a delay of only two months between foot-and-mouth disease and ensuing diabetes in cows.

9-C. MARKERS OF PRE-DIABETES.

Islet-cell antibody is clearly a useful marker of pre-diabetes, although its predictive value remains quantitatively uncertain, as discussed towards the end of III-7-C-i-a. No metabolic abnormality has been shown to occur early enough and consistently enough in the pre-diabetic phase for use as a marker, nor is it clear whether very minor differences in glucose homeostasis between 'susceptible' and 'non-susceptible' groups of siblings of Type 1 diabetic children (Ginsberg-Fellner et al, 1982; Orchard et al, 1986) might be genetically determined (see III-11-D-iii) or alternatively might reflect islet B-cell damage during the pre-diabetic phase of pathogenesis

9-D. IMPLICATIONS.

At the beginning of the Barts-Windsor Family Study (1978) a long pre-diabetic phase - years rather than months - before onset of symptoms in childhood Type 1 diabetes seemed unlikely, except in occasional cases. There was thus a great need for a prospective, large-scale study of clinically normal children, investigated for presence of a marker such as islet-cell antibody, and followed up for long periods. This would not only clarify the duration of the pre-diabetic phase, but would also allow estimation of the risk of future diabetes associated with presence of the marker.

PART III - CHAPTER 10

HETEROGENEITY WITHIN TYPE 1 DIABETES

CONTENTS

A Introduction · 127
B Analysis for heterogeneity · 128
  i Criteria for heterogeneity · · · · · · · · · · · · · · · · · · · 128
  ii Common sources of error in heterogeneity analysis · · · · · · 128
    a Multiple comparisons and Type 1 errors · · · · · · · · · · 128
    b Dependent variables; bias; errors of interpretation. · 129
  iii Methods of testing for heterogeneity · · · · · · · · · · · · · 129
    a Multiple two-way comparisons · · · · · · · · · · · · · · · 129
    b Multi-way comparisons · · · · · · · · · · · · · · · · · · · 130
C Published evidence of heterogeneity in Type 1 diabetes · · · · · · 131
  i Histology and age at onset · · · · · · · · · · · · · · · · · · · 131
  ii Season at onset of symptoms versus age · · · · · · · · · · · · · 131
  iii Genetic and immunological heterogeneity · · · · · · · · · · · · 132
    a Familial versus sporadic Type 1 diabetes · · · · · · · · · 133
    b Genetic markers versus age at onset of Type 1 diabetes 133
    c Type 1B diabetes · 133
    d Islet-cell antibody in the very young patient · · · · · 135
  iv Heterogeneity with respect to environmental factors · · · · 135
D Conclusion · 136

10-A. INTRODUCTION

> '***Heterogeneous**. 1. Diverse in nature; of completely different characters... 2. Composed of diverse elements or constituents; not homogeneous...*' (Shorter Oxford English Dictionary)

That primary diabetes mellitus is heterogeneous is not in doubt: Type 1 diabetes is distinct from Type 2 diabetes in many different respects (III-1, III-5). The question here is whether Type 1 diabetes itself is 'composed of diverse elements or constituents', which is not necessarily the same as simply showing variability of some clinical, epidemiological or other feature.

Many claims of heterogeneity in Type 1 diabetes occur in the literature. Few stand up to rigorous criticism. This chapter first outlines (from a basis of general statistical principles rather than individual references) the particular problems inherent to the type of analysis involved, and then brings together the evidence reported, much of it already mentioned.

10-B.ANALYSIS FOR HETEROGENEITY.

10-B-i. Criteria for heterogeneity

Evidence of heterogeneity requires more than mere demonstration that some attribute (HLA-B15, for example) is lacking in one subgroup of patients but present in another, or that a quantitative variable differs in magnitude between the two (e.g. age at onset of diabetes). It must also be shown that the subgroups defined by virtue of one attribute or variable also differ in some other way, that is they are non-randomly distributed with respect to a second variable. This may be tested statistically against the null hypothesis that distribution is random. A significant result indicates that heterogeneity is likely, provided the two variables are inherently independent of each other and relevant to the diabetes.

For example, if it could be shown that early age at onset were a distinctive and characteristic feature of diabetes in HLA-B15-positive, as opposed to B15-negative, subjects, and if that association could not be explained by various common errors (see below), then it would be reasonable to conclude that heterogeneity existed. Even then, confirmation would usually be necessary in a fresh sample of patients, and differences in respect of other attributes would be sought.

10-B-ii. Common sources of error in heterogeneity analysis

10-B-ii-a. Multiple comparisons and Type 1 errors

The search for heterogeneity frequently involves making many comparisons. It is a fundamental principle of statistics that the larger the number of comparisons made, the greater is the probability that one or more will give a 'significant' result, even though in the population

sampled no deviation from random in fact exists. Such 'Type 1 errors' are in the real world more likely to be detected (and published) than 'negative' findings, unless the analysis is rigorous and designed to minimize this type of error. Probably for this more than any other reason, many examples of apparent heterogeneity reported in the literature fail to stand up to further investigation in enlarged samples from the same population or in different populations.

10-B-ii-b. Dependent variables; bias; errors of interpretation.

It might be shown, for example, that Type 1 diabetic children with onset before age 5 are less frequently positive for varicella viral antibodies at the time of diagnosis than those older at onset. This would not necessarily say anything about the diabetes, since frequency of varicella antibodies is related to age in any group of children; the group selected by virtue of one attribute (low age at onset) would be biased in respect of the other because the variables concerned are dependent. Only if the association between age and varicella antibody frequency in diabetic children were significantly different from that in non-diabetic children (i.e. a second-order interaction, one variable influencing the interaction between the other two variables) would the finding be relevant to diabetes at all. Even then, to assume that this indicated aetiological heterogeneity of diabetes itself ('age' affecting the interaction between 'varicella antibody' and 'diabetes') might be an error of interpretation; an effect of diabetes on host responses to varicella, or on exposure to the virus, would also have to be considered.

10-B-iii. Methods of testing for heterogeneity.

10-B-iii-a. Multiple two-way comparisons.

Many reports of heterogeneity are based on apparently 'significant' two-way comparisons (e.g. presence or absence of HLA-B15 versus age at onset) within samples of diabetic patients. This has the merit of simplicity, and is appropriate when seeking confirmation, in a new sample, of heterogeneity already indicated by analysis of another, in which case conventional tests of significance apply. However, when the aim is to identify new types of heterogeneity, selective reporting of two-way analyses can be misleading because of type 1 errors. Correction

of 'p-values' by a factor equal to the number of comparisons made reduces the risk of such errors, provided the comparisons to be made are either exhaustive or selected a priori (rather than a posteriori because the data 'look interesting'!).

A disadvantage of such correction of p-values is that it may hide heterogeneity which does exist in the population sampled, particularly where more than one of a set of individual comparisons gives results significant at the 5% level. It may therefore be better to report results of all comparisons, accepting as significant only those giving a very low p-value (<0.001); borderline results are then available for further investigation in fresh samples. A further disadvantage is that second-order interactions (see above) indicating genuine heterogeneity may go undetected.

10-B-iii-b. Multi-way comparisons

In order to avoid errors due to multiple comparisons it is in general preferable to perform initially a single, multi-dimensional analysis, for example by categorizing a sample of patients simultaneously with respect to HLA status (using mutually exclusive and exhaustive classes), age at onset of diabetes, sex, and some measure of organ-specific immunity such as thyroid antibody status. Any variable which then proves to be independent of the others in the sample (i.e. distribution is random with respect to it) may then be eliminated by 'collapsing' the array of data over that variable, simplifying further analysis - a technique particularly useful where data from several different populations or centres are to be analysed, since if distribution is independent of the variable 'centre' then the sets of data may simply be pooled. Furthermore, log-linear modelling techniques permit detection of second- (and higher-) order interactions (see above) (Everitt, 1977; see also METHODS, VI-3-C).

Very few published studies adopt this approach, probably for three reasons: its conceptual complexity, the difficulty of computation, and especially the great subdividion of samples it entails, necessitating

large samples in the first place if the proportion of 'empty cells' in the array is to be kept reasonably small.

10-C. PUBLISHED EVIDENCE OF HETEROGENEITY IN TYPE 1 DIABETES

10-C-i. Histology and age at onset

A significant (inverse) relationship between frequency of insulitis and age has been demonstrated, but the implications of this are unclear - possibly nothing more than a slower progression of islet B-cell destruction in adolescents than in young children (III-3-C-i). The suggestion that clinical Type 1 diabetes beginning beyond age 30 years may be histologically quite distinct from that in younger subjects is interesting but unconfirmed (III-3-C-i).

10-C-ii. Season at onset of symptoms versus age

The seasonal onset of symptomatic Type 1 diabetes generally found in children of school age, but not in younger children, has been mentioned (III-4-B-i). Such variation does not necessarily indicate heterogeneity of the Type 1 diabetes disease process itself. Thus, timing of onset of symptoms may be related only loosely, if at all, to initiation of pathogenesis, seasonality then being a non-specific effect (III-4-B-1; III-8-A). In very young children the pre-diabetic phase is likely to be shorter than in others (III-9), and the relative lack of seasonality in them may simply be because the final stages of development of metabolic abnormality are temporally more closely linked to a non-seasonal initiating event, and thereby less likely to be influenced by non-specific environmental factors, than in older children.

Furthermore there is a curious geographical inconsistency in seasonality differences with age (III-4-B-i) which would be unlikely if they reflected fundamental differences in pathogenesis. On the other hand, if (as is likely) any of a number of different environmental factors are capable of triggering symptomatic carbohydrate intolerance once the majority of islet B-cells had been destroyed, if some but not all of these factors are seasonal, and if there is differential exposure to seasonal and non-seasonal factors with respect to age, then geo-

graphical variation in the prevailing patterns of such factors could easily explain such inconsistencies.

Nevertheless the possibility remains that different types of genetic susceptibility interact specifically with different groups of environmental factors to produce B-cell damage, and that susceptibility to such factors varies with age.

10-C-iii. Genetic and immunological heterogeneity.

Some authors term as genetic heterogeneity the contribution of more than one gene to disease susceptibility, as stated clearly by Barbosa & Rich (1984). That usage unnecessarily and confusingly includes the concept of bi- or multifactorial (as opposed to single-gene) inheritance of an homogeneous entity of susceptibility - homogeneous, that is, in qualitative if not in quantitative terms (cf. III-5-C). The meaning adopted here is more precise, and accords with the dictionary definition at the head of this chapter: genetic susceptibility to Type 1 diabetes is heterogeneous if it is 'composed of diverse elements or constituents' in that different genes, or different sets of interacting genes, may each determine susceptibility, presumably by different mechanisms.

Thus it is not necessary to invoke heterogeneity to explain the fact that some Type 1 diabetic children type positive for HLA-DR3 but not DR4, some for DR4 but not DR3, many for DR3 and DR4 together, and a few for neither; even by the simplest hypothesis that the same two susceptibility genes underlie all cases (III-6-E-iii), incomplete linkage disequilibrium between each of the two putative susceptibility genes and the corresponding DR genes will have the same result. On the other hand, if the HLA-region gene-products themselves contribute to susceptibility (III-6-E-v, III-11-E), then there must be considerable heterogeneity, which may well be reflected in clinical, epidemiological, histological or immunological variation among Type 1 diabetic patients.

The problem of Type 1 errors (see above) is especially troublesome here, because of the very large number of statistical comparisons which may be made in relation to a highly polymorphic system such as HLA. This

probably explains the controversy surrounding many aspects of possible genetic heterogeneity.

10-C-iii-a. Familial versus sporadic Type 1 diabetes.

This has been considered in detail (III-6-E-i); there is no convincing evidence for heterogeneity, but further studies are needed, remembering that the greater the number of individuals affected, the more likely are multiple copies of susceptibility genes to occur within families; patterns of HLA genotypes might vary on that account. Greenberg & Anderson (1983) did, however, recently show compatibility of pooled observations from several studies (excluding the Barts-Windsor study) with the hypothesis that 'simplex families represent primarily a recessive form of Type 1 diabetes while multiplex families represent two other subforms: one showing three-allele inheritance/and the other not'.

10-C-iii-b. Genetic markers versus age at onset of Type 1 diabetes

Early suggestions that in Type 1 diabetes early age at onset was asssociated with HLA-B15 were not confirmed by extended studies, but remain controversial (III-6-E-i). Further analyses on large samples would still be relevant. There has been rather more agreement about a similar association of early onset with Bf-F1 (III-6-D), but again further confirmation on a large sample is needed. Combined analysis of published data from different centres, if correctly done, may help overcome the problem of small numbers of observations, but remains open to the problem of biased reporting.

10-C-iii-c. Type 1B diabetes

Nissley et al (1973) observed that the frequency of organ-specific antibodies in first-degree relatives of juvenile-onset diabetics who were themselves positive for these antibodies was significantly greater than that in relatives of antibody-negative diabetics. He believed this to indicate genetic heterogeneity of the diabetes, but as explained above this is an error of interpretation, since familial occurrence of the antibodies is to be expected irrespective of diabetes. He failed to find any other differences between the diabetics so grouped.

There is better support for the concept of Type 1B diabetes as described already (III-6-E-i; III-7-C-i-a), with the characteristics of late age at onset of diabetes, predominantly female sex, particularly marked association with 'S1' HLA specificities (B8 and DR3) rather than 'S2', and persistence of islet-cell antibody long after diagnosis, as well as association with various measures of organ-specific autoimmunity. Even here the evidence is less clear-cut than at first appears, for the following reasons.

The frequency of autoimmunity increases with age. Organ-specific autoimmune diseases are often associated with HLA-B8 and female sex. Furthermore, it is likely that the whole pathological process in Type 1 diabetes is slower when clinical onset is late in life than with onset in childhood, without any evidence for pathogenetic heterogeneity in this respect (see above). It follows that any sample selected for presence of organ-specific autoimmunity will be biased towards female sex, increasing age and HLA-B8 positivity, regardless of whether or not there is diabetes; if there is, the bias in achieved age will bias age at onset of diabetes, and will increase the likelihood of islet-cell antibody persisting for longer than average, if this antibody does indeed reflect the disease process in a temporal sense. Insufficient details are given in published reports to exclude effects of age and ascertainment bias, nor does a rigorous multidimensional analysis appear (Bottazzo et al, 1978a)

Nevertheless, the age at onset of diabetes in patients with thyroid disease was too great (mean 36 - 40y) to be explicable by bias when compared with that in Type 1 diabetes generally (III-4) (Bottazzo et al 1978a). Moreover, HLA-B8 was significantly more frequent in patients with diabetes and Graves' disease than in those with either disease alone; this is compatible with heterogeneity of the diabetes, but also with coexistence of two distinct, B8-associated susceptibility genes, one for diabetes and the other for Graves' disease. Long-persistent islet-cell antibody in the absence of clinical or biochemical endocrine autoimmunity was associated with higher frequency of B8 than found in ICA-negative Type I diabetic patients matched for duration (Op. cit.).

Here biased ascertainment due to selection for autoimmune disease does not apply, provided the two groups were ascertained in the same way (not stated).

Type 1B diabetes does therefore appear to be a distinct variant, or one end of a spectrum, of Type 1 diabetes. How far the long persistence of islet-cell antibody is a feature of this variant rather than merely a reflection of slow disease progression before onset in adult life is not yet clear.

10-C-iii-d. Islet-cell antibody in the very young patient.

The suggestion that islet-cell antibody is less common in Type 1 diabetes with onset at ages under 5 years than at school age is not supported by careful analysis of published data; just as for insulitis frequency, such variation as does exist is probably explicable by more rapid disease progression in the very young (III-7-C-i-a).

10-C-iv. Heterogeneity with respect to environmental factors.

In view of the large number of potentially diabetogenic infectious and toxic agents known (III-8), and the great variability of environmental influences, it seems most unlikely that all cases of Type 1 diabetes should be identical with respect to environmental events leading up to islet B-cell destruction. Indeed, it has been shown that, in certain strains of mice susceptible to diabetes, the disease may be brought about by any of a number of different viruses. Furthermore, diabetes may develop as a result of cumulative insults, viral and toxic, each individually insufficient to cause diabetes (Toniolo et al, 1980). Thus, human Type 1 diabetes is likely to be environmentally multifactorial, but not necessarily truly heterogeneous in the sense of the definition at the head of this chapter, i.e. 'composed of diverse elements or constituents'. The concept of a purely 'viral' form of Type 1 diabetes has been around for a long time (Irvine et al 1977), but even the rare fulminating cases most obviously belonging to such a category may have features suggesting a more prolonged autoimmune pre-diabetic phase than would fit that concept (III-8-B-ii-b).

If it could be shown that one particular environmental factor (or well-defined set of factors) interacted specifically with one type of genetic susceptibility to produce diabetes but not with another, distinct type of susceptibility, whereas a second factor did produce diabetes in subjects with the second type of susceptibility, then the definition of heterogeneity above would be satisfied. Such evidence is so far lacking.

10-D. CONCLUSION

Fundamental to many topics in this chapter is the negative effect of age at diagnosis on the rate of progression of the Type 1 diabetic disease process. It is instructive to view this relationship in the opposite sense: the more rapid is the disease process, the younger is likely to be the patient at onset of clinical Type 1 diabetes. Could it be that in general pathogenesis begins very early in life, so that the age at diagnosis is actually determined by the rate of progression? In that case, age at diagnosis might well be a manifestation of genetic heterogeneity. If on the other hand pathogenesis is initiated by environmental events during life, and age modifies the rate of progression, then the evidence for heterogeneity is weaker. Thus the duration of the pre-diabetic phase (III-9) is of great importance.

PART III — CHAPTER 11

ACTIONS OF SUSCEPTIBILITY GENES

CONTENTS

A Introduction · 137
B Immune response genes · 138
i Do they exist? · 138
ii A role in susceptibility to Type 1 diabetes? · · · · · · · · · · · · 138
C Susceptibility to viral infection · · · · · · · · · · · · · · · · · · 139
i Animal work · 139
ii Evidence in man · 140
D Other possible mechanisms · 140
i Susceptibility to toxins · 140
ii Islet B-cell regenerative capacity · · · · · · · · · · · · · · · · · 141
iii Impaired insulin secretion · 141
E Direct involvement of MHC-complex gene-products? · · · · · · · · · · 141
F Conclusion · 142

11-A. INTRODUCTION

It is widely considered likely that the mechanisms by which HLA-linked genes may determine susceptibility to Type 1 diabetes are modification of immune response so as to facilitate autoimmunity directed against islet B-cells, and alteration of host responses to virus infections. Both seem inherently probable, because of the nature of the HLA complex (III-7-A) and the evidence implicating auto-immunity and probably viruses as well in pathogenesis (III-7 and 8). Other possibilities might include susceptibility to substances toxic to the insulin-producing cells, impairment of islet B-cell regenerative capacity, and perhaps (as a contributory factor) reduction in insulin secretory function. This chapter reviews relevant evidence, much of it already mentioned.

11-B. IMMUNE RESPONSE GENES

11-B-i. Do they exist?

In animals there is evidence of immune response genes in the major histocompatibility complex (corresponding to the human HLA region), shown for example in breeding experiments (Munro & Waldman 1978). Such genes (Ia genes) in the H-2 system of inbred mice are considered to correspond to human DR genes (Silver et al, 1979).

Human breeding experiments are impossible so the evidence has to be more indirect, using genetic markers (including HLA). Injected exogenous insulin is an antigenic substance, the immune response to which has been extensively studied in man. Some reports claim a positive association between insulin antibody response and 'S1 axis' HLA specificities (Schernthaner, Ludwig & Mayr, 1979) but others have described different associations instead, including a possible weak negative one with B8, not quite significant at the 5% level (Bodansky et al, 1982), and, assuming only unrelated subjects were studied by Kahn et al (1982), a positive one with the combination A2,Bw44,Dr7. More recently, Schernthaner et al (1983) have changed their views, supporting a negative association with HLA-DR3 and a positive one with DR4, using highly-purified porcine and human insulins. The problems of statistical type 1 errors and selective reporting apply here as in other HLA and heterogeneity work (III-10-B). Controversy continues.

Other antigen-response systems showing evidence of HLA-linked or HLA-associated variation include leprosy (de Vries et al, 1976), influenza vaccination (Munro & Waldman 1978) and various allergies (Levine, Stembar & Fotino, 1972; Marsh, Bias & Hsu, 1973; Kapoor-Pillarisetti et al, 1981).

11-B-ii. A role in susceptibility to Type 1 diabetes?

Although of limited relevance to human diabetes, some animal models are interesting in this respect. Low-dose-streptozotocin-induced diabetes in congenic strains of mice appears to be under genetic control

of the major histocompatibility system, H-2 (III-3-C-iv). In the BB rat (III-7-D-iii), the major determinant of susceptibility to diabetes is probably a recessive gene;this is not linked to the major histocompatibility system, but its action appears to be dependent on the latter - an epistasis effect (Howard, 1983).

It is possible that some of the immunological abnormalities present in patients with Type 1 diabetes reflect primary effects of an immune response gene, but it is unlikely that such effects could be distinguished either from non-inherited immunological features of the pathogenetic process itself, or from metabolically-induced disturbances (III-7-A). Hence the importance of studies of non-diabetic relatives, using genetic markers.

In non-diabetic relatives of human Type 1 diabetic patients, not only diabetes but also organ-specific autoimmunity is more common than in relatives of control subjects (III-7-B), suggesting either a common susceptibility gene (not necessarily HLA-linked) or a haplotypic association (linkage disequilibrium) between an 'autoimmunity' gene and a different diabetes susceptibility gene. Since HLA-DR3 and associated specificities are associated with both Type 1 diabetes (III-6) and many forms of organ-specific autoimmunity, a simple hypothesis is that the same DR3-associated gene might have both effects (III-7-B). This hypothesis was tested for the first time in HLA-genotyped families as part of the present work (Gorsuch et al, 1980; see RESULTS, VIII-5-B-i).

11-C. SUSCEPTIBILITY TO VIRAL INFECTION

11-C-i. Animal work

Virus-induced diabetes in laboratory animals depends on the genetic background (strain) of the host (Bouchier & Notkins, 1973; III-8-B-I-b). Susceptibility of mice to diabetes induced by encephalomyocarditis virus (cf. III-11) behaves in cross-breeding experiments as a recessive trait (Ross et al, 1976). However, if this does have any bearing on human Type 1 diabetes, it can have little to do with the postulated HLA-linked S2

gene, since there is no evidence of linkage to the homologous H-2 system in the mouse (Craighead, 1981).

The mechanism of variation in susceptibility of mice to viral diabetes probably involves general host responses to viral infection (e.g. interferon) rather than more specific islet B-cell effects (Craighead, 1981; Wilson et al, 1983).

11-C-ii. Evidence in man

Hauptman (1980) postulated that, in man, genetically-determined abnormalities of complement components might be important in HLA-linked genetic susceptibility to Type 1 diabetes. This idea arose because the unexpressed allelomorphs of C4 now termed C4-AQ0 and BQ0 showed increased relative risks for the disease in certain populations (III-6-D), while C4 normally has a role in neutralization of sensitized virus (Daniels et al, 1969; Leddy et al, 1977). The concept is explored by Rittner & Bertrams (1981). Some support comes from the report that not only Type 1 diabetics but also their non-diabetic identical twins have abnormally low levels of circulating C4 (Vergani et al 1983).

A tenuous but suggestive recent observation is that, among 173 patients with congenital rubella syndrome, the 21 who developed diabetes showed the same disturbances of HLA-DR frequencies as are present in Type 1 diabetes generally (Rubinstein et al 1982).

Attempts to investigate this whole question in man are difficult because of the complexity of interacting genetic and environmental factors affecting host responses. So far the mode of action of the S2 diabetes gene remains an open question. The only other condition known to be associated with HLA-DR4 is seropositive rheumatoid arthritis, the aetiology of which remains equally obscure.

11-D. OTHER POSSIBLE MECHANISMS

11-D-i. Susceptibility to toxins

Genetically-determined variation in susceptibility to environmental toxic agents, just as to infectious agents, would be expected in any

species as a result of natural selection in different environments. Relevant here is the demonstration of such variation in susceptibility of laboratory animals to diabetes induced by low-dose streptozotocin (III-3-C-iv).

11-D-ii. Islet B-cell regenerative capacity

See III-3-C-ii.

11-D-iii. Impaired insulin secretion

Investigation in HLA-genotyped families had not been undertaken at the time of the present work. More recent reports are mentioned above (III-9-C).

11-E. DIRECT INVOLVEMENT OF MHC-COMPLEX GENE-PRODUCTS?

MHC-complex gene-products might act by impairing immune response generally, as suggested above (III-11-C) for C4-null alleles and susceptibility to viral infection. Alternatively, or additionally, they might directly influence interactions between the islet B-cell membrane and immunocytes or viruses; it is relevant that the islet B-cell can express DR gene-products, at least in presence of insulitis (Alejandro et al, 1981; Bottazzo et al, 1984) and the BB rat (Dean et al, 1985).

Direct involvement of several gene-products in susceptibility implies either heterogeneity or epistasis, or, in the present context, both (i.e. different sets of interacting genes leading to different types of susceptibility - see 10-C-iii). This could explain the rather low relative risks found for association between individual DR antigens and Type 1 diabetes as a whole, an observation generally considered incompatible with the view that susceptibility is determined directly by DR3 and DR4 alone (III-6-C) - although a further possible explanation is that the DR antigens recognized serologically may not in fact be homogeneous: Owerbach et al (1983) showed differences at the DNA level between DR4 in Type 1 diabetic subjects and that in control subjects.

A prediction of the 'direct involvement' hypothesis is that true genetic heterogeneity must exist, since Type 1 diabetic patients differ so widely in their HLA and complement types. Conversely, good evidence of such heterogeneity would favour the hypothesis (III-10).

11-F. CONCLUSION

Current understanding of the mechanism by which susceptibility genes might act is still largely conjectural. While new techniques in molecular genetics will undoubtedly lead to great advances, there is still a need for statistically sound investigation at the level of the whole organism, so as to determine the differences between a susceptible and a non-susceptible individual. One approach is embodied in the present work.

PART IV

INDICATIONS FOR PRESENT RESEARCH

CONTENTS

Introduction · 143
A Basic genetic and immunological observations · · · · · · · · · · · · 144
B Genetic heterogeneity within Type 1 diabetes; mode of inheritance 145
C The pre-diabetic period: prospective studies · · · · · · · · · · · · 146
D Quantitative risk of Type 1 diabetes in siblings of affected · · · 145
children · 146
E Nature of susceptibility · 146

INTRODUCTION

The Literature Survey raises many questions concerning aetiology of Type 1 diabetes. Perhaps the most fundamental are the following:

- What is the nature of inherited susceptibility?
- Can the risk of future diabetes developing in a given child be quantified?
- How long before onset of clinical diabetes does the disease process begin, what environmental factor(s) initiate it, and by what mechanism do those factors interact with genetic susceptibility to cause islet B-cell destruction?

Such questions need to be answered before the long-term aim of prevention of the disease is likely to be achieved, but hitherto most attempts to answer them have been hampered by lack of effective markers of susceptibility or of the pre-diabetic phase of pathogenesis, by the effect of diabetes itself on immunological and other variables, and especially by the retrospective nature of studies directed at those already diabetic.

The principal aim of the present work was to establish a cohort of non-diabetic individuals in whom to investigate prospectively events occurring during the pre-diabetic phase. Given the population incidence of Type 1 diabetes in childhood (of the order of 10^{-4} per year) this would necessitate an impracticably large cohort of at least 10^4 individuals, if selected at random. Using the non-diabetic siblings of diabetic children instead would bring two advantages: first, the incidence in such siblings is increased about 25-fold (Lit. Survey, III-6-F), so that several new cases of diabetes should occur over a few years in a cohort of only a few hundred; secondly, HLA-genotyping within the families would allow identification of subgroups of siblings with high, intermediate and low probabilities of being susceptible, according to the number of haplotypes in common with the diabetic probands (III-6-C-ii).

The decision to concentrate on HLA-genotyped families, each with one or more Type 1 diabetic child, was strengthened by the opportunity thus afforded to investigate many outstanding questions, which form subsidiary indications for the study and are detailed below.

A. BASIC GENETIC AND IMMUNOLOGICAL OBSERVATIONS

Further information was needed on the following:

<u>HLA-genetics</u>

- Linkage disequilibria within the HLA-region (III-6-B-iii), sought or confirmed by gene-counting techniques in the genotyped families, without assuming Hardy-Weinberg equilibrium (see B below);
- Zygotic assortment of HLA haplotypes within families (III-6-B-iv): random or not?
- Associations between Type 1 diabetes and genetic markers on chromosome 6 (III-6-C-i; III-6-D) in the defined population concerned.

<u>Autoimmunity</u>

- Occurrence of clinical diabetes and other organ-specific autoimmune conditions among relatives of patients with Type 1 diabetes (III-5; III-6; III-7-B);

- Occurrence of islet-cell, thyroid and other autoantibodies in Type 1 diabetic subjects and their relatives (III-7-C-i-a; III-7-B);
- Correlation between results of assays for islet-cell antibody by complement-fixing and by conventional techniques (III-7-C-i-a).

B. GENETIC HETEROGENEITY WITHIN TYPE 1 DIABETES; MODE OF INHERITANCE

Careful investigation for evidence of genetic heterogeneity within childhood Type 1 diabetes (III-9) formed a major indication for this work, and would be facilitated by availability of supplementary observations from large numbers of diabetic multiplex sibships and other diabetic subjects ascertained by colleagues and by the author before and during the course of the main study (III-6-C-ii).

Further information on the following would be relevant to mode of inheritance of susceptibility:

- disturbances of Hardy-Weinberg equilibrium within the HLA system in diabetic subjects, such as might arise if the gene in question, or one in tight linkage disequilibrium with it, made a major contribution to susceptibility either in a recessive mode (excess of homozygotes) or exclusively in the heterozygous state (e.g. DR3,4: III-6-E-iii);
- differences in degree of linkage disequilibrium within the HLA system between 'diabetic' and 'non-diabetic' haplotypes, to be expected if HLA-complex gene-products themselves contribute to susceptibility by epistasis (III-6-E-v);
- distribution within diabetic sib pairs of haplotypes from parents negative for both DR3 and DR4 (III-6-E-v);
- occurrence of Dw2/DR2 in diabetics, and frequency of these on 'diabetic' and 'non-diabetic' haplotypes from parents negative for both DR3 or DR4 (III-6-E-iv: a protective gene?)

Even in the absence of HLA-DR data, HLA-B and complement factor studies might yield useful information.

C. THE PRE-DIABETIC PERIOD: PROSPECTIVE STUDIES

Information was needed on the following:

- The value of islet-cell antibody (conventional and complement-fixing) as a marker of pre-diabetes in children (III-7-C-i-a);
- Duration of the pre-diabetic phase in children (III-9);
- Whether viral infections, if relevant at all, initiate pathogenesis or merely trigger metabolic decompensation late in the prediabetic period (prospective investigation needed) (III-8-B);
- Whether (and when) simple metabolic tests can identify pre-diabetic changes in carbohydrate tolerance (III-9-C).

D. QUANTITATIVE RISK OF TYPE 1 DIABETES IN SIBLINGS OF AFFECTED CHILDREN

Further estimates of risk in siblings, taking into account age, geographical variation in incidence and HLA genotypes, would be valuable in genetic counselling and future attempts at prophylaxis (III-6-F).

E. NATURE OF SUSCEPTIBILITY

Investigation of HLA-genotyped siblings of children with Type 1 diabetes may provide evidence for the following postulated mechanisms of action of susceptibility genes:

- Modification of autoimmune responses (III-11-B);
- Modification of immune responses to virus infection (III-11-D);
- Certain aspects of cell-mediated immunity (III-7-C-ii);
- Differences in islet B-cell function (III-11-D-iii);
- Direct effects of MHC-complex gene-products (III-11-E).

PART V

WORK DONE

CONTENTS

A	The Bart's-Windsor Family Study	147
	i Conception, 1977	147
	ii Planning, 1977-1978	148
	iii Ascertainment and fieldwork, 1978-80	148
	iv Laboratory work	148
	v Data handling	149
	vi Continuity; work included in this Dissertation	149
B	Study of newly-diagnosed Type 1 diabetes	149
C	Extended analyses requiring additional data	150

This chapter outlines the origins, planning, execution and analysis of the work on which the Dissertation is based, and indicates the author's own contribution.

A. THE BART'S-WINDSOR FAMILY STUDY

A-i. Conception, 1977

As a result of his observations in Liverpool the late Professor (then Dr.) Andrew Cudworth believed that a prospective family study would help elucidate the pathogenesis of Type 1 diabetes (III-6-C-ii; Cudworth, 1977; see IV INDICATIONS). Dr. John Lister's records of childhood (and adult) diabetic patients in East Berkshire offered a nucleus of families for such a study. Responsibility for the gestation, birth, naming and maturation of the Bart's-Windsor Family Study was undertaken by the author, who had been clinically involved with many of the patients for up to three years.

A-ii. Planning, 1977-1978

With the advice of Professor Cudworth and others (see XI ACKNOWLEDGMENTS) the author planned the following aspects: ascertainment area, method and criteria; numbers of families needed; liaison with physicians clinically responsible; communication with families; laboratory investigations indicated; data handling. He also obtained financial support (MRC project grant) and local ethical committee approval.

A-iii.Ascertainment and fieldwork, 1978-80

All aspects of contact with patients and their families in the first two years were organized, documented and personally undertaken by the author: liaison with physicians concerned; searches of clinical records; initial approach to families; initial visits to family homes to obtain informed consent, relevant clinical and family histories, and blood samples from household members (and sometimes more distant relatives); and regular follow-up visits for further information and blood samples. The first family visit was in April 1978.

A-iv. Laboratory work, 1978-1980

As well as venesection, initial processing, storage, distribution and documentation of all samples was done by the author until mid-1980. Because of the scale of the project it was impossible for the same person to perform all laboratory investigations unaided; collaboration was therefore necessary.

All subjects were HLA-ABC-typed. The author worked in the HLA laboratory at St. Bartholomew's Hospital and was closely involved with this aspect, in collaboration with Dr. Eva Wolf, Professor Cudworth, and initially Professor H. Festenstein (The London Hospital). Selected families were also HLA-DR-typed during this period (see F below).

Screening for autoantibodies was performed on all samples at the Middlesex Hospital in collaboration with Dr. G-F Bottazzo and Dr. Betty Dean.

Collaborative arrangements were made for the following assays with which the author was not directly involved (VI-2): viral serology; HLA-Dw by mixed lymphocyte culture; red-cell glyoxalase and HbA1; complement C2 allotypes; Bf allotypes. Investigations of cell-mediated immunity in the families were made in association with Dr. P. Pozzilli (St. Bartholomew's Hospital), but fall outside the scope of this thesis.

A-v. Data handling, 1978-1985

The author is directly responsible for all aspects of data handling in connection with this work. Written records of all information obtained were kept, and many statistical analyses were made from these. However, because of the complex nature and very large volume of data produced by the family study, a computer-based system was soon seen to be necessary. A new, specialized system, GENOBASE, was created jointly by the author and Dr. D. Ingram (VI-3-B). All data used here were entered and checked by the author. Statistical packages MINITAB and GLIM, and, for specialized analyses, microcomputer/calculator programs written by the author, were used in conjunction with GENOBASE. Further processing of data and preparation of this Dissertation were undertaken by the author with the aid of word-processing, spreadsheet and graphics software.

A-vi. Continuity; work included in this Dissertation

The Family Study continued after the author withdrew from direct involvement with the families in 1980. Certain aspects of the planned work were incomplete at that time and are therefore not included here, being covered by successors, Drs. K. Spencer and A. Tarn. This applies in particular to HLA-DR results (except where they clarify otherwise ambigous haplotype segregation), C4 allotypes, changes in organ-specific and viral antibodies with time, and detailed metabolic studies within families.

B. STUDY OF NEWLY-DIAGNOSED TYPE 1 DIABETES

New cases of Type 1 diabetes developing during the main study in the same geographical area were of interest for several reasons: they

facilitated investigation of immunological and other phenomena before insulin treatment, they could supplement numbers of diabetic subjects available for heterogeneity and other analyses, they provided a source of new families to enter the family study, and they permitted a new estimate of local incidence.

The author arranged to be informed of any new case known to the physicians caring for diabetic children in the area, and to obtain blood samples as soon as possible after diagnosis for HLA and immunological studies. In most cases he supplemented this by a personal visit. Clinical and family histories were obtained. In several cases the family joined the main study.

C. EXTENDED ANALYSES REQUIRING ADDITIONAL DATA

Heterogeneity analyses in particular necessitated larger numbers of observations than were available from A and B above. The UK study of diabetic multiplex sibships, based on the Paediatric Register of the British Diabetic Association and carried out from St. Bartholomew's Hospital by Walker and Cudworth (1980), was deliberately planned for compatibility of data with the Bart's-Windsor Family Study. Professor Cudworth asked the author to include results of that study, and of his earlier work in Liverpool (Cudworth, 1977), in appropriate analyses. This has been done.

PART VI — METHODS

CHAPTER 1

FIELDWORK

CONTENTS

A Ascertainment and recruiting of subjects studied · · · · · · · 151
i Barts-Windsor Family Study · · · · · · · · · · · · · · · · · 151
a Ascertainment criteria · · · · · · · · · · · · · · · · · · 151
b Ascertainment method, and geographical area · · · · · · 152
c Ethical aspects · 154
d Recruitment procedure · · · · · · · · · · · · · · · · · · 155
e Ascertainment rate and bias · · · · · · · · · · · · · · · 155
ii Newly-diagnosed Type 1 diabetic patients · · · · · · · · · 157
B Information and specimens obtained at first visit · · · · · · 157
i Clinical history · 157
ii Family structure · 158
iii Family history of diabetes and other conditions · · · · · 158
iv History of exanthemata and vaccinations · · · · · · · · · 158
v Blood samples at first visit · · · · · · · · · · · · · · · 158
C Prospective study · 159
i Frequency and duration of follow-up · · · · · · · · · · · 159
ii Record diary · 159
iii Blood samples · 159
iv New development of diabetes · · · · · · · · · · · · · · · 159

1-A. ASCERTAINMENT AND RECRUITING OF SUBJECTS STUDIED

1-A-i. Barts-Windsor Family Study

1-A-i-a. Ascertainment criteria

Families were selected for inclusion in the main study on the basis of the following criteria:

- European Caucasoid origin. Most surnames were British.

- Clinical evidence of Type 1 diabetes in the proband. Insulin treatment and history of weight-loss mandatory. Records of significant non-fasting ketonuria (++ on routine testing with Acetest or Ketostix) were sought in all cases and confirmed in most.

- Availability of at least one sibling aged < 20 years at time of entry to study, whether or not diabetic. Siblings older than this would

be less useful for prospective study because they might move away and, if not diabetic, would have passed through most of the period of high risk for diabetes. Families where the only sibling was aged < 5 years were included only at the parents' particular request.

- Availability of at least one parent, and of both parents where the proband had only one sibling. This was to reduce the likelihood of ambiguous HLA haplotype segregation results.

- Residence and intention to remain in a geographically accessible area for at least two years.

- Absence of severe psychosocial barriers to participation, in the opinion of parents and medical attendants (three families from the central, East Berkshire district were rejected for this reason).

- Informed consent (see below).

1-A-i-b. Ascertainment method, and geographical area.

The geographical area concerned is shown in Figure VI-1. Some demographic details are given in Appendix 6-1.

The majority of families were ascertained through the patient index of juvenile and adult diabetics maintained continuously in East Berkshire (Windsor and Slough clinics) since the 1950's. Throughout this period all childhood cases of diabetes admitted had been managed jointly by Dr. Lister and his paediatric colleagues, and followed up in the diabetic rather than the paediatric clinic. The index included not only patients currently attending the clinics but also any others who had been admitted or attended as outpatients in the past. All insulin-treated patients on the index were considered. Information relevant to ascertainment criteria was obtained by direct questioning of those still attending the clinics, by perusal of clinical case records, occasionally by contact with the general practitioner, and finally by direct contact with the family at home (see below). New cases diagnosed during the first year or so of the study were also considered for inclusion (see below, A-ii).

Physicians caring for juvenile and adolescent diabetic patients in two adjoining districts kindly allowed inclusion of families from N.W. Surrey (Dr. Jean Scott and colleagues at St. Peter's Hospital, Chertsey)

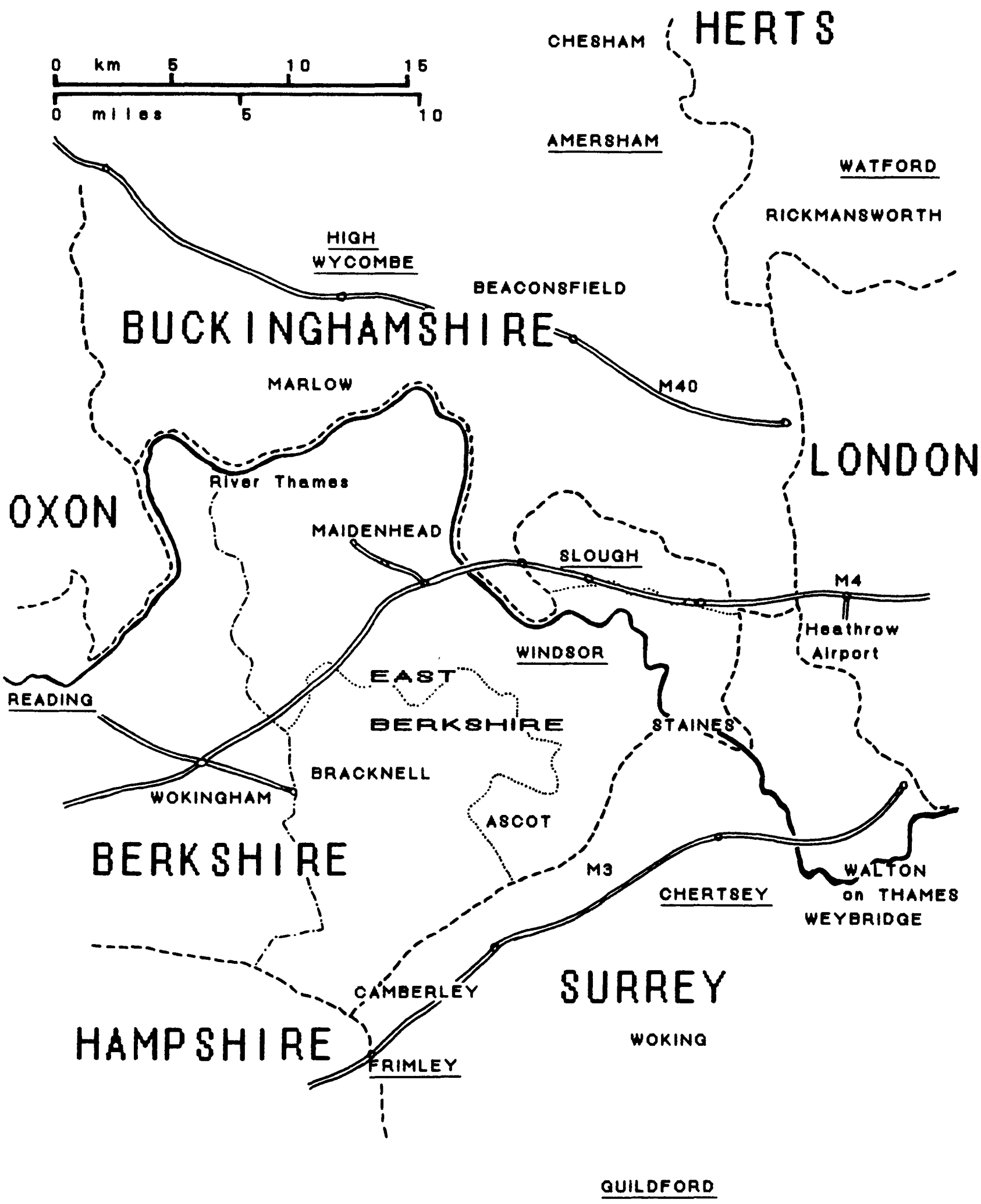


FIGURE VI-1. MAP OF EAST BERKSHIRE AND SURROUNDING DISTRICTS.
Towns underlined have hospital clinics attended by diabetic children.
Key: ---county boundary; ·····district boundary; -·-·East Berks boundary.

and Wycombe and Amersham Health District, Buckinghamshire (Drs. Bruce Thompson, Tessa Westcott, C. Cheetham and D. Garrow).

A few families who had never lived or attended clinics in the main study area came to the attention of the author by various means. Two of these, particularly keen to take part and readily accessible to the author at their homes just to the north of the area, were included in the prospective study.

Newly-diagnosed patients notified to the author during the first year of the study (1-A-ii) were included if criteria were met. One suitable subject was 'missed' because not notified until later.

1-A-i-c. Ethical aspects

The protocol was submitted to and approved by the Ethical Committee of the East Berkshire Health District in 1978. One potentially controversial point, the taking of blood samples from healthy children, was considered fully acceptable in the presence of the parents and with their continued consent, given that there would be a diabetic child in each family. In practice most children were surprisingly tolerant of occasional venepuncture by the author, even if few relished it. Only a small minority never took part or withdrew from followup (see VII). Similar ethical approval was readily obtained by collaborating physicians in adjoining districts.

The prior agreement of the hospital physician caring for the diabetic proband was invariably present. General practitioners were informed and their comments invited before ascertainment of individual families; none expressed disapproval, while a number were actively supportive. Any medically important finding in family members (anaemia or thyrotoxicosis, for example) was immediately notified to the GP.

Informed consent was a condition of recruitment and is considered below.

1-A-i-d. Recruitment procedure

Families apparently satisfying ascertainment criteria received an introductory letter from the hospital consultant, and a more detailed one from the author. After a few days this was followed by a telephone call, when a few families proved unsuitable because too few family members were available (3 families in central area) or no reply could be obtained despite repeated attempts (1 in central area). Most of the remainder agreed to a home visit; however, a small number including two in the central area declined for unspecified reasons.

The author visited each family by arrangement, at home on a Sunday or a weekday evening. The aims and implications of the study were again explained and understanding ensured. It was agreed with the family that, should tests suggest a sibling to be at particular risk of becoming diabetic, this would be disclosed only if diabetes developed or (on request) after withdrawal from the study. There were several reasons for this rule. First, the information would bring no direct benefit since no safe and effective prophylaxis was yet available. Secondly, knowledge of such high risk in a particular sibling might cause him deleterious stress; thirdly, that stress might itself increase the likelihood of progression to diabetes; and finally, if it did so, then this would bias any estimate of predictive value of the tests concerned.

Only one family (in the central area) decided not to participate at this stage, the remainder giving informed consent.

1-A-i-e. Ascertainment rate and bias

This was not planned as an epidemiological study. Nevertheless, the question of possible bias toward multiply-affected families is relevant to the calculations on risk of recurrence within sibships (VIII-D). Such bias is to be expected where the ascertainment rate is low: if the probbility of any diabetic child being picked up is a, then the probabilty for a simplex family (one diabetic child) is also a, but the probability for a family with two affected children, for example, is $a + a(1 - a)$. This tends to 2a as a tends to zero, and to a as a tends to 1.0.

In the peripheral areas (N.W. Surrey; Wycombe and Amersham) the ascertainment rate is undoubtedly low. By chance, however, only one of the families ascertained in these areas included more than one diabetic child at the time; thus no such bias exists among them.

In the central study area (East Berkshire) the ascertainment rate is probably high, for the following reasons. First, for two decades Dr. Lister was the only physician in that district with an interest in juvenile diabetes. All insulin-dependent patients at that time, and certainly all children, would have been admitted to paediatric or adult medical wards for initial education and treatment, and if not initially under his care would have been under his care solely or jointly with paediatric colleagues. Thus patients diagnosed in the district are unlikely to have escaped inclusion on the patient index (see above). Patients moving into the district after diagnosis and not seeking specialist care locally would be approximately balanced by those moving a short way out of the district but still appearing on the patient index, with the possible exception of Bracknell, a 'New Town' with considerable growth in the 1970's. Similarly, cases from within the district referred to specialists outside would be roughly balanced by the converse.

Secondly, despite widespread publicity (including coverage in a BBC Television 'Horizon' feature in 1979) no further families inadvertently omitted from earlier ascertainment came to light within the central area. Thirdly, among the few families on the patient index who were otherwise suitable but excluded because of unavailability of family members, one included two diabetic children. This balanced the inclusion of a single family with more than one affected child which had moved outside the district.

Finally, the white Caucasoid population aged <20 years in East Berkshire in 1981 was approximately 82,200, with mean age 10.1 years (VIII-4-B; Appendix 6-2). The cumulative frequency of diabetes is probably between 2.1 and 2.75 per thousand at age 20, and about half this at the mean age of 10 (Figure III-3; Literature Survey, III-4-A-v). Thus the expected number of diabetic children in the population concerned is

between 86 and 113 (cf. VIII-4-B). Assuming that 1 in 10 is an only child, and a further 1 in 10 has no sib aged <20 years, it is reasonable to expect some 90 diabetic children aged <20 years in East Berkshire satisfying the ascertainment criteria for the BWFS. In fact 85 such children were ascertained.

Thus, even though the ascertainment rate is not known exactly, it is clearly high in the only area where bias towards multiplex families might otherwise have resulted.

1-A-ii. Newly-diagnosed Type 1 diabetic patients

The author arranged to be informed of any previously-undiagnosed case of Type 1 diabetes in a child or young adult admitted to any of the hospitals in the study area. Local clinical staff usually obtained blood samples at medically-indicated venepuncture, often before insulin treatment was started. The author generally visited the hospital within 12 hours or so to collect the samples and obtain clinical and family information. Later, the families of suitable patients were invited to join the Family Study as above, and such cases diagnosed during the first year (up to April 1979) are included in the family study analyses presented here.

1-B. INFORMATION AND SPECIMENS OBTAINED AT FIRST VISIT TO FAMILY

1-B-i. Clinical history

The following information concerning any case of diabetes was obtained by perusal of hospital case records, and by questioning of the family:

- Date (and age) at diagnosis of diabetes;
- Nature and duration of symptoms preceding diagnosis;
- Events believed to have preceded and perhaps precipitated onset;
- Past and present treatment, including insulin dose;
- History of ketoacidosis or ketonuria;
- Presence of complications.

1-B-ii. Family structure

A family tree was drawn and annotated in presence of the family, to include all known pregnancies (including miscarriages) of the proband's mother, with date of birth, sex, birth history, and (in most families) whether breast-fed. The proband's second-degree and some third-degree relatives were added.

1-B-iii. Family history of diabetes and other relevant conditions

A detailed family history was obtained, with special emphasis on diabetes (with age at onset and whether insulin-treated), thyroid disease (characterized so far as possible), pernicious anaemia ('treated by regular injections'), Addison's disease, coeliac disease, vitiligo, premature greying of hair ('substantially grey by age 30'), and rheumatoid arthritis where this could be distinguished from degenerative arthropathies. In general it was not practicable to consult medical records, but so far as possible the diagnoses were substantiated by appropriate questions. Frequently a parent contacted a more distant relative for confirmation. The date of onset was obtained for conditions within the nuclear family.

1-B-iv. History of exanthemata and vaccinations

For all children a history of exanthemata and vaccinations was elicited, usually from the mother. In a few families this information was clearly unreliable or lacking. In many it was surprisingly detailed, dates being provided. Particular emphasis was placed on the following: mumps, measles, chicken-pox, german measles, glandular fever, whooping cough; immunizations against diphtheria, tetanus, pertussis, polio, measles, smallpox, rubella and tuberculosis.

1-B-v. Blood samples at first visit

From all available family and household members blood was taken for baseline investigations: samples for isolation of mononuclear cells for HLA-ABC typing, DR typing in some, and storage for HLA-Dw and other work; clotted blood to provide serum for autoantibody screen, viral serology, HLA antibody screening in the mothers, storage against future requirements, and in many cases complement allotypes; whole blood in EDTA

for glyoxalase isoenzymes and glycosylated haemoglobin assay. Some families were not tested for all these, but all were HLA-ABC typed and tested for organ-specific autoantibodies.

1-C. PROSPECTIVE STUDY

1-C-i. Frequency and duration of follow-up

Only families in which there was one or more non-diabetic member were followed up prospectively. It was planned that visits should normally be repeated at intervals of about three months, but more frequently where a non-diabetic member was found to be positive for islet-cell antibody. The actual frequency and duration of follow-up achieved are summarized under RESULTS (VIII-3-B-i).

1-C-ii. Record diary

Each family was given a notebook divided into sections, one for each family member, with instructions. This was intended for recording of any events, particularly any which might reflect viral infections. Non-diabetic members were also requested to test the urine with Diastix (Ames) at monthly intervals and to record the results. Any positive urine test was to be reported immediately to the author, who would then make an extra visit in case (presymptomatic) diabetes had developed.

1-C-iii. Blood samples

At every visit clotted blood samples of about 12 ml were taken from all available family members, aliquots of serum being destined for auto-antibody testing, for storage with a view to estimation of serial viral antibody titres, and to be held in reserve. Extra samples were taken from time to time for supplementary genetic, immunological or other work.

1-C-iv. New development of diabetes

When non-diabetic subjects were known to be or to have been positive for islet-cell antibody, or had reported glycosuria, random blood glucose and glycosylated haemoglobin levels were assayed on most occasions (see below). Mildly abnormal results were an indication for more detailed metabolic studies at St. Bartholomew's Hospital, but if decompensating diabetes was present immediate local admission was arranged.

PART VI - CHAPTER 2

LABORATORY METHODS

CONTENTS

A Immunogenetics 160
i HLA-A,B,C 160
a Separation of viable mononuclear leucocytes 161
b Typing sera 162
c Typing method 162
d Quality control 163
ii HLA-DR 163
iii HLA-D 164
iv Factor B (Bf) 164
v Complement C2 164
vi Glyoxalase I (GLO) 164
B Autoantibodies 164
i Islet-cell antibody 164
ii Gastric parietal-cell and adrenal-steroid-cell antibodies . 166
iii Thyroid microsomal and thyroglobulin antibodies 166
iv Antinuclear antibodies 166
C Viral serology 167
D Metabolic indices 167
i Urine glucose 167
ii Blood glucose 168
iii Glycosylated haemoglobin 168
iv. Islet B-cell secretory capacity 169
E Prevention and detection of errors of substitution 169

2-A. IMMUNOGENETICS

2-A-i. HLA-A,B,C

In collaboration as stated above (V-A-iv), the standard micro-lymphocytotoxicity technique (Terasaki et al, 1978) was used, with minor modifications according to contemporary practice in the laboratories concerned.

2-A-1-a. Separation of viable mononuclear leucocytes

Peripheral blood lymphocytes were used. In general, aliquots of 10ml venous whole blood were drawn into sterile plastics containers each containing 2ml of sodium citrate solution (3.8%), and as soon as practicable an equal volume of culture medium was added to maintain viability overnight. The medium was a mixture of TC199 (Wellcome Reagents, Dartford) 10ml, sodium bicarbonate (4.4%) 2ml, HEPES buffer (1.0 mol/l) (Wellcome Reagents) 1ml and antibiotic solution 1ml (containing ampicillin, cloxacillin and gentamicin). Where possible 20ml of blood were used so that sufficient cells would be available for storage and/or HLA-DR typing, but this was reduced in small children (indeed, one baby was typed on 1-2 ml heparinized blood).

Samples were kept at room temperature (not less than 14°C) since lower temperatures in cold weather reduced yield. On arrival in the laboratory they were defibrinated by continuous agitation in conical flasks after addition of 1-2 drops bovine thrombin (50 IU/ml in phosphate buffered saline (PBS; Oxoid, Basingstoke)) and 0.3 ml of calcium gluconate solution (10%) for each 1 ml of citrate used.

In a few instances samples were defibrinated immediately on venesection, but it was impracticable for an unassisted operator to do this routinely during a family visit. The alternative procedure gave satisfactory results.

Defibrinated samples were layered onto a mixture of Ficoll (Pharmacia, Uxbridge) and Triosil (Vestric, Enfield) with a specific gravity of 1.080, before centrifugation for approximately 10 min. at 600g. Mononuclear cells were aspirated from the density junction by Pasteur pipette, washed twice by centrifugation in barbitone buffer (Oxoid) and counted in a haemocytometer. Cells not immediately required for typing or cell-mediated immunity experiments were adjusted to $4x10^6$/ml and aliquots frozen for storage in liquid nitrogen in an equal volume of medium consisting of dimethyl-sulphoxide (DMSO) 20%, group AB human serum 40% and TC199 40%.

2-A-i-b. Typing sera

Cells were typed using a minimum of 120 well-characterized sera, some obtained locally and some by exchange with other laboratories. Most local sera were provided several weeks or months post-partum by voluntary donors found by screening in local antenatal clinics. In addition the Barts-Windsor Family Study was itself a source of sera. Mothers were screened by cross-matching their sera against cells of their husbands or children. Positive results were surprisingly common in view of the long period which had typically elapsed since the latest pregnancy, and were defined on screening plates of cells of known specificities. Several mothers with useful antibodies agreed to donate serum during routine visits by the author.

As more sera became available and new specificities were characterized internationally, most families were re-typed on fresh (or occasionally stored) cells. In some cases this was not possible because a family had withdrawn from the study. Thus a number have never been typed for the 'newer' HLA-Cw antigens Cw6, 7 and 8. In the majority, however, all the HLA-A, B and C specificities recognized by the VIIIth Histocompatibility Workshop have been tested for (see III Literature Survey 6-B-i, Figure 4).

2-A-i-c. Typing method

Viable mononuclear cells (adjusted to 1.5×10^6/ml in culture medium) were added from a Hamilton multiple microdispenser in aliquots of 1.0 µl to the wells of Terasaki typing plates (Falcon, Scientific Supplies, London), previously charged with an array of typing sera (1.0 µl per well) under mineral oil and stored at -40°C. Plates were incubated at room temperature (about 22°C) for 1 hour, then 5 µl non-toxic rabbit complement (Buxted) were added to each well, and after further incubation at room temperature for 2 hours the plates were read by an experienced

and by re-typing on the same and, later, on fresh samples from the same subjects. No antigen was counted positive unless indicated by at least two different sera.

2-A-i-d. Quality control

All results were checked by a senior member of the laboratory staff (Dr. Wolf) with extensive personal knowledge of HLA serology, cross-reactions and the characteristics of the sera used. During the period of this study the laboratories concerned participated in Histocompatibility Workshops (Bodmer et al, 1977; Terasaki, 1980). A further guarantee of accurate typing in the Family Study was the ability to check consistency of results in different family members, and the opportunity to retype at follow-up visits; thus most of the families have been typed at least twice.

2-A-ii. HLA-DR

HLA-DR results are used here only to confirm otherwise ambiguous haplotype segregation within families, and to investigate possible intra-HLA recombinants suggested by crossovers between HLA-B and GLO loci (see V-A-vi). Typing was performed in collaboration, initially with Prof. H. Festenstein at the London Hospital and later with Dr E. Wolf at St. Bartholomew's Hospital.

The two-colour double immunofluorescence technique of van Rood et al (1976) was used. Briefly, viable mononuclear cells isolated as above were added to wells on a Hamax 60-well glass-bottomed plate (Medicell, London) charged with well-characterized anti-DR sera. The B-lymphocytes present were labelled by fluorescein-conjugated sheep anti-human-Ig serum, the cell suspension was applied to the typing plates as before (0.5 µl per well), the plates were incubated before and after addition of complement, and ethidium bromide was added to show cell-death. Plates were examined by combined phase-contrast and UV microscopy. Positive reactions were shown by presence of cells fluorescing red (dead) with a green 'cap' (B-lymphocytes). Specificities tested were those of the VIIIth Workshop (see above).

2-A-iii. HLA-D

Diabetic probands selected for presence of HLA-Bw62, B8 or B7 were kindly typed by Dr. J. Sachs using mixed lymphocyte culture between responder cells from the probands and a panel of well-characterized homozygous typing cells; proliferation, shown by uptake of tritiated thymidine, indicated that the specificity of the typing cells was absent from the responder cells (Sachs et al, 1980). Specificities Dw1 to 10 were tested for.

2-A-iv. Factor B (Bf)

Selected families (principally those with a haplotype positive for HLA-B18) were kindly typed by Dr. J. Markwick, who subjected serum to thin-layer agarose-gel electrophoresis in veronal buffer at pH 8.6 before immunofixation using antisera specific for properdin factor B (Alper et al, 1972).

2-A-v. Complement C2

Sera from most probands were typed for C2 by Dr. L. Wells, using isoelectric focussing in polyacrylamide gel followed by complement-dependent lysis in an overlay from a donor congenitally deficient in C2 (Meo et al, 1976). Selected families, principally where an intra-HLA recombinant was suspected, were fully typed.

2-A-vi. Glyoxalase I (GLO)

Red-cell glyoxalase I (GLO; EC 4.4.1.5) catalyses irreversible conversion of glutathione and methylglyoxal to S-lactoyl-glutathione. Two iso-enzymes occur. Most families were typed by Dr. S Welch, using starch-gel electrophoresis of haemolysates followed by incubation with methylglyoxal and glutathione and addition of iodine to stain the starch gel in areas where glutathione had been consumed (Parr et al, 1977).

2-B. AUTOANTIBODIES

Serum samples were frozen immediately after separation, as soon as practicable after venesection, and stored at -20°C until analysis.

2-B-i. Islet-cell antibody

The standard immunofluorescence 'sandwich' test was used (see III-7-C-i-a): undiluted serum was applied to unfixed cryostat sections of fresh Group O+ human pancreas and incubated for 1 hour at room temperature. Slides were then washed, incubated with sheep anti-human-IgG-fluorescein-isothiocyanate conjugate (Wellcome), washed again, dried, and examined by Ploem UV epi-illumination microscopy. Brightness of islet fluorescence was graded from weak positive (w+) to very strongly positive (+++), and agreement obtained between two experienced observers unaware of the diabetic status and previous islet-cell antibody results on the subject concerned. Titres were not routinely determined because of the very large numbers of tests to be performed and the limited supply of substrate pancreas.

For the complement-fixing test (CF-ICA), the modifications described in III-7-i-b were used. In many cases a 'selective' pattern of fluorescence was observed (Figure VI-2).

Rare instances of fluorescence confined to a few peripheral islet cells only were shown to represent glucagon-cell or somatostatin-cell antibody by counterstaining with an appropriate antiserum conjugated with a different fluorescent marker. These were not counted positive for ICA.

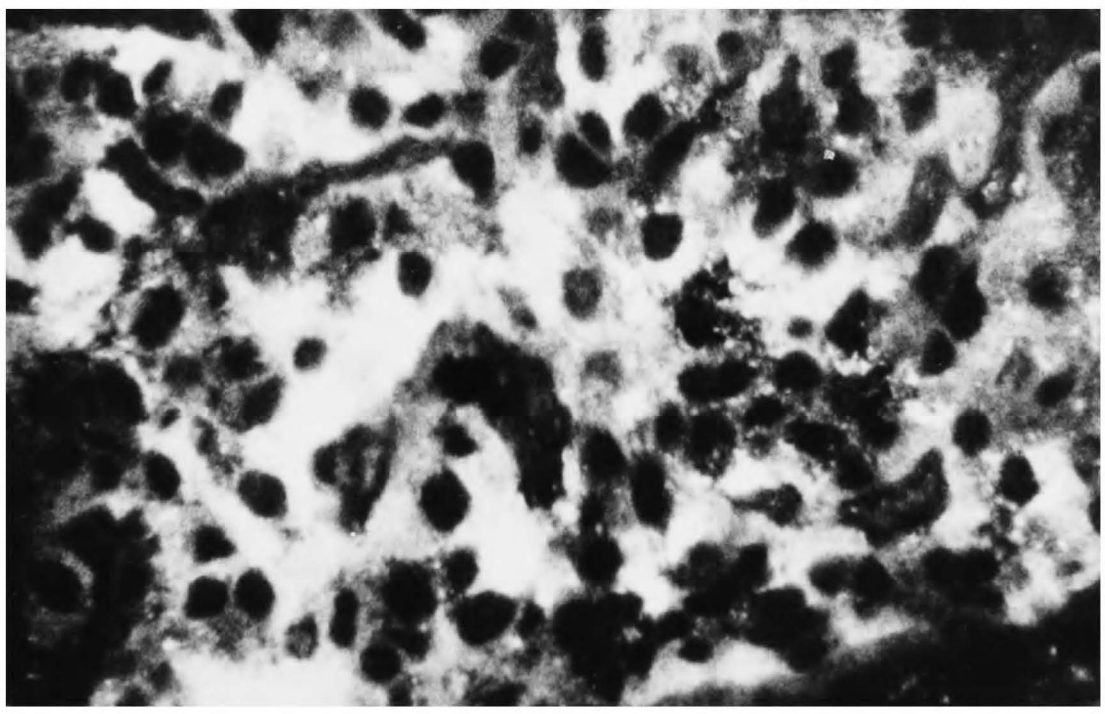


FIGURE VI-2. COMPLEMENT-FIXING ISLET-CELL ANTIBODY. A positive immunofluorescence reaction.

2-B-ii. Gastric parietal-cell and adrenal-steroidal-cell antibodies

The immunofluorescence method used was identical to that described for islet-cell antibody except that the substrate was unfixed human group O gastric mucosa or adrenal cortex. Positive reactions, graded as for ICA, were shown by characteristic patterns of fluorescence (Figure VI-3).

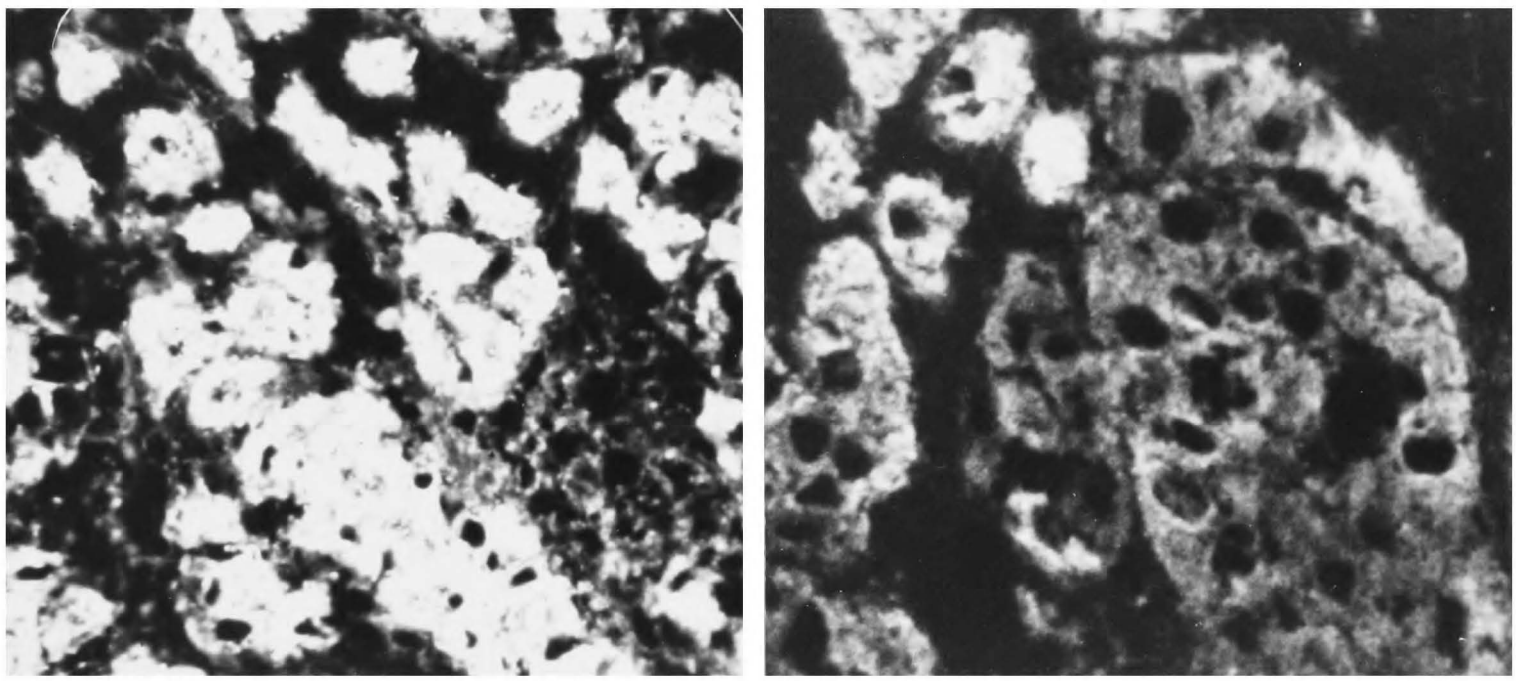


FIGURE VI-3. (Left) GASTRIC PARIETAL-CELL and (right) ADRENAL STEROIDAL-CELL ANTIBODIES. Positive immunofluorescence reactions.

2-B-iii. Thyroid microsomal and thyroglobulin antibodies

Titres were estimated using commercial haemagglutination kits (Wellcome Reagents) of formolized red blood cells coated with the appropriate antigen. Reciprocal titres of at least 400 (20^2) for microsomal and 20 for thyroglobulin were counted positive.

2-B-iv. Antinuclear antibodies

It was not the original intention to test for these, but nuclear fluorescence was frequently apparent on cryostat sections of pancreas during the immunofluorescence test for islet-cell antibody (see above), and was reported by brightness and by pattern of distribution: diffuse, speckled or membranous (Figure VI-4).

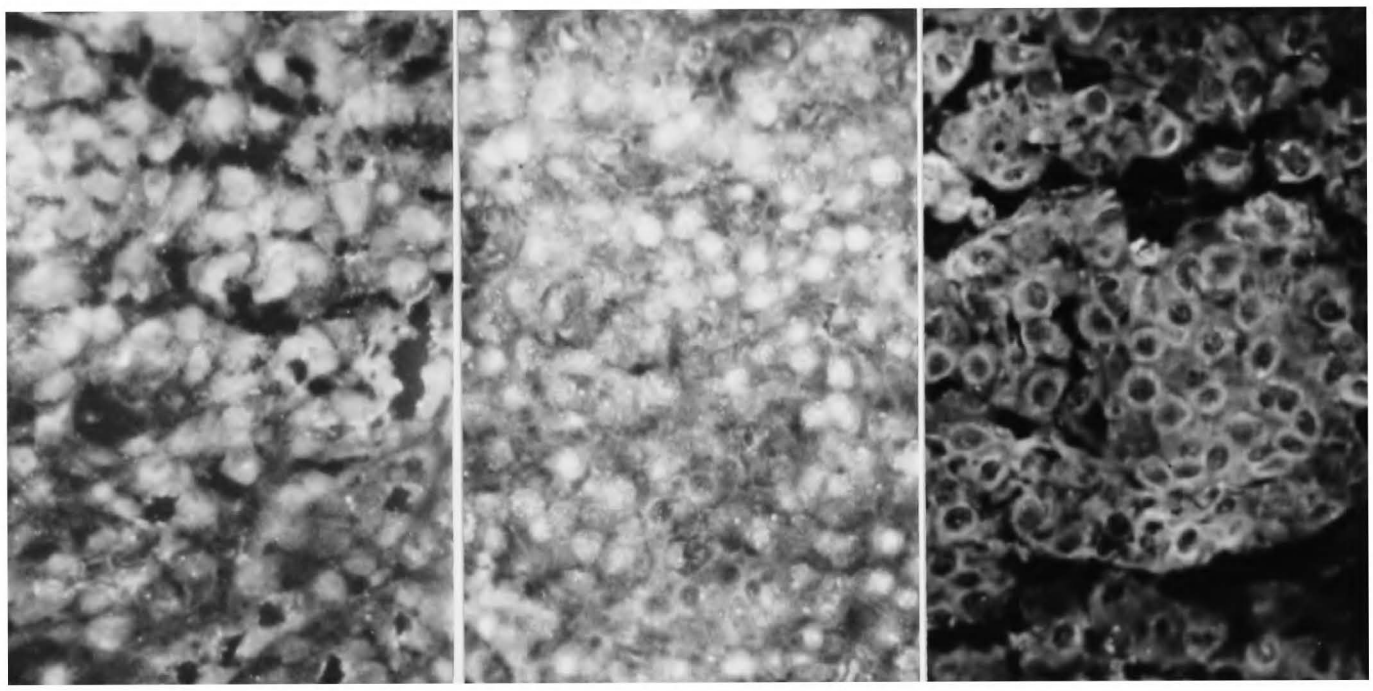


FIGURE VI-4. ANTI-NUCLEAR ANTIBODIES. Diffuse (left), speckled (middle) and membranous (right) immunofluorescence patterns on sections of human Group O pancreas.

2-C. VIRAL SEROLOGY

Samples taken on entry to the study were tested at the Public Health Laboratory, West Park Hospital, Epsom, by courtesy of Dr. DR Gamble. In all families for whom samples were available at the beginning of 1979 titres were determined by standard techniques for antibody to poliovirus types 1, 2 and 3 (neutralization titres), Coxsackie B virus types 1-5 (neutralization titres; Gamble et al, 1969a), and rubella virus (haemagglutination inhibition titres). Complement fixation titres to mumps virus were attempted in a few families, but non-specific anticomplementary activity was such a frequent and marked feature of these sera that this was abandoned.

Serial samples of serum from all individuals followed up were stored at -20°C for future serial analyses.

2-D. METABOLIC INDICES

2-D-i. Urine glucose

In most cases families tested urine with the reagent normally used

by the diabetic member(s), either Clinitest for reducing substances or, more commonly, Diastix for glucose (Ames, Stoke Poges; some supplies were kindly donated by the manufacturers).

2-D-ii. Blood glucose

Venous samples were taken into fluoride-oxalate specimen containers and stored at 4°C until analysis up to three days later. Blood glucose was estimated by automated glucose-oxidase methods in hospital routine laboratories at St. Bartholomew's Hospital or at King Edward VII Hospital, Windsor. In many cases on-site screening was also performed using fresh venous blood and a reagent-strip method, generally Dextrostix (Ames, Stoke Poges) with a 'Glucochek' reflectance meter, but latterly by visual reading of BM-Test Glycémie 20-800 (BCL, Lewes). Neither screening method was sufficiently accurate to be used in statistical analyses.

2-D-iii. Glycosylated haemoglobin

Total HbA_1 was estimated by Dr. S. Welch (Welch & Boucher, 1978) using microcolumn chromatography under strict temperature control (±0.5°C), on undialysed haemolysate from venous blood taken into EDTA and stored at 4°C. Each available family member was tested at least twice. All results on samples stored for more than 10 days were omitted from the analysis since a few such results early on in the study were found to be unexpectedly and non-reproducibly high, even though other workers have reported no significant deterioration with storage up to 28 days (Gabra & Mitchell 1981). This inconsistency may relate to the fact that the latter study involved large volumes of blood in a sealed bag, whereas here the quantity of blood in each specimen tube was small in comparison with the volume of air.

At the time of these assays (1979)the importance of labile adducts appearing on chromatography with stable HbA_1 was not appreciated (Svendsen et al, 1979). However the usual practice of storing diluted haemolysate at 4°C overnight before analysis would probably have allowed dissociation of labile adducts.

2-D-iv. Islet B-cell secretory capacity

In a few subjects B-cell secretory capacity was assessed by measurement of plasma glucose, insulin and C-peptide responses to a standard breakfast (65gm carbohydrate, partly refined; 20 gm fats; 17 gm protein) followed by intravenous glucose (50gm) after an overnight fast. Plasma insulin (Morgan & Lazarow, 1963) and C-peptide (Heding, 1975) were kindly estimated at St Bartholomew's Hospital by Dr. S. Medbak. Detection thresholds and coefficient of variation were respectively 1mU/l and 6.2% for insulin, 0.05 nmol/l and 7.7% for C-peptide.

2-E. PREVENTION AND DETECTION OF ERRORS OF SUBSTITUTION

Because of the possibility of inadvertent substitution of one sample for another during simultaneous handling of fairly large numbers of samples under field conditions and subsequent processing, blood sample tubes for each family were labelled with both names and identity codes and were laid out in order of code number before venesection, the name being checked with each subject. Serum samples were likewise double-labelled with identity codes, date and the ordinal number of the visit to the family. At laboratory level the usual safeguards applied; in addition the fact that most subjects were tested at least twice for HLA and autoantibodies, and the simultaneous testing of whole families, meant that inspection of results allowed identification of the very rare occasions when a switch had occurred.

Any antibody detected in only one (and not the latest) of several samples tested in the same individual is always counted negative, in case a substitution error is responsible. Any doubt over HLA results after two samples was resolved by a third; only one such error of substitution was detected.

PART VI — CHAPTER 3

DATA HANDLING & STATISTICAL ANALYSIS

CONTENTS

A Manual records · 171
i Identity codes · 171
ii Card index · 171
iii Results Books and flow sheets · · · · · · · · · · · · · · · · · · · 171
iv Other records · 171
v Limitations of manual methods · · · · · · · · · · · · · · · · · · · 172
B Computer-based data handling: 'GENOBASE' · · · · · · · · · · · · · · 172
i Background · 172
a Requirements · 172
b Approach · 173
ii Computer facilities used · 173
a Hardware · 173
b Software · 174
c Personnel · 175
iii Outline of structure of GENOBASE · · · · · · · · · · · · · · · · · 175
iv Structure of data array · 177
a General, and HLA genotype record · · · · · · · · · · · · · · · · · 178
b Non-HLA genetic data · 181
c Baseline Autoimmunity record · · · · · · · · · · · · · · · · · · · 181
d Diabetes record · 181
e Baseline Virology record · 181
f Follow-up records: autoimmune and clinical data 182
v Data entry · 182
vi Summarized display of family data · · · · · · · · · · · · · · · · · 184
vii Analysis facilities and output files · · · · · · · · · · · · · · · 185
a Definition of variables · 185
b Selection · 186
c Output · 187
viii Security · 187
ix Documentation · 188
x The future · 189
C Statistical analyses · 189
i Analysis of frequency data · 189
a Two-dimensional contingency tables · · · · · · · · · · · · · · · · 189
b Multi-dimensional contingency tables · · · · · · · · · · · · · · · 191
ii Continuous and quasi-continuous variables · · · · · · · · · · · · 191
iii Specialized (genetic) and miscellaneous analyses · · 192

3-A MANUAL RECORDS

Throughout the study written records were kept in several forms. Every item of information appeared in at least two physically separate places secure against unauthorized entry.

3-A-i. Identity codes

Consecutive families entering the study were given consecutive identity numbers ('IDF'). Within each family individual subjects were given identity numbers ('IDS') according to the following system: the father and mother of the sibship including the proband were numbered 1 and 2 respectively, and members of the sibship, living or not, were numbered consecutively from 3 in order of live birth. Further conventions adopted for other members of families are not important here. The same system was used for families in the UK (British Diabetic Association) Multiplex Family Study, for miscellaneous families HLA-typed but not fulfilling ascertainment criteria for these studies, and for newly-diagnosed diabetic individuals and their families.

3-A-ii. Card index

A record card was completed for each family during the ascertainment process and was taken on field visits. Clinical and family data, certain laboratory results useful for reference in the field, and a note of visits made and samples obtained were recorded on the card.

3-A-iii. Results Books and flow sheets

All information obtained was collated and recorded in Results Books organized by family and by individual subjects. Results were also summarized by family in the form of flow sheets to facilitate analysis.

3-A-iv. Other records

Laboratory records or photocopies thereof were kept to facilitate error checking.

3-A-v. Limitations of manual methods

There were three main problems. First, the volume of data generated was very large. Nearly 800 individuals were included, each producing up to 120 items of information on entry to the study as well as a further 10-20 items at every follow-up visit, giving an estimated total of over 100,000 items to be included here, with further expansion planned.

Secondly, the genetic information was complex, allowance having to be made for expaternity, ambiguous haplotypes due to homozygosity in a parent, genetic recombination, incomplete typing of some individuals in whom results could nevertheless be deduced from relatives, and other problems.

Thirdly, it was necessary to count, and extract quantitative data on, individuals classified not only by their own attributes but also by those of their relatives. For example, poliovirus titres were to be examined in non-diabetic, full siblings of diabetic probands, in relation not only to their own age, sex and vaccination status, but also to the concordance between their own and the probands' haplotypes (RESULTS, VIII-5-B-i-a). Again, linkage disequilibria in HLA haplotypes occurring in probands were to be compared with those in haplotypes occurring in the families but not in any diabetic family members (VIII-RESULTS-1, 5), necessitating a search through all members before any given haplotype was counted.

Clearly, a computer-based system was needed.

3-B. COMPUTER-BASED DATA HANDLING: 'GENOBASE'

3-B-i. Background

3-B-i-a. Requirements

The following requirements were identified:

1. To store electronically HLA and other genetic information on members of 'nuclear families', together with other baseline data to be analysed in conjunction with the genetic data such as measures of autoimmunity, family history, clinical details of any diabetes, and possible indices of

immune response. Complexities of genetic information to be accommodated exhaustively (see above), and editing facilities to be provided.

2. To store as soon as available information obtained at consecutive visits to families during the prospective study, and to make such follow-up records available for viewing and editing.

3. To extract stored data rapidly and selectively according to any of a wide range of criteria determined by the operator, these criteria often involving attributes not only of the individual concerned but also of other family members; and to provide such output in a form suitable for analysis by manual and computer-based techniques.

4. To display in summarized form the data on a given family.

5. To permit future incorporation of further classes of information as well as additional families.

3-B-i-b. Approach

Professional advice was sought of Dr. D Ingram of St. Bartholomew's Hospital. No ready-made database system available could provide the necessary flexibility and complexity of analysis. The author and Dr. Ingram therefore jointly developed a completely new system, GENOBASE, which has grown in scope over several stages since 1980-1981. Many of the analyses reported here were done with earlier versions; major changes in 1985-6 are noted in the following description.

3-B-ii. Computing facilities used

3-B-ii-a. Hardware

GENOBASE was developed and is used on DEC 11/23, 11/44 and 11/73 minicomputers (Digital Equipment Corporation, Boston, Mass., USA) at St. Bartholomew's Hospital. Mass storage is provided by large DEC 'winchester' hard disk drives. Terminals are, in general, DEC VT/100 types or similar. A DEC Professional 350 microcomputer with a winchester drive has been used in terminal-emulation mode in conjunction with the

11/44 minicomputer, and could indeed run GENOBASE independently. The system could be transported to other types of computer, given sufficient dynamic memory (about 100 kilobytes user-accessible) and rapid-access mass storage (>1 megabyte, depending on the amount of data to be stored). However, minor modifications to the display might then be necessary.

Certain complex algorithms were developed and tested by the author using a Casio FX-702P pocket computer / programmable calculator, before incorporation in the main suite of programs.

Output from GENOBASE is entirely in the form of standard characters rather than graphics. Almost any printer can therefore be used, or data may be exported to other computers. For example, a packet-switching data link via the PAD modem and the Joint Universities Network (JNT) enabled some large-scale multivariate analyses to be performed on the GEC OS4000 mainframe computer (Euclid) at University College, London, under control from a terminal at Barts. Tabulated output data were also transferred to floppy disk in IBM format for further processing on the personal computer used by the author, an Opus PC-II IBM-PC-compatible microcomputer with 1024 kbyte RAM (384k configured as a Ramdisk) and two 360k floppy disk drives, driving a Panasonic KX-P1080 Epson-FX-compatible 9-pin dot-matrix printer.

3-B-ii-b. Software

Source code for GENOBASE was written in the FORTRAN-IV programming language, and compiled. Those algorithms developed and tested by the author were written and 'debugged' in BASIC before translation to FORTRAN.

An advanced word-processing system, LEX (Ace Microsystems Ltd., London) was used extensively during development and greatly facilitates operation of GENOBASE. It can readily be used to process data files, source code and command files forming input to or output from other software systems, including GENOBASE and statistical packages.

Direct data transfer between various computers was achieved using

the Kermit protocol, many versions of which are widely available free of charge.

Statistical processing of output data from GENOBASE was done mainly with the aid of MINITAB and GLIM (see 3-C) on the Barts computers and on Euclid. Output from these packages and from GENOBASE itself was further processed under LEX or line-editors on these larger computers, and using a variety of software on the Opus microcomputer: PC-Write version 2.55, a versatile and very fast word processor (Quicksoft, Seattle, USA); PC-Outline, an even faster outliner and wordprocessor with excellent import-export facilities (Softworks Development, Mountain View, California, USA); and Ability, an integrated package including spreadsheet, simple graphics, database, wordprocessor and communications (Migent UK, London).

3-B-ii-c. Personnel

Dr. D Ingram is Senior Lecturer in Medical Physics and Computing at St. Bartholomew's Hospital. Mr. W. Howard, a student of computing, assisted him in the latest (1985-6) phase of development of GENOBASE.

3-B-iii. Outline of structure of GENOBASE

The structure of the system is summarized in Figure VI-5. Data are stored on disk in a series of short records of various types, as indicated in Figure VI-5. At any time, further records may be added for any individual subject, or additional subjects may be incorporated. The data on the disk may be re-organized or 'condensed' by means of a utility program.

Various types of analysis and output are possible. In all cases all the data required for a given family are first read by the system from the disk records and transferred as a whole to a temporary (buffer) area of dynamic memory, where they form a strictly-organized array (Fig VI-5; 3-B-iv). Data from this array may then be displayed on the screen (3-B-vi), transferred non-selectively in tabular form to output files suitable for printing or processing under LEX, or searched by the analysis program (3-B-vii).

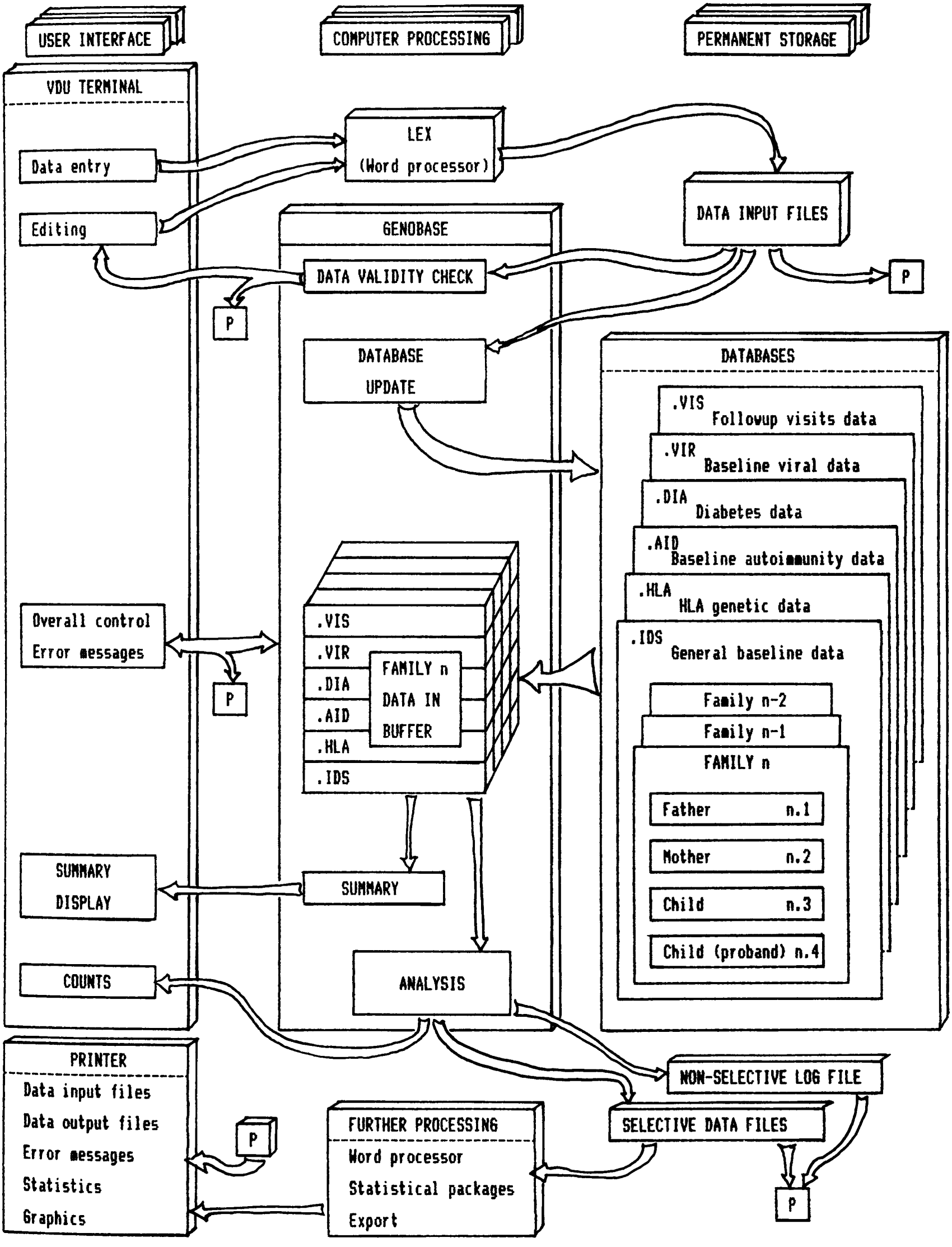


FIGURE VI-5. GENOBASE: OVERALL STRUCTURE AND FACILITIES

GENOBASE is a menu-driven system: any of the above options, with the exception of the 'condense' utility, may be selected from options displayed on the screen.

3-B-iv. Structure of data array

Figure VI-6 explains in more detail the structure of the family data array as it exists in the buffer area of memory during processing. It can be envisaged in three dimensions: dimension x corresponds to the position of the variable concerned (e.g. thyroid microsomal antibody titre) within a disk record, dimension y gives the type of record (not all types need be present in buffer together, as shown here), while z gives the IDS of the individual whose record it is. In effect the family number (IDF) represents a fourth dimension, as families are fed successively through the buffer during processing.

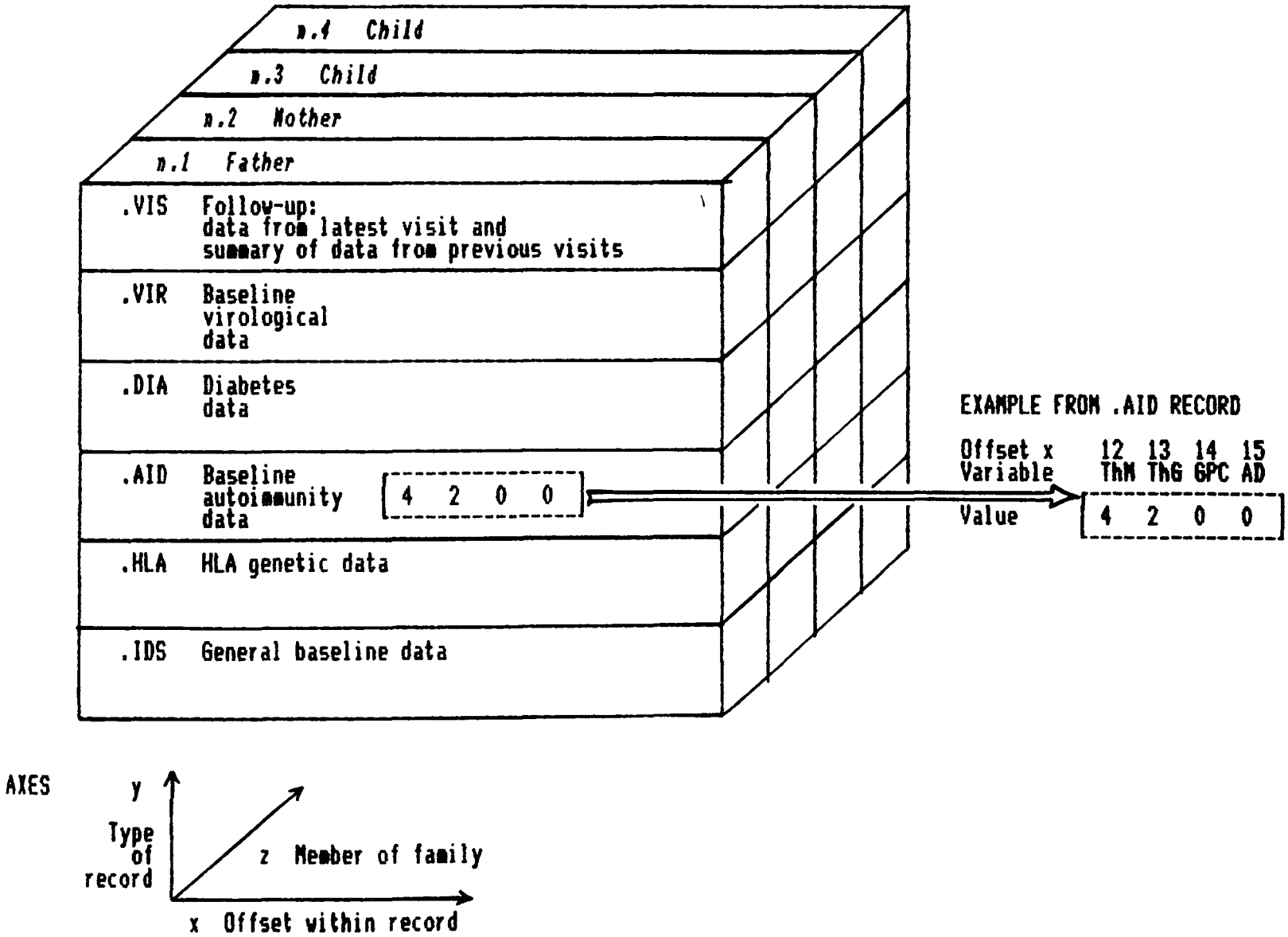


FIGURE VI-6. GENOBASE: SCHEMATIC REPRESENTATION OF FAMILY DATA ARRAY. Part of the contents of an autoimmunity record is shown to illustrate how data are organized. KEY: ThM thyroid microsomal Ab, ThG thyroglobulin Ab, GPC gastric parietal-cell Ab, AD adrenal steroidal-cell ab.

Users have access to full documentation of the way in which information is represented in the database. Certain features of particular importance or interest are explained below.

3-B-iv-a. General, and HLA genotype record

Exclusion codes

The 'General' record includes a series of exclusion codes for rapid identification of families and individuals who may commonly need to be excluded from analyses because of expaternity of the proband, ambiguity of HLA haplotypes, or non-availability for testing.

Dates

Date of birth, like most other dates in GENOBASE, is stored in 'julian' form as number of days since an arbitrary zero (in fact 31st December 1600 by the current calendar), so as to facilitate arithmetical manipulations. During input and output procedures dates are converted between julian and standard forms (day-month-year) by a subroutine (JDAT3) developed by the author, generally applicable and with full allowance for leap-years (Appendix 5-a).

Other items in the General record concern family structure; for example, twins are identified as identical or non-identical.

HLA genotypes - outline

The HLA genotype record has the most complex structure. The genotype of a parent is entered as two full haplotypes, and that of a child in general as coded haplotype abbreviations (a,b,c,d) to simplify entry of data and assessment of haplotype concordance within sibships. The full genotype is however 'reconstituted' during processing, using algorithms largely written and tested by the author before incorporation in GENOBASE. Provision is made for non-straightforward cases: intra-HLA and HLA-GLO crossovers, 'deduced' data, equivocal assignment of one or more specificities to haplotypes, expaternity, and 'splits', as explained below. GLO is not treated as part of the haplotype but is entered

separately for each individual, because crossovers between this locus and HLA are common.

HLA-region loci

Loci accommodated are as follows; only those marked '*' are included in the present analysis.

HLA Class I	Class III	Class II	Nearby
A B Cw	Bf C2 C4A C4B	Dw DR DR52/53 DP DQ	GLO
* * *	* *	(*)	*

Expaternity

Occasionally HLA-typing shows expaternity of a child. In such a child the paternal haplotype abbreviation is recorded as "e" (or "f" or "g" in families multiply afflicted), and the full paternal haplotype actually present in the child is entered in the child's record on disk. During processing the system automatically uses this, rather than one of the haplotypes of the father of the rest of the sibship, to synthesize the reconstituted genotype. Presence of expaternity is also flagged in family exclusion codes.

Recombinant haplotypes

Intra-HLA recombinants are accommodated in the same way as expaternity, except that a different code is used for the haplotype abbreviation according to the position and sense of the crossover.

Crossovers between GLO and HLA-B are coded in special variables in each child's record. Coding takes account of the fact that, while there may clearly be such a crossover in a family, it may not be possible to determine whether it is paternal or maternal, nor even in which child it occurs. Non-informative meioses are coded as such.

Deduced genotypes and equivocal assignment to haplotypes

For each locus, two 'qualifying bytes' are available so that attention may be drawn to limitations of the genetic information available on an individual, and so that analyses may if appropriate

exclude such individuals. The first is used to indicate that the alleles recorded at any locus so marked have only been deduced as present since they form part of the same haplotype in other family members, the individual concerned not having been tested. In input and output files and displays a "*" appears under the data at the locus concerned.

Sometimes an individual is known to be heterozygous at a locus, but the two alleles cannot be unequivocally assigned, one to each haplotype. For the purpose of recording the data the alleles are arbitrarily assigned to haplotypes, and the second qualifying byte is used to indicate that assignment is equivocal: an "=" sign appears at the locus concerned.

These qualifying bytes are a feature of the latest version of GENOBASE. The previous system was more economical on storage space but conceptually complex. Algorithms defined by the author were used for automatic conversion from old to new systems.

Homozygosity

Where a parent is homozygous (or presumed to be so in that only one specificity is detectable and appears in all the children) at all loci tested, the degree of haplotype concordance between proband and other children is ambiguous. In such cases the proband is given the haplotype abbreviation "a" (or "c" if the mother is homozygous), while the other children are given "1", their reconstituted genotypes deriving from the parent's "a" (or "c"). In addition exclusion codes are set to identify such families and individual parents.

Splits, and changes in HLA nomenclature

In general the latest version of GENOBASE stores actual (i.e. non-coded) genetic data according to the standard nomenclature in force. If available sera detect 'splits' (see III-6-B-i and Fig. III-4) not recognized at the most recent Histocompatibility Workshop, temporary new codes are assigned. When recognized at a later Workshop the relevant data can easily be recoded. Analyses required in terms of the parent broad antigen can take account of temporary or recognized splits.

3-B-iv-b. Non-HLA genetic data

Further records may be added for storage of data on genetic systems not linked to HLA, such as Gm (three loci, two haplotypes for each individual) and the insulin-gene restriction-fragment-length polymorphism. These are not described here.

3-B-iv-c. Baseline Autoimmunity record

The following measures of immunological status at entry to the study are included in this AI record:

1. History of allergy or probable allergy, coded according to clinical picture described (asthma, eczema, hay fever, contact allergy, drug or food allergy, urticaria or angioneurotic oedema, or combinations of these);
2. Autoantibodies (islet-cell Ab by conventional and complement-fixing techniques, gastric parietal-cell and adrenal steroidal-cell antibodies, all coded for brightness (and ICA for distribution) of fluorescence; thyroid microsomal and thyroglobulin Ab (log haem-agglutination titres); antinuclear Ab detected by immunofluorescence test on pancreas, coded as for ICA;
3. Autoimmune endocrine and other clinical conditions of interest in this contest, and the duration of each at entry to the study;
4. HLA antibodies detected during screening for typing sera, coded for class of antigen reacting (mothers only).

3-B-iv-d. Diabetes record

The date (in julian form), month of year, age and duration of symptoms at the time of diagnosis and the type of diabetes are recorded. A summary of family history of diabetes and other relevant conditions in first, second and third-degree relatives is also included. There is scope for addition of further information.

3-B-iv-e. Baseline Virology record

Reported polio vaccination status and log-transformed antibody

titres to certain viruses at entry to the study are stored here (see VI-2-C).

3-B-iv-f. Follow-up records - autoimmune and clinical data

The following information obtained at each successive visit during the prospective study may be entered into a Follow-up record as soon as results are all available: date and arbitrary code number of visit to family; up to six 'events' since the previous visits, coded by type of event (e.g. 'C' for colds and other upper respiratory tract infections); numbers of negative and of positive urine tests since the previous visit; blood glucose and glycosylated haemoglobin results, if tested; and autoantibodies as above (3-B-iv-c). Duration of breast-feeding of children in infancy is included in the first record (here rather than elsewhere for historical reasons).

For routine analyses in relation to genetic information, detailed analysis of all follow-up records would be intolerably slow. Therefore, whenever a new follow-up record is entered, a 'summary' record for the same individual is automatically updated to include maximum and minimum values of continuous or quasi-continuous variables and cumulative sums of counts. Only the latest follow-up record and the summary record are normally transferred to the buffer during processing for analysis or summary display (baseline autoimmune results also being available; 3-B-iv-c). The whole sequence of follow-up records for an individual may be examined as a special option.

3-B-v. Data entry

Formerly an interactive routine was used for entering baseline data, but now batch entry of tabulated data is used excusively. LEX may be set up to define automatically, in the form of 'rulers', any format of tab stops required to assist keyboard entry of tabulated data. Different types of documents, each defined by its own three-character extension to the filename of the documents, may have different rulers (numbering up to five per document type). For use in this context, each type of record in GENOBASE has its corresponding LEX document format. Figure VI-7 shows the rulers and column-headings used, with examples of data. The complex

Column number	5	10	5	20	5	30	5	40	5	50	5	60	5	70	5	80
Database type	*.IDS General baseline data															
Column headings	IDF	IDS	Calendar date of ..... Birth			Entry to study			Con-sult-	NT	EX 1-4	F:NChil M:NFPreg C:Sex		F:NPrd M:Twins C:NTwin		
Ruler defining format	.####	##	##	##	####	##	##	####	T	T	T	##	##	##	##	##
Example of data	123	1	12	12	1938	25	3	1979		0	0000	2		3		
	123	2	21	3	1938	25	3	1979		0	0000		1		0	
	123	3	17	7	1963	25	3	1979		0	0000			1		
	123	4	4	9	1965	25	3	1979		0	0000			2		
Database type	*.HLA HLA GENETIC DATA															
	FULL HAPLOTYPE: A,B,C,D in parents, E in child (Ruler 1)															
Column headings	IDF	IDS	Data	A	B	Cw	Bf	C2	C4A	C4B	Dw	DR	D5	DP	DQ	GLO
Ruler defining format	.####	##	T	##	##	T	T	T	##	##	##	##	T	T	T	T
Example of data	123	1	A	3	B	7	F	C	20	0	3	3	2	1		2
	123	1	B	2	62	3	S	N	3	48	9B	9B	N	N		1
	123	1	QD								*	*	*	*		
	123	1	QE					=								
Database type	SUMMARIZED GENOTYPE in child (Ruler 2)															
Column headings	IDF	IDS	Data	PHAP		PGLO	PGLOX					MHAP		MGLO	MGLOX	
Ruler defining format	.####	##	T	T		T	T					T		T	T	
Example of data	123	3	P	a		2	0					c		1	0	
	123	4	S	a		1	X					d		2	0	
Database type	*.AID Baseline autoimune data															
Column headings	IDF	IDS	Date of record			Hyp HLA Ab	Autoantibodies ICA 123	Th MT	GA	ANA 12	AI dis. (Up to six)	Durations of respective AI diseases				
Ruler defining format	.####	##	##	##	####	T	T	T	T	T	T	##	##	##	##	##
Example of data	123	1	25	3	1979	H	000	20	20	000	t	5				
Database type	*.DIA Diabetes data															
Column headings	IDF	IDS	Date of diagnosis			WksD	Fam.Hx.Diabetes Sib	Par	R2	R3	Fam.Hx.Autoimmunity DisR1	AbR1	DisR2	DisR3		
Ruler defining format	.####	##	##	##	####	##	T	T	T	T	T	T	T	T		
Example of data	123	3	15	7	1973	4	0	0	1	0	t	1	PT	t		
Database type	*.VIR Virological data															
Column headings	IDF	IDS	Date of sample			Data	Po vac	Polio titre 1	2	3	Coxsackie B titre 1	2	3	4	5	Rubella titre
Ruler defining format	.####	##	##	##	####	T	T	##	##	##	##	##	##	##	##	##
Example of data	123	4	25	3	1979		1	3	3	1	0	1	0	4	0	2
Database type	*.VIS Follow-up visit data															
Column headings	IDF	IDS	Nvis	Date of visit			Brst -fed	Events 123456		UNeg	UPos	Glucose Value	M	HbA1	ICA 123	Th MT	GA	ANA 12
Ruler defining format	.####	##	T	##	##	####	##	T		##	##	##	T	##.#	T	T	T	T
Example of data	123	4	1	25	3	1979	0								000	00	0	00
	123	4	2	13	6	1979		CMCCXC		3	0			9.9	000			00
	123	4	3	15	9	1979		C							000			00

FIGURE VI-7. GENOBASE: DATA INPUT DOCUMENT FORMAT. Data are entered using LEX (word-processor). At any time column headings may be called on to the screen. Obscure headings are explained as follows. *.IDS records: IDF = family no.; IDS = individual no.; NT = not tested; EX1-4 = exclusion codes (see text); NChil (father only) = no. of children; NPrd = IDS of proband; NFPreg (mother) = no. failed pregnancies; Twins = no. sets of twins; NTwin (child) = IDS of co-twin. *.HLA records: Data = type of record (A-E for full haplotype, P(roband) or S(ib) for summarized genotype, QD to tag data deduced from other family members, QE to tag equivocal assignment to haplotype); D5 = DR52/53; PHAP = paternal haplotype, MGLOX = maternal HLA-GLO recombinant. *.AID: Hyp = hypersensitivities; GA = gastric and adrenal antibodies. *.DIA: WksD = weeks of symptoms before diagnosis. *.VIS: UNeg = no. negative urine tests; Glucose M tags reflectance-meter result.

nature of the HLA data necessitates simultaneous use of several different formats, the variable 'DATA' indicating that in force on a given line. Thus a, b, c and d indicate the four parental haplotypes in a family, e indicates an expaternity or other full haplotype in a child, P and S indicate summarized haplotypes in children (probands and siblings respectively), while QE and QD define qualifying bytes.

Advantages of this method of data input include the following:

1. If a data column is entered by use of the TAB key, a '##' in the ruler at the head of the column allows only numerical data to be entered, helping to reduce keyboard errors.
2. Entry of repetitive data (such as dates when several families were visited on the same day) is enormously speeded by the word-processor's text-copying and manipulation facilities.
3. The tabular layout of the document helps to show up errors.
4. A document may be printed out (or viewed on the screen) to be checked for accuracy of data, and errors corrected, before risking corruption of the database.

The next stage in data entry is to run the document created under LEX through the data-checking facility included in GENOBASE. This verifies, for example, that dates of birth are chronologically and biologically feasible, that the values of variables fall within permitted ranges, that probands are indeed diabetic, and that any new follow-up documents have dates later than those already in the database. Only then are the new records added to the database.

Some manipulation of data takes place during data entry. Thus, in the Diabetes record, only the date of diagnosis is actually entered into the system; the computer stores the month of diagnosis separately, converts the date to julian form, and uses the date of birth already present in the database to calculate age at diagnosis.

3-B-vi. Summarized display of family data

A 'family tree' may be displayed to show the HLA-genetic structure of the family, the contents of parental haplotypes, diabetic status, sex

and presence of Baseline Autoimmunity records, as well as other information. The operator may select various classes of information to be displayed in more detail in the lower section of the same screen for any individual. Figure VIII-1 (RESULTS, VIII-1-B-iii) is derived directly from screen-dumps of the summarized genetic data on four families.

3-B-vii. Analysis facilities and output files

According to the requirements defined above (3-B-1-a) the operator must be able to specify precisely both the criteria by which individuals are selected for analysis, and the data to be extracted from the records relating to those individuals. Selection criteria often involve long or complex series of logical tests on the values of some of the hundreds of variables in the database.

The algorithms used by GENOBASE to direct such analyses are defined in 'command files', written by the operator under LEX. A simple, annotated example is shown in Figure VI-8. This and the following description relate to the search and analysis system used up to 1985; since then W Howard has completely re-written this part of GENOBASE to give more flexibility and ease of use.

Command files for GENOBASE are written in a symbolic language specially designed for precision and economy. Each command file consists of several sections delineated by colons (:) and full stops (.), the function of each section being explained in the Figure.

3-B-vii-a. Definition of variables

In the pre-1986 version, variables are identified in command files by their position in the data array. First is a symbol for family member, A representing the member currently under examination, F the father, M the mother and P the proband. Next is the 'offset', a number derived from the 'y' and 'x' dimensions, which is tabulated for reference (in fact, |offset| = 100y + |x|). Negative sign indicates a byte variable, absence of sign an integer variable counting as two bytes.

```
:C      IRPS11.CMD
:C      POLIO VACCINATION & TITRES
:C      'INCLUDED' NON-DIABETIC HLA-IDENTICAL SIBLINGS OF PROBANDS,
:C      BY AGE AT ENTRY
:C      PRINTED: SEX, AGE ENTRY, POLIO 1-3
```

	:C indicates a comment
:F	Range of families included:
1 132	IDF = 1 to 132
.	End of 'range' section.
:SA	'Select' individuals satisfying the following group of tests, 'ANDed' together:
F-13 - !P	Family expaternity exclusion code not 'P' ...
F-13 - !T	... not 'T' (both show expaternity of proband).
F-15 - !H	Family genotype exclusion code not indicating
F-15 - !D	ambiguity of haplotypes due to homozygosity etc.
A2 > !2	Individual identity code (IDS) >2 (i.e. child)
A2 - F10	IDS not that stored (under father) as proband's IDS
A-11 - !0	Not diabetic
A-13 - !X	Individual expaternity exclusion code not 'X'
A-12 = !0	Tested for HLA etc.
A-56 = !1	Vaccinated against polio
A29 > !-1	Polio titres tested
A-49 = P-49	Paternal HLA haplotype same as that of proband
A-52 = P-52	Maternal HLA haplotype same as that of proband
:P IRPS21.DAT	'Print' data on selected individuals in file named.
A9	Sex
A11	Age at entry
A29	Polio 1 titre
A30	2
A31	3
.	End of 'print' statement
::	End of command file

FIGURE VI-8. GENOBASE: A COMMAND FILE FOR THE DATA ANALYSIS SYSTEM. Text outside the frame is explanatory and does not form part of the command file itself.

A more 'friendly' identification of variables by name is a feature of the 1986 version.

3-B-vii-b. Selection

The :S section (selection of individual family members) is based on Boolean logic applied to tests of the values of variables in the buffer array. Each test determines whether a variable is identical to (=), different from (-), less than (<) or greater than (>) either another variable or a value defined after '!'. Such statements may be 'ANDed' together in groups which may then be 'ORed' with further groups, or vice-versa.

A further facility not included in this example is that any of up to five 'flags' may be set or cleared when a group of logical tests is satisfied, the state of each flag being evaluated after all members of a family have been tested and a count incremented if the flag is 'set'.

3-B-vii-c. Output

Total counts obtained from analyses are displayed on screen, while data selectively transferred to a file are automatically tabulated in a form suitable for viewing and editing under LEX or input to a statistics package.

Data representing contents of complete records (e.g. all HLA data, or all follow-up records) on single or multiple individuals may be transferred to files in the same format as documents used for data entry (3-B-v), to facilitate updating.

All data from major sections of the database on selected individuals may be dumped into a log file for printing. Figure VI-9 is a photograph of printout representing all available data concerning the 166 families studied here. The large volume of information stored is obvious.

3-B-viii. Security

Access to GENOBASE is protected by password and account number, and is restricted to individuals approved by the Steering Committee of the Barts-Oxford Family Study (the expanded derivative of the Barts-Windsor study), of which the author is a member. Only summarized output data may be exported to other systems.

Data may be used only for research, and may not be published in any form which would reveal confidential information and would allow identification of families from, for example, sex and dates of birth of individuals - a restriction which applies to this Thesis.

Data within the system are not readily attributable to individuals by name, address or any other means, nor would the vast majority of the

FIGURE VI-9. GENOBASE: THE VOLUME OF DATA. Photograph of printout from log files. On the left are the follow-up data for 1978-1980, printed at six lines per inch, each line representing one .VIS record (see Figure VI-7). On the right are the remaining data on 162 families, in summarized form. Total length of printout approximately 60 metres.

information be comprehensible (or, indeed, of any interest) to any but the most highly-informed person in this field of study. Thus the risk of breach of confidentiality by 'hackers' is minimal.

GENOBASE has been registered as required by the Data Protection Act (1984).

3-B-ix. Documentation

The facilities of GENOBASE and the coding of data are fully documented by the author and W Howard.

3-B-x. The future

The structure of GENOBASE allows future addition of further types of data records, and of effectively unlimited numbers of families. Programming support is currently available, and it is hoped that a 'user-friendly shell' can be built around the Analysis facilities, so that the more simple and frequently-needed analyses may be more readily accessible to users than at present.

3-C. STATISTICAL ANALYSES

3-C-i. Analysis of frequency data

3-C-i-a. Two-dimensional contingency tables

Two-by-two cross-classification (contingency) tables were tested for significance of deviation from random distribution by standard techniques: Fisher's exact probability test where there were low 'expected' frequencies in one or more cells, or the Chi-squared test with Yates' correction (Everitt, 1977). Woolf-Haldane approximate relative risk (RR) (with confidence limits and correction for bias) and population attributable risk are calculated as described in the Literature Survey, III-6-B-v. Calculations related to predictive value of a test T with respect to a disease D were performed as follows, with reference to the decision matrix giving numbers of individuals in a random sample (McNeil, Keeler & Adelstein, 1975; Møller-Petersen, 1985):

	T+	T-
D+	a	b
D-	c	d

Pv (Prevalence of D) = (a+b) / (a+b+c+d)
Qv (Prevalence of non-D) = 1 - Pv
Specificity = a / (a+b)
Sensitivity = d / (c+d)
L_{pos} (likelihood ratio for positive result, disease pos/neg)
= sensitivity / (1 - specificity)
L_{neg} (likelihood ratio for negative result, disease neg/pos)
= specificity / (1 - sensitivity)

By analogy with RR, the natural logarithms of Lpos and Lneg are approximately normally distributed. Their variances may be calculated by a similar method (see III-6-B-v):

$\mathrm{Var}(\log_e L_{pos}) = 1/a + 1/c + 1/(a+b) + 1/(c+d)$

$\mathrm{Var}(\log_e L_{neg}) = 1/b + 1/d + 1/(a+b) + 1/(c+d)$

and their 95% confidence limits derived in the usual way, the exponential derivatives of these giving the confidence limits of L_{pos} and L_{neg}.

PV_{pos} (predictive value of positive result for disease; probability that an individual who is T+ is also D+)

$$= \frac{Pv}{Pv + Qv/L_{pos}}$$

PV_{neg} (predictive value of negative result for non-disease)

$$= \frac{Qv}{Qv + Pv/L_{neg}}$$

Approximate 95% confidence limits for PV_{pos} and PV_{neg} are derived by substituting the confidence-limit values of L_{pos} and L_{neg} in these equations.

Efficiency = Pv x sensitivity + Qv x specificity

All these calculations are included in the author's BASIC program SPK2X2 (Appendix 5-b). In a few cases, where a very small sample was compared with a very large one, the standard binomial test was used, and the variance of a binomial frequency estimate $\hat{p}$ was calculated by the standard formula

$$\mathrm{Var}(\hat{p}) = \hat{p}(1-\hat{p}).$$

Larger r x c contingency tables were tested for significance by the χ^2 test; as a measure of association Goodman & Kruskal's gamma statistic was used, while order effects were investigated where appropriate by testing for linear trend using a regression method (Everitt, 1977; all available in author's SPKRXC program, Appendix 5-c).

3-C-i-b. Multi-dimensional contingency tables

Three-dimensional tables with up to 24 cells were tested for mutual independence of the three variables and for first-order interactions by the author's Chi-squared program SPKRCL, running on a pocket computer (Appendix 5-d; Everitt, 1977).

Further investigation of these, and analysis of more complex tables, necessitated log-linear modelling techniques available on larger computers in the General Linear Model package GLIM (NAG Publications, Oxford University Computer Laboratory). Only hierarchical models were considered; that is, higher-order terms for interactions between three or more variables were included only if all related lower-order terms were also included. Selection of models to be tested for goodness of fit, and criteria for simplification of data structure by 'collapsing' over a variable, followed the guidelines of Bishop, Fineberg & Holland (1975; p.47, Theorem 2.5-1):

> *'Suppose the variables in an s-dimensional array are divided into three mutually-exclusive groups. One group is collapsible with respect to the u-terms involving a second group, but not with respect to the u-terms involving only the third group, if and only if the first two groups are independent of each other (i.e. the u-terms linking them are zero).'*

3-C-ii. Continuous and quasi-continuous variables

The following standard tests were used as appropriate, almost invariably in the package MINITAB (Pennsylvania State University, USA) running on the Barts computers. Occasionally the author's BASIC programs were used with a pocket or desk computer for comparison or correlation between two small data-sets.

Parametric tests: Student's t-test, analysis of variance, regression, Pearson (product-moment) correlation coefficient.

Non-parametric tests: Mann-Whitney 'U' test; Kruskal-Wallis test; Spearman rank correlation coefficient.

3-C-iii. Specialized (genetic) and miscellaneous analyses

Measures of linkage disequilibrium (/\, r/\) were computed using the formulae given in III-6-B-iii, and their significance assessed using χ^2. Differences in linkage disequilibrium between 'diabetogenic' and 'non-diabetogenic' haplotypes were assessed by multidimensional analysis (see above).

Tests for deviation from Hardy-Weinberg equilibrium were based on the χ^2 statistic, i.e. the sum of the squares of the differences between observed genotype frequencies and their Hardy-Weinberg expectations. For the highly-polymorphic loci each allele was considered separately, the formula (6-1) in III-6-B-ii being used and giving two degrees of freedom. The author's BASIC computer program HARDY1 (Appendix 5-e) was used to convert raw gene counts and phenotype counts into gene frequency, phenotype frequency, the observed and expected genotype distributions, and χ^2.

The observed distribution of diabetic siblings among HLA haplotype concordance groups were tested by χ^2 for goodness of fit to the distribution expected by single-gene recessive and dominant hypotheses, the expected values being obtained from tables published by Thomson & Bodmer (1977). The author's BASIC program TOMBOD (Appendix 5-f) was used to calculate and plot graphically the values of χ^2 for 21 different gene-frequencies over the range 0.00 to 1.00 (see RESULTS, VIII-B-i; Figure VIII-2).

PART VII

SUBJECTS STUDIED

CONTENTS

A Barts-Windsor Family Study · 193
i Numbers of families studied · · · · · · · · · · · · · · · · · · 193
ii Numbers of individuals; sex and age · · · · · · · · · · · · · 194
a Probands · 194
b Siblings · 194
c Parents · 196
iii Details of diabetes in probands · · · · · · · · · · · · · · · 197
a Age at diagnosis · 197
b Duration of symptoms before diagnosis · · · · · · · · · · · · 197
c Season at reported onset of symptoms · · · · · · · · · · · · 197
d Duration of diagnosed diabetes at entry to study 197
B Newly-diagnosed Type 1 diabetic subjects · · · · · · · · · · · · 199
i Numbers and sex · 199
ii Age, season and duration of symptoms at diagnosis · · 199
C Other series providing supplementary data · · · · · · · · · · · 199
i UK (British Diabetic Association) Multiplex Study · · 200
ii Liverpool Study · 200
iii Miscellaneous · 202
D Control subjects · 202

A. BARTS-WINDSOR FAMILY STUDY

A-i. Numbers of families studied

In the first 2.5 years of this study the author recruited 166 families. One proband was first cousin to another, but otherwise they were, to the best of their knowledge, mutually unrelated. Ascertainment rate is considered under VI-METHODS-1.

Expaternity of the diabetic proband was shown by HLA typing in 6 families. These are included only in certain simple analyses (e.g. population frequencies of HLA phenotypes in diabetics) and as a source of 'control' haplotypes. Most family-based analyses therefore include a maximum of 160 families (159 unrelated).

Prospective follow-up is defined as availability of information (including serology) on at least one non-diabetic sibling, obtained at two or more separate visits. Full siblings of the probands in 151 families were followed up. Of the remaining 9 families with full siblings (as opposed to half-siblings), four were 'multiplex' for diabetes and had no unaffected child aged <20 years, while five families withdrew after the first visit because of lack of motivation, dislike of venesection or removal away from the area.

A-ii. Numbers of individuals; sex and age

A-ii-a. Probands

The proband in a nuclear family is here defined as the Type 1 diabetic child first to have been diagnosed.

Of the 166 probands, 91 (55%) were male; the M:F ratio of 1.21 is not significantly different from that of 1.05 in the general population of E. Berks (Office of Population Censuses and Surveys, 1982). Their ages at entry to the study were approximately normally distributed (Figure VII-1), and similar in both sexes. Details are in Appendix 1-A. The six showing expaternity did not differ significantly from the remainder in age or sex (5 were male).

A-ii-b. Siblings

The total number of siblings and half-siblings born alive in the 166 families was 321, with a male:female ratio of 1.15 (172/149), which was not significantly different from the ratio of 1.05 in the general population.

Sixteen siblings (5 male) were not HLA-typed during 1978-80 because they were geographically inaccessible, died from a cause unrelated to diabetes (1 sibling), or witheld consent to venepuncture. Their ages (mean ± SD) were, in males and females respectively, 27.4 ± 4.5 and 23.4 ± 5.2 years, significantly greater than the ages of the 95% of sibs who were actually tested ($p < 0.001$).

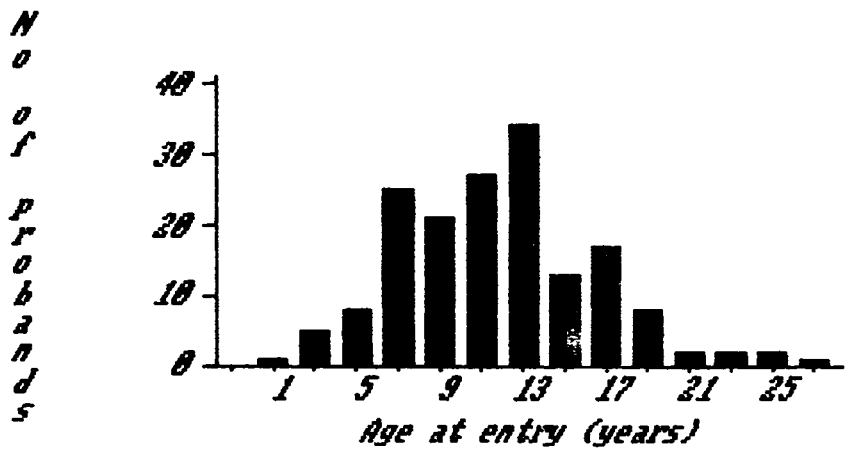


FIGURE VII-1. AGE OF 166 BWFS PROBANDS AT ENTRY TO STUDY. Each bar represents the number of probands within a 2-year tranche of age centred at the point marked on the horizontal axis.

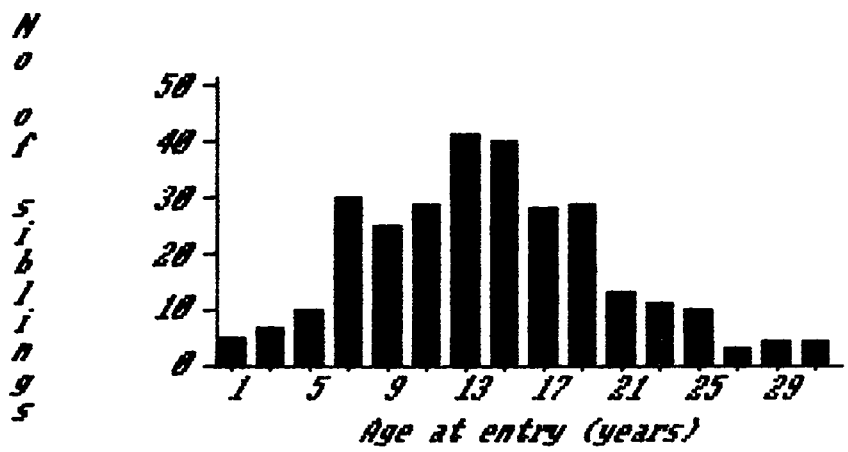


FIGURE VII-2. AGE OF 289 BWFS SIBLINGS AT ENTRY TO STUDY. Each bar represents the number of siblings within a 2-year tranche of age centred at the point marked on the horizontal axis.

Ten siblings, in addition to the 166 probands, were already diabetic and treated with insulin at entry to the study. Two of these were male and 8 female (not significant: $p = 0.06$). One lives in Canada and has not been HLA-typed, but is an identical twin of her proband.

Details of age and sex for the 289 HLA-genotyped full siblings of probands in 160 families are summarized in Figure VII-2.

Among the 280 initially non-diabetic siblings entering the study, 264 were followed up (with at least one further blood test) and 241 continued as full participants for at least 6 months. Most of the remainder remained in intermittent contact.

A-ii-c. Parents

Figure VII-3 the age distributions of the available parents, 154 fathers and 159 mothers, in the 160 families without expaternity of the proband. Seven fathers and two mothers had Type 1 diabetes at entry to the study.

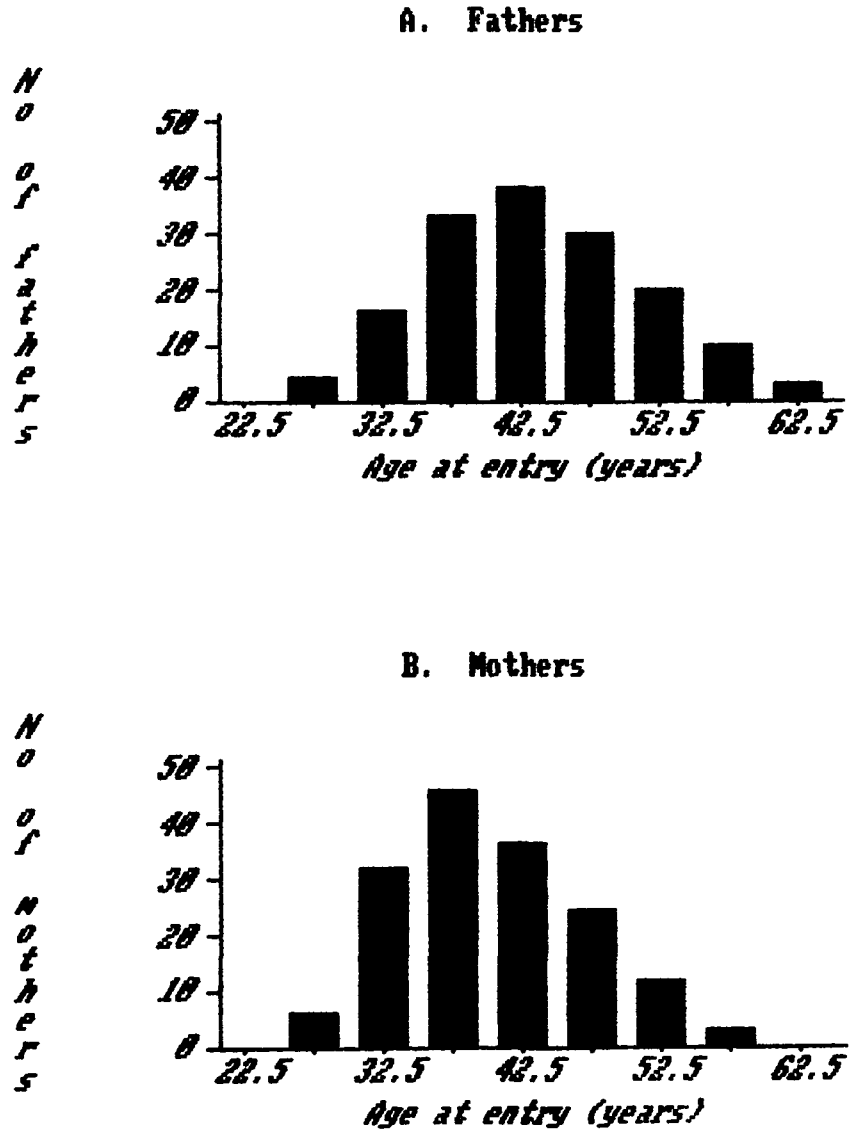


FIGURE VII-3. AGE OF 154 FATHERS (A) AND 159 MOTHERS (B) AT ENTRY TO BWFS. Each bar represents the number within a 5-year tranche of age centred at the point marked on the horizontal axis.

A-iii. Details of diabetes in probands

Appendix 1 gives relevant data.

A-iii-a. Age at diagnosis

The median (range) age at diagnosis of diabetes in the probands is 8.5 (0.7-18.7) years; the distribution (Figure VII-4) shows the expected maxima around ages 4 and 10-12 years, but is biased because of preferential ascertainment of younger families (specifying a sibling aged under 20 years at entry). Age at diagnosis is considered again in the context of heterogeneity (VIII-2-B-iii-c).

A-iii-b. Duration of symptoms before diagnosis

Retrospective information, based on hospital clinical records and recollection by parents, was available for 151 probands (84 male) and is shown in Figure VII-5. Median reported duration of symptoms in boys was 3.5 weeks, significantly less than the 6.0 weeks in girls ($p = 0.032$, Mann-Whitney 'U' test). The apparent bimodality of the distribution may be artefactual: long-standing and insidious symptoms would tend to be reported as present for 'about a year' and coded as 50 weeks.

A-iii-c. Season at reported onset of symptoms

The distributions of calendar month and of season (i.e. three-month period, or quarter) at reported onset of symptoms show the seasonal pattern described in other populations (III-4-B-i). The distribution overall is not significantly different from random, but that in probands aged >5 years at diagnosis is significant ($p < 0.025$). Seasonality is considered further in the context of heterogeneity (VIII-2-B-vi).

A-iii-d. Duration of diagnosed diabetes at entry to study

Median duration of diabetes at the time of recruitment was 4.3 (range 0 - 24.3) years, with the distribution shown in Figure VII-6. Eleven probands were first tested within two weeks of diagnosis.

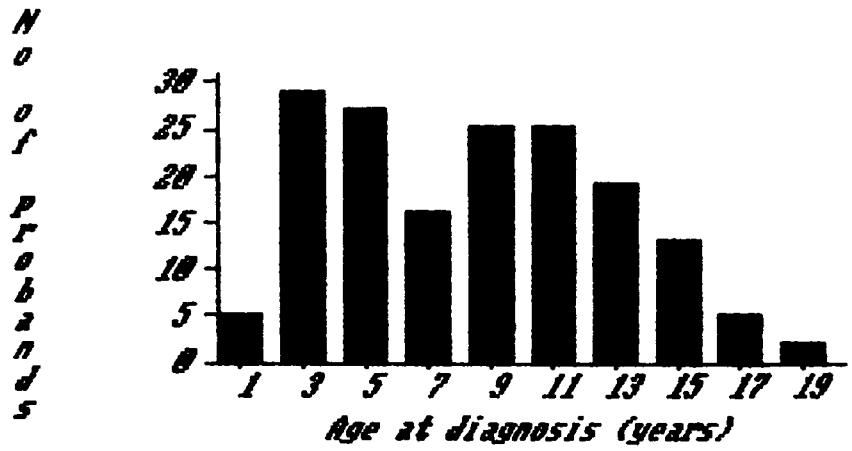


FIGURE VII-4. AGE AT DIAGNOSIS OF DIABETES IN 166 BWFS PROBANDS. Each bar represents the number of probands within a 2-year tranche of age, centred at the point marked on the horizontal axis.

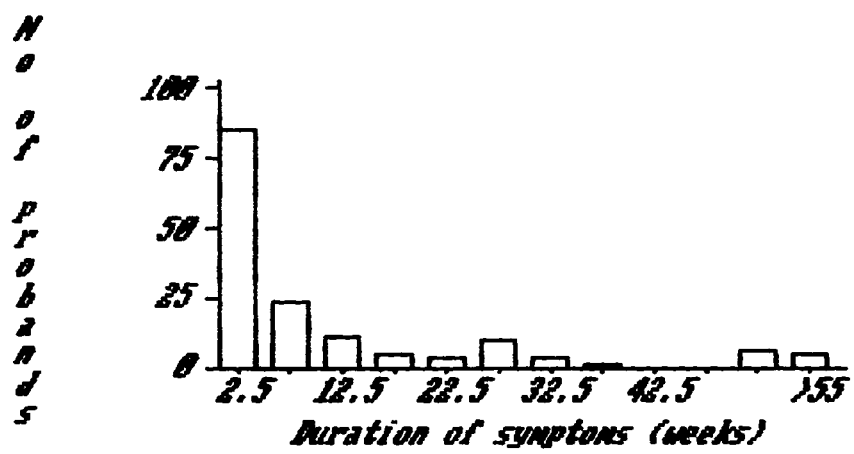


FIGURE VII-5. DURATION OF DIABETIC SYMPTOMS REPORTED BY 151 BWFS PROBANDS. Each bar represents the number of probands within a 5-week tranche of symptom-duration, centred at the point marked on the horizontal axis.

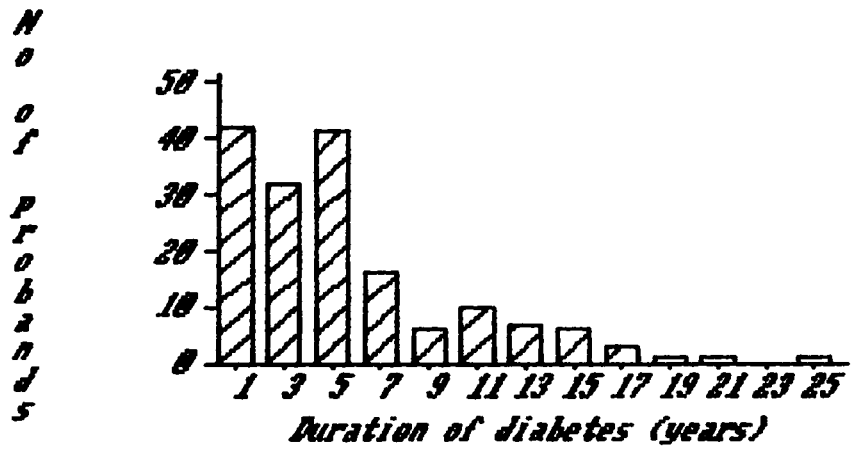


FIGURE VII-6. DURATION OF DIABETES IN 166 PROBANDS AT ENTRY TO BWFS. Each bar represents the number of probands within a 2-year tranche of age, centred at the point marked on the horizontal axis.

B. NEWLY-DIAGNOSED TYPE 1 DIABETIC SUBJECTS

Included here are white Caucasoid subjects developing Type 1 diabetes by age 20 years, notified to the author within 1 week of diagnosis in the period August 1978 - July 1980 inclusive, and living in the same geographical area as families in the Family Study.

Relevant data are given in Appendix 3.

B-i. Numbers and sex

A total of 45 new cases (28 male) were ascertained during the two-year period, including 4 among families aleady participating in the BWFS. Four further cases were identified but were not ascertained because they were too young (girls aged 2 and 3), consent was witheld (boy aged 6) or notification was too late (boy aged 4). The incidence rate in the Windsor, Maidenhead and Slough districts is calculated in RESULTS, VIII-4-B-i-b (see also Appendix 6-1).

Of the 45 new cases ascertained, 11 entered the Family Study as probands during the period concerned here, while four were siblings already taking part. The following details concern the remaining 30.

B-ii. Age, season and duration of symptoms at diagnosis

The mean age of these cases was 12.35 ± SD 3.5 years, significantly greater than age at diagnosis of diabetes in probands in the family study ($p<0.0001$). Duration of symptoms and seasonality were however comparable to those in the family study. Details are given in Appendix 1, and are included in analyses for heterogeneity (VIII-2-B-ii).

C. OTHER SERIES PROVIDING SUPPLEMENTARY DATA

For heterogeneity analyses, the numbers of patients ascertained by the author in the Barts-Windsor Family Study and as newly-diagnosed cases were insufficient, requiring supplementation from the sources outlined below. In total 589 informative, unrelated Type 1 diabetic subjects are available.

C-i. UK (British Diabetic Association) Multiplex Study

The Paediatric Register of the British Diabetic Association (BDA) consists of cases of diabetes newly-diagnosed at ages <16 years from 1972 onwards throughout the United Kingdom (Gamble 1980a), and includes information on affected siblings. Most such multiplex families were visited in 1979 by Dr. Angela Walker, the information and laboratory results obtained being so far as possible the same as in the Barts-Windsor Family Study. Only insulin-treated multiplex sibships were included. Published observations (Walker & Cudworth, 1980) included some multiplex families from the Barts-Windsor Study and the Liverpool series (see below) but excluded some with ambiguous genotypes, so the sample of 117 white Caucasoid families from the UK study recorded here differs slightly from that published. The first-diagnosed child in each sibship is again taken as the proband, except that in four cases where some information is missing the second-diagnosed is used for heterogeneity analyses (VIII-4).

In sex (M:F = 1.29), age at diagnosis (7.9 ± 4.5 years), age at time of study (15.5 ±5.7 years) and season at diagnosis the probands in the UK Multiplex study were comparable to those in the Barts-Windsor Family Study. The second-diagnosed members of the UK multiplex sibships had a mean age at diagnosis of 10.1 ± 3.7 years, and a similar M:F ratio (1.40).

In the heterogeneity analyses a further 8 families (see below, C-iii) are included. The age distribution of all 125 probands is depicted in Figure VII-7.

C-ii. Liverpool Study

Existing and new cases of Type 1 diabetes diagnosed at ages < 30 years, mainly in the Liverpool area, were ascertained and HLA-typed by the late Professor Cudworth in 1974-6 (Cudworth & Woodrow, 1976).

At his specific request the following details, included in his PhD thesis (Cudworth, 1977), are used here in heterogeneity analyses (VIII-2-

B-ii): age at onset (13.8 ± 7.0 years; Figure VII-8), sex (M:F = 156:132, ratio 1.19), HLA type and number of diabetic members in sibship. In 14 cases the latter is not known; 46 of the remaining 274 sibships are multiplex.

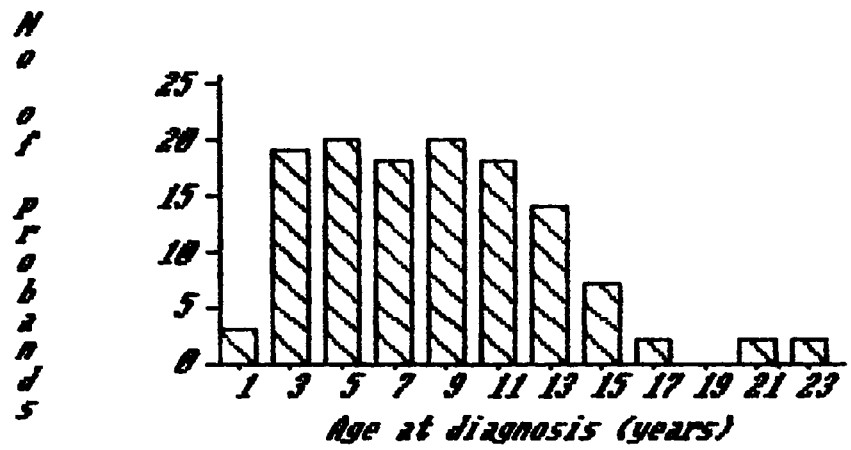


FIGURE VII-7. AGE AT DIAGNOSIS OF DIABETES IN 125 PROBANDS IN THE UK (BDA) MULTIPLEX STUDY. Each bar represents the number of probands in a two-year tranche of age, centred at the point marked on the horizontal scale.

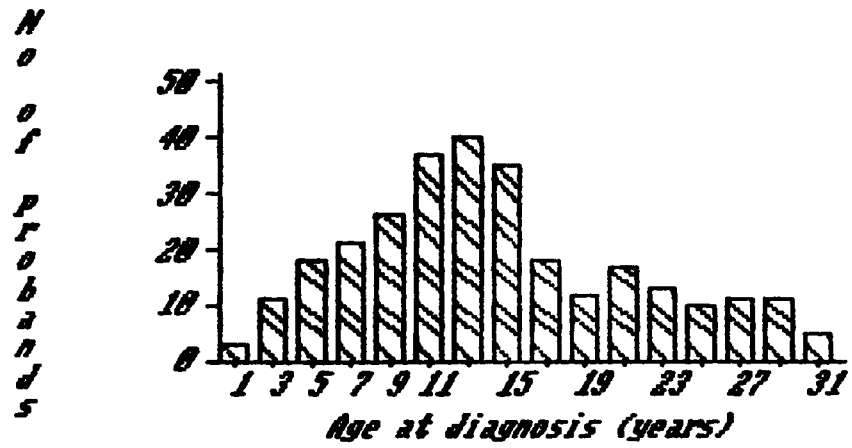


FIGURE VII-8. AGE AT DIAGNOSIS OF DIABETES IN 288 PROBANDS IN THE LIVERPOOL STUDY. Each bar represents the number of probands in a two-year tranche of age, centred at the point marked on the horizontal scale.

C-iii. Miscellaneous

Three HLA-genotyped multiplex sibships were ascertained by the author but did not qualify for the BWFS because of age, and a further five were ascertained by Professor Cudworth. Ages at diagnosis of Type 1 diabetes in the 8 probands (4 male) were 13, 7, 8, 7, 15, 12 and 20 years. These are grouped with the BDA Multiplex Study cases for heterogeneity studies, ascertainment criteria being similar.

D. CONTROL SUBJECTS

HLA phenotype frequencies in diabetic probands are compared with those in a panel of 170 healthy, white Caucasoid subjects typed at St. Bartholomew's Hospital. Most were volunteers (usually pregnant women and their husbands) being screened for HLA typing sera. Some were unrelated adult household members or friends of Family Study participants, others were laboratory staff, their colleagues or acquaintances. The majority came from London rather than the nearby area around Windsor.

Testing of control families (such as neighbours of Family Study participants) was considered, but was rejected for logistic and ethical reasons. Thus there is no ideal control sample of fully-genotyped families. So far as possible, internal controls are used here; for example, various attributes of HLA haplotypes appearing in diabetic probands are compared with those of haplotypes present in the same families but not appearing in any diabetic member. Similarly, siblings are grouped according to the degree of haplotype concordance with probands for comparison of immunological and other variables.

PART VIII. RESULTS

CHAPTER 1

BASIC GENETIC AND IMMUNOLOGICAL OBSERVATIONS

CONTENTS

A Introduction · 204
B Chromosome-6 genetic markers in the BWFS probands · · · · · · · · 204
  i Phenotype and gene frequencies in BWFS probands · · · · · · 204
    a HLA-A,B,Cw · 204
    b HLA-D · 208
    c Properdin Factor B (Bf) · · · · · · · · · · · · · · · · · · 208
    d Complement C2 · 209
    e Glyoxalase-I (GLO) · 210
  ii Linkage disequilibrium in the BWFS · · · · · · · · · · · · · 211
    a First-order (two-locus) effects · · · · · · · · · · · · · · 211
    b Second-order linkage disequilibrium, B18:Cw5:BfF1 · · · · 211
  iii Genetic recombination in the BWFS · · · · · · · · · · · · · 213
    a HLA-A-B recombination · · · · · · · · · · · · · · · · · · · 213
    b HLA-B - GLO recombination · · · · · · · · · · · · · · · · · 213
    c HLA-B-DR recombination · · · · · · · · · · · · · · · · · · 213
    d Discussion of observations on recombination · · · · · · · 215
  iv Zygotic assortment of HLA-B8: random or not? · · · · · · · · 215
    a Background; the question of bias · · · · · · · · · · · · · 215
    b Approach · 218
    c Observations · 218
    d Analysis · 220
    e Discussion · 224
  v Siblings grouped by haplotype concordance with probands · 225
    a Distribution of siblings between groups · · · · · · · · · 225
    b Distribution of diabetic siblings · · · · · · · · · · · · 226
    c Age distribution across haplotype concordance groups 226
    d Discussion · 226
C Basic immunological observations · · · · · · · · · · · · · · · · 228
  i Autoimmune and other diseases in diabetic probands · · · · 228
  ii Family history of relevant diseases · · · · · · · · · · · · 228
  iii Auto-antibodies in Type 1 diabetic subjects · · · · · · · · 230
  iv Auto-antibodies in non-diabetic subjects · · · · · · · · · 232
  v Correlation between ICA results by the two techniques · · 234
  vi Discussion of basic autoimmunity results · · · · · · · · · 235

1-A. INTRODUCTION

This RESULTS chapter confirms that the probands in the Brts-Windsor Family Study (BWFS) show genetic and immunological features already established as characteristic of Type 1 diabetes. In addition it explores further possible HLA associations, and examines some questions raised in the Literature Survey and summarized under IV-INDICATIONS-A.

Because of the large number and variety of different observations, their implications are discussed within each section as appropriate rather than in a single, separate Discussion section.

1-B. CHROMOSOME-6 GENETIC MARKERS IN THE BWFS

For reasons of space and confidentiality, (see VI-3-B-viii), full results on individual probands and families are not given here, but may be examined at the discretion of the Steering Committee of the Barts-Oxford Family Study (See Appendix 4).

1-B-i. Phenotype and gene frequencies in BWFS probands

Appendix 2 summarizes phenotype and gene-count observations for all specificities tested in the 165 unrelated probands of the BWFS, as well as those in Barts controls and in 'control haplotypes' from the families (VII-D). 'Probable-homozygous' blanks are counted as homozygous for gene frequencies at HLA-A and B loci, an expedient which is justified by the absolute lack of 'true blanks' confirmed in family members with different genotypes. This does not apply to the Cw locus (see below).

Results of particular interest are examined and discussed below.

1-B-i-a. HLA-A,B,Cw

Table VIII-1 shows results for selected specificities. The established associations of Type 1 diabetes with HLA-A1, B7 (negative) B8 and Bw62 are confirmed. Associations reported elsewhere but not significant here include A2, A3, A11 and B18.

TABLE VIII-1. SOME HLA-A,B,Cw PHENOTYPE AND GENE FREQUENCIES IN 165 UNRELATED PROBANDS OF THE BWFS, with control observations. For full results on all specificities see Appendix 2.

HLA	PHENOTYPE FREQUENCY						GENE FREQUENCY					
	Prob (%)	Cont (%)	2p uncorr	Rel.risk mean	95%cl	PAR	Prob'd	Cont'l	2p uncorr	'Rel.risk' mean	95%cl	'PAR'
A 1	48	34	0.008	1.8	1.2-2.9	0.23	0.2697	0.1564	<0.001	2.0	1.3-2.9	0.13
2	56	45	NS	-	-	-	0.3091	0.2736	NS	-	-	-
3	21	22	NS	-	-	-	0.1061	0.1433	NS	-	-	-
11	7	10	NS	-	-	-	0.0364	0.0619	NS	-	-	-
24	19	14	(0.07)	-	-	-	0.0970	0.0847	NS	-	-	-
B 7	13	27	0.002	0.4	0.2-0.7	-0.19	0.0638	0.1629	10^{-4}	0.3	0.2-0.6	-0.12
8	54	27	10^{-6}	3.1	2.0-4.9	0.37	0.2909	0.1205	$<10^{-6}$	3.0	2.0-4.5	0.19
18	9	9	NS	-	-	-	0.0485	0.0391	NS	-	-	-
39	11	4	0.029	2.7	1.1-6.6	0.07	0.0547	0.0163	0.019	3.2	1.3-8.3	0.02
44	20	29	(0.06)	-	-	-	0.1061	0.1824	0.008	0.5	0.3-0.8	-0.01
57	2	8	(0.05)	-	-	-	0.0121	0.0391	0.045	0.3	0.1-0.9	-0.03
60	16	14	NS	-	-	-	0.0818	0.0358	0.028	2.3	1.1-4.6	0.05
62	28	8	10^{-6}	4.7	2.4-9.0	0.23	0.1545	0.0423	$<10^{-5}$	3.9	2.1-7.2	0.12
Cw3	45	24	10^{-4}	2.5	1.6-4.0	0.27	0.2445	0.0882	10^{-6}	3.3	2.1-5.3	0.17
7	66	55	(0.086)	-	-	0.28	0.4159	0.2902	0.007	1.7	1.2-2.6	0.18
* *-*	(not applicable for phenotypes)						0.0651	0.2080	10^{-5}	0.3	0.1-0.5	-0.18

Note that controls for phenotype frequencies do not come from the same geographical area as the probands, whereas controls for gene frequencies are from the same families as the probands (see text). The two control samples therefore differ slightly in genetic constitution.

Gene frequencies incorporate corrections for probable-homozygous blanks and for incomplete typing for Cw6,7,8 (see text). Cw phenotype frequencies only are derived indirectly by calculation from corrected gene frequencies using the formula for Hardy-Weinberg equilibrium (III-6-B-ii).

The association with HLA-Bw60 reported in some populations is not significant here when phenotype frequencies in the Windsor area family study probands are compared with those in the London controls, nor when the thirty newly-diagnosed cases, of whom six (20%) are positive for Bw60, are included (Appendix 3). However, comparison of gene frequencies in probands with those in 'control' haplotypes from the Windsor area families does support such an association.

The association with Bw39 was at first discounted as a Type II error due to multiple comparisons. At the B locus, 13 specificities gave counts large enough (>7) to merit comparisons, and correction of the p-values by the factor 13 is appropriate for previously-unreported associations. However, three (10%) of the newly-diagnosed patients are Bw39-positive, giving an overall p-value of 0.016 for phenotype frequency among family probands + newly-diagnosed cases. Furthermore, a similar association at a much higher level of significance has recently been reported in Finland for the parent broad antigen B16, which is more common there than in UK (Mustonen et al, 1985). The deviation from random distribution observed here is therefore probably genuine after all.

HLA-Aw24 is in linkage disequilibrium with B39, and the parent antigen A9 is also increased in Finnish diabetic patients, but Aw24 is not significantly associated with diabetes in the present study.

The weak negative association of Type 1 diabetes with HLA-Bw44 is probably secondary to the positive association with B8, B18, Bw39, Bw60 and Bw62: since one or other of these occurs on 58% of 'diabetic' haplotypes, only 42% of such haplotypes are free for randomly-distributed HLA-B alleles, compared with 79% of 'control' haplotypes. Inclusion of the independent sample of newly-diagnosed cases (Appendix 3) strengthens rather than weakens the Bw44 association (combined data: phenotype frequency 20%, 2p = 0.049, RR = 0.6 (0.3-1.0), PAR = -0.133). The same applies to Bw57 (combined data: phenotype frequency 2%, 2p = 0.021, RR = 0.3 (0.1-0.8)

Analysis for association involving the Cw locus is more problematic because of the substantial proportion of blanks. When the families were first typed in 1978-9, HLA-Cw6,7 and 8 had not been defined, although some sera did identify Cw7 (then locally termed 'blue'). Initially, therefore, blanks accounted for nearly half of the gene-counts and occurred in most probands. Later some 75% of the families with blanks were re-typed for Cw6, 7 and 8. Thus the observed phenotype and gene counts (Appendix 2-3-a) include two classes of serological blank, "-"

where typed with sera covering Cw1-Cw8, and "=" where typed only for Cw1-5, as well as the 'probable-homozygote' blank 'x'.

Correction of gene frequencies for 'x' was done first. Among 'control' haplotypes, 20 included Cw'x', paired as shown below. To avoid bias it was assumed that the proportion representing true serological blanks (Cw'-') was the same as that overall, i.e. with a frequency of approximately 0.2. Only one 'x' was paired with '=' and thus not typed for Cw6-8. Two were paired with Cw5, and the probability of each being '=' was assumed equal to the overall frequency of '=', about 0.2. Therefore a 'count' of 1 + 0.4 = 1.4 was added to the observed count of '='. Seventeen occurrences of 'x' among control haplotypes were paired with Cw6-8. These were counted as '-' or as the paired antigen in the proportion 0.28:0.72, since the frequency of Cw'-' among the 210 haplotypes typed for these antigens is approximately 0.28. Among probands' haplotypes 5 included Cw'x'; a similar method of correction was used.

The overall gene-counts were incremented to correct for 'x' as shown here:

HLA-Cw specificity	5	6	7	8	'-'	'='
Control haplotypes						
Pairings with 'x'	2	2	14	1	0	1
Final increment	1.6	1.4	10.1	0.7	4.8	1.4
Probands' haplotypes						
Pairings with 'x'	1	0	4	0	0	0
Final increment	0.8	0	3.7	0	0.3	0.2

In order to calculate gene frequency for Cw"-", 6, 7 and 8, it was necessary to exclude those typing "=", and to avoid bias by reducing proportionately the total number of haplotypes typed to give a nominal figure (N') for use both as the denominator of the fraction representing genotype frequency and in statistical testing. The formula used is given in Appendix 2-3-b. Phenotype frequency could not be directly corrected in this way because of homozygosity; it was most simply derived by calculation from gene frequency assuming Hardy-Weinberg equilibrium, while the nominal total number of subjects tested for Cw6-8 was taken to be half the nominal number of haplotypes so tested. This procedure

enabled reasonably valid comparison with phenotype data from the London control subjects.

The only really significant HLA-Cw result in Table VIII-1 is confirmation of an association between Type 1 diabetes and Cw3. In view of the assumptions involved in correction for incomplete typing for Cw6-8, the apparent association with Cw7 when gene frequencies in probands and in control haplotypes are compared, and even the apparently highly-significant excess of blanks among probands' haplotypes, cannot be considered established from these data but require further investigation.

1-B-i-b. HLA-D

Fifty-eight probands from the Family Study and 43 healthy control subjects were selected for presence of one of the following HLA-B specificities: B7, B8 and Bw62. These and their families were typed for HLA-D to investigate mode of inheritance of susceptibility. Results have been published (Sachs et al, 1980) but are not detailed here, except for the finding that 6 probands were positive for Dw2, all six being among the 26 selected for presence of B7.

1-B-i-c. Properdin Factor B (Bf)

Fifty-three of the 165 unrelated families, or parts of families, were Bf-typed. The principal aim was to determine whether BfF1 showed an association with Type 1 diabetes independent of HLA-B. To increase the chance of including all occurrences of this uncommon allotype, all families having a haplotype positive for HLA-B18 were included, taking advantage of the linkage disequilibrium between Bf-F1 and HLA-B18. Thus B18 occurs in 28 (4.6%) of the 607 probands' and 'control' haplotypes overall, and in 27 (13.8%) of the 196 haplotypes typed for Bf ($2p = 0.0004$). Because of this bias, comparison of phenotype frequencies with randomly-ascertained controls is inappropriate. However, comparison between gene frequencies in proband and 'control' haplotypes from the same families is valid if the bias is taken account of.

Table VIII-2 shows the gene frequencies observed. The association between Bf-F1 and Type 1 diabetes is confirmed. It is not attributable to

the bias towards B18 since, in this sample, B18 is not associated with Type 1 diabetes (16/106 diabetic and 11/90 control haplotypes, typed for Bf, were positive for B18; X^2 = 0.014, 2p > 0.5)

Bf-S does not appear to be significantly associated with diabetes in a simple 2x2 comparison of gene frequencies in proband and 'control' haplotypes. However, when B8 is included in a 2x2x2 comparison a weak negative association appears (p <0.025) and the simplification by collapsing the table over B8 is shown not to be valid (see VIII-1-B-ii-a and VIII-2-B-iii-e).

Linkage disequilibrium between Bf-F1, Cw5 and HLA-B18, and associations with diabetes, are explored below (1-B-ii).

TABLE VIII-2. COMPLEMENT FACTOR B (Bf) GENE-COUNTS IN UNRELATED PROBANDS AND IN 'CONTROL' HAPLOTPES FROM THE BWFS

Bf	PROBANDS (n=53)		CONTROL HAPLOTYPES		ASSOCIATION		
	Gene count	Gene frequency	Gene count	Gene frequency	2p	'RR'	'PAR'
Bf-F	12	0.1132	9	0.1000	NS		
Bf-F1[a]	13	0.1226	3	0.0333	0.039	3.6	0.09
Bf-S	81	0.7642	78	0.8667	NS		
Bf-S1	0	0.0000	0	0.0000			
Total	106	1.0000	90	1.0000			

[a] Frequency of Bf-F1 in both probands and control haplotypes is positively biased by preferential selection of HLA-B18-positive families for typing (see text).

1-B-i-d. Complement C2

A random sample of 106 diabetic probands was typed for C2. Unfortunately no local control results are available for direct comprison of phenotype frequencies, but published figures from West Germany (Dewald & Rittner, 1979) are included in Table VIII-3 for comparison. Geographical differences between that population and UK are most unlikely, since gene frequencies in the German study and in 800 healthy Newfoundlanders (Dornan et al, 1981) were identical (respectively, C2-A 0.000 and <0.001; C2-B 0.035 and 0.03; C2-C 0.965 and 0.97).

In addition full genotypes were determined in selected families, principally those where there were GLOxHLA-B, or HLA-AxB, recombinants. The number of 'control' haplotypes is therefore small, but overall it is quite clear from Table VIII-3 that there is no important association between Type 1 diabetes and C2 allotypes.

TABLE VIII-3. COMPLEMENT C2 PHENOTYPE AND GENE FREQUENCIES IN TYPE 1 DIABETIC PROBANDS OF THE BWFS, WITH CONTROL OBSERVATIONS.

ALLOTYPE FREQUENCIES				GENE FREQUENCIES		
C2 type	Probands n=106 No. %	Controls[a] n=289 No. %	Diff. Fisher 2p	Proband Count Freq.	Control[b] Count Freq.	Diff. Fisher 2p
A	0 0	0 0		0 0.0000	0 0.0000	NS
B	10 9	20 7	NS	10 0.0472	1 0.0370	NS
C	106 100	289 100		202 0.9528	26 0.9630	NS
Total	106	289		212 1.0000	27 1.0000	

NOTES: [a] Control phenotype frequencies from West Germany (see text)
[b] Control gene frequencies from haplotypes within the BWFS but not occurring in any diabetic subject.

1-B-i-e. Glyoxalase-I (GLO)

GLO allotypes were determined in 155 families in the laboratory of Dr. SG Welch. In this case 'control haplotypes' from the same families are of questionable value because of the frequency of recombination between GLO and HLA. The control figures shown in Table VIII-4 therefore represent healthy subjects tested in the same laboratory (Parr et al, 1977).

TABLE VIII-4. GLYOXALASE (GLO) ALLOTYPES IN UNRELATED TYPE 1 DIABETIC BWFS PROBANDS AND IN LONDON CONTROL SUBJECTSa

GLO	ALLOTYPE FREQUENCY			GENE COUNTS		
	Probands (n=155) No. %	Controls[a] (n=296) No. %	Diff. 2p	Probands No. Freq.	Controls No. Freq.	Diff. 2p
1	102 66	203 69	NS	138 0.4451	263 0.4443	NS
2	116 75	236 80	NS	172 0.5548	329 0.5557	NS
Total				310 0.9999	592 1.0000	

NOTES: [a] Control data from the same laboratory (Parr et al (1977)

1-B-ii. Linkage disequilibrium in the BWFS

Linkage disequilibria between selected specificities in the HLA region were examined in 637 haplotypes: 330 present in probands and 307 not occurring in any diabetic family member. The principal aim was examination of possible epistatic effects in genetic susceptibility to Type 1 diabetes by use of log-linear modelling (see VIII-2-B-iii-c). Therefore only certain specificities associated with Type 1 diabetes in this or other studies, and occurring sufficiently frequently for multivariate analysis, are considered here. HLA-Cw7 is not included because of incomplete typing (see above).

Again, to avoid bias, 'possible homozygote blanks' are counted as if true homozygosity were present. Being based on gene counts these calculations do not assume Hardy-Weinberg equilibrium. Because of missing data the total number of haplotypes studied is less than 637 for some pairs of specificities.

1-B-ii-a. First-order (two-locus) effects

Table VIII-5 gives details of some linkage disequilibria studied. In most of these cases it was shown that the 'diabetic' and 'control' haplotypes could be pooled for this purpose (see VIII-2-B-iii-c). However this was not so for the association HLA-B8 vs Bf-S, which is therefore examined separately in the two samples.

1-B-ii-b. Second-order linkage disequilibrium, B18:Cw5:BfF1

Table VIII-6 shows the 2x2x2 contingency table for pooled 'diabetic' and 'control' haplotypes classified according to presence or absence of each of the allelotypes HLA-B18, HLA-Cw5 and Bf-F1. Inspection shows that BfF1 occurs predominantly with both B18 and Cw5. This three-way association was shown to be significant by use of the GLIM package. Omission of the second-order interaction term from the log-linear model gave an unacceptable fit (X^2_L= 8.901, 1 degree of freedom; $p < 0.01$). The three-locus Δ value (equation 6-7 under III-6-B-iii) is 0.0469. (See also VIII-2-B-iii and Table VIII-41.)

TABLE VIII-5. CERTAIN FIRST-ORDER LINKAGE DISEQUILIBRIA AMONG HAPLOTYPES IN THE BWFS. Except where stated, haplotypes present in probands are pooled with those present only in non-diabetic subjects (see text). Each haplotype is counted only once per family.

Alleles	Observed numbers of haplotypes				Significance (Fisher)	/\	/\,
First	+		-				
Second	+	-	+	-			
HLA-A1 -B8	96	41	37	463	<<0.0001	0.1058	0.6456
HLA-A3 -B7	23	56	48	510	<0.0001	0.0242	0.2889
HLA-Aw24-Bw39	11	47	11	568	<0.0001	0.0141	0.4499
HLA-Bw62-Cw3	60	4	48	522	<<0.0001	0.1223	0.7392
HLA-Bw60-Cw2	1	37	25	571	NS		
HLA-B18 -Cw5[a]	16	11	16	153	<0.0001	0.0591	0.5131
HLA-B18 BfF1[a]	14	13	2	167	<<0.0001	0.0602	0.8550
HLA-Cw5 BfF1[a]	14	18	2	162	<<0.0001	0.0581	0.8506
HLA-B8 BfS[b] probands	26	0	55	25	0.0006	0.0578	0.3125
'control'	9	0	69	12	NS	(0.0133)	(0.1481)

NOTES [a] These form part of the three-locus linkage disequilibrium between B18, Cw5 and BfF1; the mean values given here for the two-locus associations are not necessarily representative since they are dependent on the allele present at the third locus (see below), and subjects typed for Bf were selected (see text).

[b] The contingency table for B8 vs BfS vs Diabetes could not be 'collapsed' over diabetes (see text).

TABLE VIII-6. THE THREE-LOCUS LINKAGE DISEQUILIBRIUM BETWEEN HLA-B18, Cw5 AND BF-F1: OBSERVATIONS. Contingency table giving the number of each of the 8 possible haplotypes observed in informative families. Haplotypes present in probands are pooled with those present only in non-diabetic subjects (see text). Each haplotype is counted only once per family.

Allele	Haplotype constituents								Total
HLA-B18	+	+	+	+	-	-	-	-	
HLA-Cw5	+	+	-	-	+	+	-	-	
BfF1	+	-	+	-	+	-	+	-	
No. observed	14	2	0	11	0	16	2	151	196

1-B-iii. Genetic recombination in the BWFS

1-B-iii-a. HLA-A-B recombination

For crossovers between HLA-A and -B there were, in the families studied here, 830 informative meioses and only one recombinant, giving a recombination frequency estimate of 0.001 (95% confidence limits approximately 0.000-0.002, assuming binomial sampling). In this case HLA-Cw remained with HLA-B (Figure VIII-1), and no B-Cw crossovers were identified at all.

1-B-iii-b. HLA-B - GLO recombination

For crossovers between HLA-B and GLO there were 414 informative meioses and 34 recombinants, representing a recombination rate of 0.082 (95% confidence limits 0.055-0.110). There was no significant association between recombination rate and either sex or diabetic status.

1-B-iii-c. HLA-B-DR recombination

Many of the families with HLA-B - GLO recombinants were selected for HLA-DR typing during the period 1979-80 in an attempt to identify B-DR recombinants. Other families DR-typed during this period were selected because of homozygosity for HLA-A,B and Cw in one parent, or because two or more siblings were diabetic. Two B-DR crossovers were found among these selected families, and a third during subsequent routine DR-typing (Figure VIII-1). By late 1983 534 meioses in the 166 families were informative for B-DR, the three recombinants giving a recombination rate of 0.0056 (95% confidence limits approximately 0.000-0.012). Cases where DR types were deduced rather than determined directly are of course excluded from this analysis.

Child 4 in Family 42 is particularly interesting for two reasons. First, there is a double crossover (B-DR and DR-GLO); the obvious alternative explanation of mis-typing for DR is most unlikely, because this family has been typed twice in London while cells from Child 42-4 have also been typed in the laboratory of van Rood in Leiden, The Netherlands as part of a collaborative study. Secondly, this child and

FAMILY 5 **Proband is 5** **Twins 2 pairs**

		FATHER								MOTHER						
Locus		A	B	Cw	DR	Bf	C2	GLO		A	B	Cw	DR	Bf	C2	GLO
Haplotp.	a	26	49	7	5			1	c	1	18	5	3			2
Haplotp.	b	30	57	6	4			2	d	2	44	5	4			2
		Diabetes 0								Diabetes 0						

dztwins dztwins

IDS	3	4	5	6	7	8
Sex	f	f	m	m	f	f
Pat.Hap.	a	a	b	a	b	a
Mat.Hap.	d	c	6	d	c	c
Diab.	0	0	1	0	0	0

FAMILY MEMBER 5 SEX m

GENOTYPE	Haplotype b							Haplotype 6						
Locus	A	B	Cw	DR	Bf	C2	GLO	A	B	Cw	DR	Bf	C2	GLO
	30	57	6	4			2	2	44	5*	3			2

FAMILY 39 **Proband is 3**

		FATHER								MOTHER						
Locus		A	B	Cw	DR	Bf	C2	GLO		A	B	Cw	DR	Bf	C2	GLO
Haplotp.	a	3	27	2	2	S	C	2	c	2	8	7	3	S	C	1
Haplotp.	b	2	14	8	2	S	C	2	d	2	62	3	4	S	C	2
		Diabetes 0								Diabetes 0						

IDS	3	4	5	6
Sex	m	m	m	m
Pat.Hap.	a	b	a	a
Mat.Hap.	d	d	5	d
Diab.	1	0	0	0

FAMILY MEMBER 5 SEX m

GENOTYPE	Haplotype a							Haplotype 5						
Locus	A	B	Cw	DR	Bf	C2	GLO	A	B	Cw	DR	Bf	C2	GLO
	3	27	2	2	S	C	2	2	8	7*	4	S	C	2

FAMILY 42 **Proband is 3** **Twins 2 pairs**

		FATHER								MOTHER						
Locus		A	B	Cw	DR	Bf	C2	GLO		A	B	Cw	DR	Bf	C2	GLO
Haplotp.	a	1	62	3	1	F	B	1	c	1	8	7	3	S	C	2
Haplotp.	b	3	62	3	9	S	C	2	d	2	39	-	4	S	C	1
		Diabetes 0								Diabetes 0						

IDS	3	4	5	6
Sex	f	f	f	f
Pat.Hap.	a	b	a	a
Mat.Hap.	c	6	c	d
Diab.	1	1	0	1

FAMILY MEMBER 4 SEX f

GENOTYPE	Haplotype b							Haplotype 6						
Locus	A	B	Cw	DR	Bf	C2	GLO	A	B	Cw	DR	Bf	C2	GLO
	3	62	3	9	S	C	2	2	39	-*	3	S	C	*1

FAMILY 97 **Proband is 5**

		FATHER								MOTHER						
Locus		A	B	Cw	DR	Bf	C2	GLO		A	B	Cw	DR	Bf	C2	GLO
Haplotp.	a	25	60	3	7	S	C	1	c	1	8	7	3	S	C	2
Haplotp.	b	24	27	1	4	S	C	2	d	26	44	1	4	S	C	2
		Diabetes 0								Diabetes 0						

IDS	3	4	5	6
Sex	f	m	m	f
Pat.Hap.	a	4	b	a
Mat.Hap.	d	d	c	c
Diab.	0	0	1	0

FAMILY MEMBER 4 SEX m

GENOTYPE	Haplotype 4							Haplotype d						
Locus	A	B	Cw	DR	Bf	C2	GLO	A	B	Cw	DR	Bf	C2	GLO
	24*	60	3	7	S	C	1	26	44	1	4	S	C	2

FIGURE VIII-1. INTRA-HLA RECOMBINANTS IN THE BWFS.
On the pedigree diagram, the recombinant haplotype is coded by a numeral rather than a letter. * = site of crossover. Diab(etes) 1 = Type 1 diabetes. IDS = individual identity code.

the proband (42-3) form the only pair of diabetic siblings in the BWFS who share neither paternal nor maternal HLA-A-B haplotype. Because of the crossover they do share the maternal DR segment of the HLA region (cf. VIII-2-B-iii-a).

1-B-iii-d. Discussion of observations on recombination

The HLA AxB and BxDR recombination rates observed here are comparable with the rates of under 1% generally reported, and their 95% confidence limits definitely exclude the 12% intra-HLA rate initially reported in diabetic families by Suciu-Foca & Rubinstein (1976; see III-6-C-iii). The rate for recombination between HLA and GLO is similar to that in other reports.

1-B-iv. Zygotic assortment of HLA-B8: random or not?

1-B-iv-a. Background; the question of bias.

Early in 1979 the author noted apparent deviations from random segregation of HLA-B8 in the families genotyped for the BWFS. An initial analysis of 69 diabetic and 33 control families, each including a parent with a single A1-B8 haplotype, showed that (i) there was among the pool of offspring a significant excess of the 1-8 haplotypes as compared with the paired alternative haplotypes; and (ii) among offspring inheriting the 1-8 haplotype there was a significant male excess (Cudworth et al, 1979).

Controversy arose (III-6-B-iv) because the analysis did not allow for ascertainment bias: since families were ascertained through children with Type 1 diabetes, which is associated with A1, B8 and the male sex, inclusion of these Type 1 diabetic probands would inevitably bias the distribution in favour of 1-8 and male sex. Serjeantson (1980) stated that the probands should be excluded from the analysis, the distribution of haplotypes among the remaining offspring alone being used to give an unbiased view. In such an analysis she found no evidence of deviation from 'random'. However, careful consideration leads to rejection of Serjeantson's method since it introduces a contrary bias. A mathematical proof based on simple probability theory follows.

Let us consider a population of M_c families, each including a parent with a single 1-8 haplotype and c offspring of that parent. By our null hypothesis of random zygotic assortment, the probability p that a given offspring inherits the 1-8 haplotype is

$$p = 1 - p = 0.5 \qquad (1\text{-}1).$$

The number of occurrences of the 1-8 haplotype in a sibship of c offspring is a discrete random variable H, whose individual values h range from 0 to c with a binomial probability distribution:

$$Pr(H=h) = {}^cC_h . p^{c-h} . (1 - p)^h = \frac{c!}{h!(c-h)!} . 0{\cdot}5^c \qquad (1\text{-}2).$$

Thus the probability that a sibship includes one or more 1-8 haplotypes is

$$Pr(H>0) = 1 - Pr(H=0) = 1 - 0.5^c \qquad (1\text{-}3).$$

Let us now suppose that families where H > 0 (i.e. at least one of the offspring inherits 1-8) are ascertained to a study, and - for the purpose of illustration - that the ascertainment rate a is independent of the value of H when H > 0. A maximum-likelihood estimate for the number of families ascertained is given by

$$N_c = a.M_c.Pr(H>0) = a.M_c(1 - 0.5^c) \qquad (1\text{-}4),$$

which is derived from (1-3). These N_c families include $c.N_c$ offspring. The total number of occurrences of 1-8 in the offspring is expected to be

$$a.M_c . \sum_{h=1}^{c} h.Pr(h) = \frac{N_c}{1 - 0.5} . 0{\cdot}5^c \sum_{h=1}^{c} h.{}^cC_h \qquad (1\text{-}5).$$

Thus the proportion of the c.Nc offspring who show the 1-8 haplotype is given by

$$x_c = \frac{1}{c.N_c} . \frac{N_c}{1 - 0{\cdot}5^c} . 0{\cdot}5^c \sum_{h=1}^{c} h.{}^cC_h \qquad (1\text{-}6),$$

which can be expanded to:

$$x_c = \frac{0{\cdot}5^c}{c.(1 - 0{\cdot}5^c)} . \left[1 . \frac{c!}{1!(c-1)!} + 2. \frac{c!}{2!(c-2)!} + \ldots + c\right] \qquad (1\text{-}7).$$

If now one 1-8 positive offspring (representing the proband) is excluded from each family, 1-8 occurs in the following proportion of the remaining $N_c(c-1)$ siblings:

$$y_c = \frac{1}{N_c(c-1)} \cdot a.M_c \cdot \sum_{h=2}^{c} (h-1).Pr(h)$$

$$= \frac{0{\cdot}5^c}{(c-1)(1-0{\cdot}5^c)} \cdot \left[1 \cdot \frac{c!}{2!(c-2)!} + 2 \cdot \frac{c!}{3!(c-3)!} + \ldots + c - 1 \right] \quad (1\text{-}8).$$

The ratio of x_c to 0.5 represents the ascertainment bias affecting the proportion of 1-8 haplotypes among all the offspring. For families with two children, substituting c = 2 in Equation (1-7) gives

$$x_2 = \frac{0{\cdot}25}{2 \cdot 0{\cdot}75} \cdot [2 + 2] = 0{\cdot}67,$$

while for three children

$$x_3 = 0{\cdot}57,$$

and so on, with the bias becoming progressively smaller as family size increases.

The method proposed by Serjeantson does not give values for y_c of 0·5, as it would if it gave a satisfactory correction for the bias in x_c. Instead it leads to a contrary bias: from Equation (1-8),

$$y_2 = 0.33; \qquad y_3 = 0.36$$

In practice, in the present context the assumption that ascertainment rate a is independent of H, the number of 1-8 haplotypes in the family, is unlikely to be true. Diabetes rather than the 1-8 haplotype formed the principal criterion for ascertainment to the BWFS; the association between diabetes and 1-8 is relatively weak, so the greater is H then the more likely is diabetes to occur in a family and hence the more likely is its ascertainment. Conversely, many families ascertained through diabetes have no 1-8 haplotype while in many that do the haplotype does not occur in the proband, thereby weakening the bias. To take account of these additional factors would lead to great complexity in the mathematical model and is not appropriate here.

The conclusion remains that any analysis of segregation of haplotypes or allotypes associated with Type 1 diabetes in these families should include probands as well as their siblings, but should be so

constructed as to allow valid interpretation irrespective of ascertainment bias. An attempt at such an analysis follows.

1-B-iv-b. Approach

A relatively simple approach is to consider the relative distributions of the haplotype concerned and of the alternative haplotypes among all available offspring of heterozygous parents, with respect to the following variables or factors: the parent concerned (father or mother), the sex of the offspring, and the diabetic status of the offspring at entry to the study (proband, other Type 1 diabetic or not having Type 1 diabetes). To consider only the haplotype A1-B8 and not B8 without A1, for example, might be misleading, but to include A1 (present or absent on the haplotype) and B8 as separate variables in the analysis would lead to an unacceptable degree of subdivision of the available data in the cross-classification of 48 categories (2x2x2x2x3). Therefore only the B8 locus is considered here. Further simplification is obtained by amalgamating the group of 7 haplotypes in diabetic siblings with those in probands, since inspection shows that they are similarly distributed (see below).

The 4-way contingency table (with 16 cells) so obtained is analysed with the help of GLIM, using a log-linear model structure (Poisson error distribution). The relative strengths and statistical significance of the interaction terms are assessed, first to seek evidence for an interaction between B8 and male sex independent of the interactions of both of the latter with diabetes, secondly to determine whether the relative distribution of the paternal B8-positive (B8+) and paired B8- haplotypes differs from that expected (i.e. 1:1, since only heterozygous fathers are considered) or from that of the corresponding maternal haplotypes, and thirdly to determine whether and, if so, how the table may be collapsed to facilitate further analysis (VI-3-C-i-b).

1-B-iv-c. Observations

The numbers of haplotypes in the BWFS offspring of B8-heterozygous parents, categorized as above, were obtained using GENOBASE. The total numbers of haplotypes included in the analysis are as follows: 136

TABLE VIII-7. DISTRIBUTION OF B8+ AND B8- HAPLOTYPES AMONG OFFSPRING OF 110 PARENTS HETEROZYGOUS FOR B8, categorized by offspring's sex and presence or absence of Type 1 diabetes (DM1) at entry, and by parent originating haplotype.

B HAPLOTYPE	B8 +				B8 -			
P PARENT	Paternal		Maternal		Paternal		Maternal	
S SEX	M	F	M	F	M	F	M	F
D DIABETES								
Proband	20	18	24	18	4	7	8	10
Other DM1	2	1	1	0	0	1	0	2
Total DM1	22	19	25	18	4	8	8	12
Not DM1	27	16	30	17	20	20	37	38
Sum over D	49	35	55	35	24	28	45	50
Sum over D+S	84		90		52		95	

paternal haplotypes in the HLA-typed offspring of 49 fathers, and 185 maternal haplotypes from 61 mothers. Table VIII-7 shows the observed distribution.

There are three reasons for the apparently surprising excess of maternal over paternal haplotypes in the Table. First, haplotypes from fathers of families where the proband's genotype showed expaternity were excluded as unrepresentative, but those of mothers in these families were included; hence the total numbers of paternal and maternal haplotypes differ. Secondly, while the phenotype frequencies of B8 in fathers and mothers of diabetic probands do not differ significantly (respectively, 35% and 40%), there are somewhat fewer B8-homozygous mothers than fathers (5/66 instead of 7/56), leaving 61 B8-heterozygous mothers and only 49 such fathers whose haplotype distributions are analysed.

Thirdly, B8+ mothers have had significantly more offspring than B8- mothers (mean numbers 3.26 and 2.68; $p = 0.031$ by Mann-Whitney test). This may reflect cultural/religious and genetic differences between mothers of Irish origin and the remainder: of the 14 mothers in the BWFS

who have had >4 live-born offspring, 10 are B8+, and of these 7 are recollected by the author as recognizably Irish, compared with only one of the four B8- mothers of large families. For fathers, on the other hand, there is no association between B8 status and number of offspring (2.93 vs 2.90). The non-significant excess of offspring of the B8+ mothers over those of B8+ fathers further accentuates the difference in numbers of haplotypes available for analysis. It could in theory result in a slightly smaller ascertainment bias for distribution of maternal than of paternal B8+ haplotypes with male sex (see above), but such a difference should be compensated by the structure of the analysis.

1-B-iv-d. Analysis

The four dichotomous cross-classification variables (see Approach, above) are abbreviated to B (B8 status of haplotype), P (parent originating haplotype), S (sex) and D (diabetes).

Table VIII-8 summarizes, for each hierarchical log-linear model contributing to the argument, the terms (parameters) included in the model, the deviance as likelihood-ratio criterion χ_L^2 (representing the differences between the observed frequencies and those expected according to the model in force), and the degrees of freedom. Selection of models was guided by the standardized estimates of the various parameters, most of which are for economy not shown here. Parameter estimates for Model 5 are, however, given in Table VIII-9.

The results can be interpreted as follows. Conclusions are enumerated in square brackets.

Model 5 includes only terms for grand mean, main effects and first-order interactions. No significant further improvement in fit is possible (as shown by the low value for deviance - less than 3.84, the 5% significance level of χ^2 for 1 degree of freedom), which means that there are no significant higher-order interactions. In particular, there is no need to invoke any influence of diabetes status on the association of male sex with B8. [Conclusion 1]

TABLE VIII-8. LOG-LINEAR ANALYSIS OF DATA SHOWN IN TABLE VIII-7, USING GLIM: INTERACTION EFFECTS. Only hierarchical models were investigated. All models shown here include terms for grand mean and main effects.

Model no.	Terms included: First-order interactions						Terms included: Higher-order	Deviance X^2_L	Deg. of freedom	Overall goodness of fit: p
	D.B	D.P	B.P	D.S	B.S	P.S				
2	-	-	-	-	-	-	-	38.59	11	<<0.001
5 *	+	+	+	+	+	+	-	0.3605	5	NS
8	+	-	+	+	+	-	-	0.5598	7	NS
10	+	-	-	+	+	-	-	6.022	8	NS
11	+	-	+	-	+	-	-	2.892	8	NS
12	+	-	-	-	+	-	-	8.354	9	NS
13	+	-	-	+	-	-	-	12.97	9	NS
14	+	-	-	-	-	-	-	13.64	10	NS
15	-	-	-	-	+	-	-	33.30	10	<0.001

NOTE: * Parameter estimates for Model 5 are given in Table VIII-9.

TABLE VIII-9. LOG-LINEAR MODEL 5 (IN TABLE VIII-8): PARAMETER ESTIMATES FOR FIRST-ORDER INTERACTIONS

Parameter	Estimate	Standard error	Standardized estimate
D(2).B(2)	1.260	0.2575	4.893
D(2).P(2)	0.0606	0.2461	0.2464
B(2).P(2)	0.5283	0.2414	2.189
D(2).S(2)	-0.3474	0.2468	-1.407
B(2).S(2)	0.6329	0.2401	2.636
P(2).S(2)	-0.0810	0.2315	-0.3499

Inspection of the standardized first-order parameter estimates for this model (Table VIII-9) shows that those for the D.P and P.S interactions are small (much smaller than the value of 1.96 which may indicate significance), suggesting that the terms representing these interactions may be omitted. This is confirmed by the minor increment in deviance produced by simplification to Model 8 (Table VIII-8). In contrast, the standardized estimate for the B.S parameter is fairly high at 2.636, suggesting that the interaction between B8 and sex is more important than that between diabetes and sex; indeed the latter is unlikely to be significant since the standardized estimate for D.S is low (1.407).

Examination of further models confirms this interpretation. Model 11 is preferable (i.e. it is the simplest model over which no significant improvement is possible). Substitution of a D.S term for the B.S term gives an inferior model (Model 10), with an increase in deviance for no change in degrees of freedom. Thus, not only is the association of B8 with sex independent of diabetes (see [1] above, but it is more marked than, and may indeed explain, any association of diabetes with sex which appears when the effect of B8 is ignored. [2]

The term for association of B8 with parent has also to be included for optimal fit, as suggested by its standardized estimate (2.189) in Table VIII-9. Simplification of Model 11 by removing this term (Model 12) causes a significant deterioration in fit (by subtraction, χ_L^2 = 5.462 for 1 DF, $p < 0.025$). The implication is that the ratio of B8+ to B8- paternal haplotypes differs from that for maternal haplotypes. [3]

Model 14, in which the only interaction term is B.D, is a significantly worse fit than Model 12 (χ_L^2 = 5.286, 1 DF, $p< 0.025$) and Model 11 (χ_L^2 = 10.748, 2 DF, $p < 0.01$). This is, however, the simplest model which is acceptable in an overall sense, unlike, for example, Models 2 and 13. Thus the association of B8 with diabetes within sibships cannot be ignored. [4]

Further examination of these implications is facilitated by collapsing the 4-dimensional contingency table over one or more variables to give a simpler data structure. Since Model 11 is a good fit and lacks terms D.P, P.S and D.S, it is acceptable to collapse over P or D in order to explore the association between B8 and sex.

Table VIII-10 shows the 2x2x2 table and some log-linear modelling results obtained after collapsing over the variable P. Again there is no second-order interaction. The simplest (and preferable) model is Model 21, with a term for association of B8 with sex (as well as one for B8 with diabetes), but none for diabetes with sex. That Model 20 is not acceptable confirms more strongly conclusion [2] already stated above. Indeed, it is clear on inspection that the preferential segregation of B8

to the male is at least as marked in non-diabetic as in diabetic offspring. Further collapse of this table over the variable D is acceptable (since D is independent of at least one other variable, in this case S), giving the values shown as 'sum over diabetes' in Table VII-10-a. This 2x2 table gives an odds ratio ('relative risk') of 1.67 for the significant association of B8 with male sex (2p = 0.0288, Fisher Exact test). ([2])

TABLE VIII-10. DISTRIBUTION OF B8+ AND B8- HAPLOTYPES AMONG OFFSPRING OF 110 PARENTS HETEROZYGOUS FOR B8, categorized by offspring's sex and presence or absence of Type 1 diabetes (DM1) at entry, IRRESPECTIVE OF PARENT ORIGINATING HAPLOTYPE.

(a) CONTINGENCY TABLE

B Haplotype	B8+		B8-	
S Sex	M	F	M	F
D Diabetes +	47	37	12	20
-	57	33	57	58
Sum over D	104	70	69	78

(b) GLIM ANALYSIS: MODELS INFORMATIVE FOR INTERACTION TERMS. Only hierarchical models were investigated. All models shown here include terms for grand mean and main effects.

Model ref. no.	Interaction terms included				Deviance X^2_L	Degrees of freedom	Overall goodness of fit: p
	D.B	D.S	B.S	D.B.S			
17	-	-	-	-	32.70	4	<0.001
19	+	+	+	-	0.1315	1	NS
20	+	+	-	-	7.081	2	<0.05
21	+	-	+	-	2.463	2	NS
22	-	+	+	-	26.74	2	<0.001

If Table VIII-7 is collapsed over D and then over S, a simplification which similarly appears acceptable, the 2x2 table relating B8 to paternal versus maternal origin is represented by the bottom line of Table VIII-7. The relative distribution of B8+ and B8- haplotypes is again dependent on whether they are paternal or maternal in origin (conclusion [3] above); $p < 0.05$). Inspection shows that the reason is a

significant deviation from random distribution of the paternal but not of the maternal haplotypes. Of the 136 paternal haplotypes examined, 84 (61.8%) are B8+, whereas the proportion expected is the same as the proportion in these fathers themselves, i.e. 50% (since only B8-heterozygous fathers are included). This difference between observed and expected is significant: $\chi_{L}^{2}{}_{(1)} = 7.0662$ with Yates' correction, $p<0.01$. ([3])

1-B-iv-e. Discussion

There is no doubt that, in the families studied here, the B8+ haplotype shows non-random segregation in that it occurs in offspring of heterozygous fathers significantly more frequently than does the alternative haplotype (conclusion [3]), while from either parent it segregates preferentially to male offspring ([2]). These findings cannot be explained by the recognized associations of B8 and male sex with diabetes ([1]); indeed the latter association appears to be secondary to the association of B8 independently with diabetes and with male sex ([2]).

Such unexpected findings have profound implications. Assuming that they are true, they imply that the B8 allele, or one in linkage disequilibrium with it, enhances fitness of the gamete or zygote. It is tempting to speculate that such an effect could maintain the frequency in the population of the apparently deleterious 'S1' diabetes susceptibility gene (see also 1-B-v-d).

It is possible that the peculiar segregation pattern observed here results from a chance peculiarity of the sample of families taken in the BWFS. Neither Weitkamp (1979) nor Hansen & Gurtler (1980) found any deviation from random zygotic assortment of HLA-A1-B8 in normal families. However, Miller, Rich and Barbosa (1981) confirmed a male excess among B8+ siblings after exclusion of diabetic probands. Further studies are needed to examine this question in a new sample of families, preferably randomly-ascertained, and to extend the analysis to other alleles such as DR3 and indeed any common allele. Such studies are beyond the scope of this work.

1-B-v. Siblings grouped by haplotype concordance with probands.

Henceforth the terms used for haplotype concordance will be abbreviated to H2 (HLA-identical by descent, sharing both haplotypes with proband), H1 (haplo-identical) and H0 (non-HLA-identical, sharing neither haplotype).

1-B-v-a. Distribution of siblings between groups

The 289 HLA-typed siblings of the probands in the BWFS may be grouped according to the the degree of haplotype concordance with the proband (Table VIII-11). Half-siblings are excluded, while siblings with intra-HLA recombinant haplotypes are here classified according to the B-locus concordance with the proband. For many subsequent analyses the 17 siblings showing ambiguous haplotype concordance (due to homozygosity of a parent) are excluded.

The distribution of the total number of siblings across the three unambiguous groups differs significantly from the 'expected' ratio of 25:50:25% ($p < 0.025$, χ^2 test) because of a deficit in the H2 group. This is discussed below (1-B-v-d).

TABLE VIII-11. HLA-TYPED SIBLINGS OF PROBANDS IN THE BWFS, GROUPED BY HAPLOTYPE CONCORDANCE WITH PROBAND.

Haplotypes shared	Monozyg. twin	2	2/1[a]	1	1/0[a]	0	TOTAL
Number of:							
brothers	1	22	5	90	4	34	156
sisters	0 [b]	27	4	59	4	39	134
both sexes	1	49	9	149	8	73	289
Number diabetic:							
brothers	0	1	0	1	0	0	2
sisters	0 [b]	3	0	3	0	1	7
both sexes	0	4	0	4	0	1	9

a. In 10 families the two haplotypes in one parent were indistinguishable, so that haplotype concordance is ambiguous. 2/1 = either 2 or one, 1/0 = 1 or 0 haplotype in common with the proband.
b. One identical twin sister of a proband is diabetic, but lives abroad and is not included here.

1-B-v-b. Distribution of diabetic siblings

Type 1 diabetes was already present at entry to the study in 9 of the 289 siblings, distributed as shown in Table VIII-11, and in a tenth sibling who was not typed but is known to be an identical twin of the proband. For certain analyses she, and the non-diabetic identical twin from another family who was typed, are included as H2 siblings. In that case 5 of the 51 H2 sibs were diabetic and on insulin at entry, and the distributions of diabetic and non-diabetic siblings among the three non-ambigous groups are significantly different ($X^2_{(2)}$ = 7.157, $p < 0.05$; linear trend analysis $X^2_{(1)}$ = 5.477, $p < 0.025$). If the identical twins are excluded the difference does not quite attain significance at the 5% level (linear trend analysis, $X^2_{(1)}$= 3.750). In contrast, if the siblings known to have B-DR recombinant haplotypes (see VII-1-B-iii) are reclassified according to DR-locus concordance, then with the twins included $X^2_{(1)}$ = 8.031, $p < 0.025$, linear trend $X^2_{(1)}$ = 7.644, $p < 0.01$; while without the twins the distribution is still significantly different from random by linear trend analysis ($X^2_{(1)}$ = 5.8392, $p < 0.025$).

1-B-v-c. Age distribution across haplotype concordance groups

The distribution of age of siblings at entry to the BWFS is approximately Normal (Figure VII-2). Analysis of variance shows that the siblings grouped by unambiguous haplotype concordance with their probands differ significantly in age, with an increasing trend from H2 through H1 to H0 haplotype concordance (Table VIII-12). The fact that probands (at 13.9 ± 4.7 y) were virtually identical in age to the siblings as a whole, and to the H1 group, has a bearing on interpretation of this finding (see below).

1-B-v-d. Discussion.

The observed distribution of siblings between the H2, H1 and H0 groups is not as surprising as it would clearly appear to some authors (Bengsch et al, 1978; Platz et al, 1981). At the International Symposium on Genetics of Diabetes Mellitus held in Freiburg-im-Breisgau, West Germany in 1981, the present author explained why the 25:50:25 ratio would not in fact be expected in these siblings, but would be distorted

TABLE VIII-12. HLA-TYPED SIBLINGS OF PROBANDS IN THE BWFS: AGE AT ENTRY versus HAPLOTYPE CONCORDANCE WITH PROBAND. Ambiguous haplotype concordance groups are excluded.

A. Mean and standard deviation of age at entry.

Results				Graphical representation
Haplotp concd'ce	No.	Age at entry Mean	SD	Age (years) 12.0 14.0 16.0 18.0
2	49	13.06	5.41	(--------*--------)
1	149	13.96	6.61	(----*----)
0	73	15.93	6.30	(-------*------)

B. Analysis of variance

Source	Degrees of freedom	Sum of squares	Mean square	F	Significance
Factor	2	285.6	142.8	3.56	p = 0.025
Error	269	10780.7	40.1		
Total	271	11066.4			

by the linkage already demonstrated between diabetes and HLA. (Gorsuch et al, 1982; Appendix 7-4).

The significant trend to increasing age with decreasing haplotype concordance is of great interest. It might be argued that it represents a bias due to the linkage of diabetes with HLA, the likelihood that, other things being equal, diabetic children are older than non-diabetic children (since the incidence rate continues at a high level throughout childhood), and the exclusion of probands from consideration. However, if that were the explanation then the probands as a group should be demonstrably older than the siblings included in Table VIII-11, and at least as old as the H0 group. Otherwise exclusion of probands could not distort the relative ages of H1 and H0 siblings. On the contrary, the mean ages of probands and of siblings overall are identical. The inescapable conclusion is that within sibships there is a predominance of susceptible children among those most recently born.

This is probably another manifestation of the phenomenon underlying the tendency for diabetes within sibships to be associated with increasing maternal age (Flood et al, 1982; Wagener et al, 1983), and may well be related to the preferential zygotic assortment of B8 (VII-1-B-iv). A speculative but plausible hypothesis, extending that already suggested (1-B-iv-e), is that a paternal gene in linkage disequilibrium with (or possibly identical to) HLA-B8 confers a protective effect against rejection of the developing conceptus by a mother already immunized by previous pregnancies, and that the same gene or one in association with it contributes to susceptibility to Type 1 diabetes.

1-C. BASIC IMMUNOLOGICAL OBSERVATIONS

This section is descriptive rather than analytical. Possible bearings on genetics of Type 1 diabetes are explored later (VIII-2 and 5).

1-C-i. Autoimmune and other diseases in diabetic probands

Five probands (two male) had at entry to the study diseases other than diabetes which are known or suspected to be autoimmune in origin. All five are HLA-B8+. Two probands, a girl aged 7 and a boy aged nearly 14, were on gluten-free diets for recently-diagnosed coeliac disease. Two girls had thyroid disease: a 17-year-old with incipient hypothyroidism (raised plasma TSH) diagnosed at the time of entry, and a 21-year-old with Graves' disease four years previously. Finally, one boy of 19 had vitilligo.

Asthma, eczema, hay fever or allergies to drugs or foods were reported by 29 (36%) of 80 informative male and 13 (22%) of 60 informative female probands.

1-C-ii. Family history of relevant diseases

Table VIII-13 gives the numbers and proportions of families reporting known diabetes in different types of relatives of the proband. Estimates of prevalence in second-degree relatives are based on the rather arbitrary assumption that each proband has 10 such relatives, and

TABLE VIII-13. DIABETES IN RELATIVES OF PROBANDS AT ENTRY TO BWFS

Type of relative	Siblings	Parents	2nd degree relatives	3rd degree relatives
	No. (%)	No. (%)	No. (%)	No. (%)
Insulin-treated:				
Families reporting	8 (5)	9 (5)	21 (13)	27 (16)
Individuals	10 (3)	9 (3)	>21[a] (>1.3)[b]	>27[a] (>0.5)[b]
Not insulin-treated:				
Families reporting	1 (0.6)	3 (2)	18 (11)	33 (20)
Individuals	1 (0.3)	3 (1)	>18[a] (>1.1)[b]	>33[a] (>0.7)[b]
Treatment not known:				
Families reporting	0	0	2	8

NOTES: a. Multiple occurrences of the same type of diabetes in the same group of relatives of an individual proband are not separately recorded.
b. Estimates of prevalence are based on the assumption that each proband has 10 2nd-degree and 30 third-degree relatives, and are therefore approximate estimates.

are minimum estimates since multiple occurrences of a disease among the 2nd-degree relatives of a proband are not distinguished from single cases in the database. For third-degree relatives the same applies, but each proband is assumed to have 30. The figure for non-insulin-treated diabetes in 3rd-degree relatives is probably unreliable, because many families lose contact in adult life.

The frequency of insulin-treated diabetes in 2nd-degree relatives does appear higher than would be expected given a population incidence of about 0.3% by age 30, supporting the existence of an hereditary factor in this disease, while the frequency of non-insulin-treated diabetes is about that expected. This difference between the two clinical types of diabetes supports the earlier reports of diabetes breeding true to type with respect to age at onset (III-5-D).

The prevalence of thyroid and certain other diseases in siblings and parents, and the frequency of reports in second-degree relatives, are shown in Table VIII-14. Past or present thyrotoxicosis affected 5.5 % of mothers, with confidence limits 3.1 - 7.9%, clearly higher than would be

TABLE VIII-14. CERTAIN DISEASES OTHER THAN DIABETES IN RELATIVES OF PROBANDS AT ENTRY TO THE BWFS. Figures refer to both past and present disease. Siblings are omitted since none was affected by any of these conditions.

Type of relative	Parents (prevalence)		2nd degree relatives	
	Fathers No. (%)	Mothers No. (%)	Families No. (%)	Prevalence[a] estimate
Thyroid disease:				
All types	3 (2)	13 (8)	27 (16)	>1.6
Thyrotoxicosis	2 (1)	9 (5.5)	17 10	>1
Myxoedema	0	4[b] (2)	7 (4)	>0.4
Exophthalmos only	0	1	0	
Type uncertain	1	0	5 (3)	>0.3
Pernicious anaemia	0	1[b]	2	0.1
Vitilligo	2	3 (2)	7 (4)	>0.4
Addison's disease	1	0	0	
Rheumatoid arthritis	0	3 (2)	18 (11)	>1
Multiple sclerosis	1	1	0	

NOTES a. See note b to Table VIII-13
b. One mother has both myxoedema and pernicious anaemia.

expected in such women (aged 40 ± 7 years) in the general population, about 1% (Barker & Phillips 1984). Other autoimmune diseases are insufficiently frequent for comparisons to be meaningful.

Among informative non-diabetic relatives, allergies were reported by 25% and 29% of the 130 fathers and 137 mothers respectively, by 28% of 130 brothers and by 33% of the 102 sisters.

1-C-iii. Auto-antibodies in Type 1 diabetic subjects.

Table VIII-15 records the frequencies of positive tests for various organ-specific and antinuclear auto-antibodies in Type 1 diabetic subjects when first tested, either on entry to the BWFS or when ascertained to the supplementary study of newly-diagnosed diabetic children.

Islet-cell antibody (ICA) results are classified as positive according to two alternative criteria: first, immunofluorescence

VIII-15. ORGAN-SPECIFIC AND ANTINUCLEAR AUTO-ANTIBODIES IN DIABETIC SUBJECTS AT ENTRY TO THE BWFS, OR AT DIAGNOSIS IN SUPPLEMENTARY STUDY

A. Islet-cell antibodies

		BARTS-WINDSOR FAMILY STUDY					NEWLY-DIAGNOSED		
		Children			Parents		(not in BWFS)		
	Sex	M	F	Both No. (%)	Fat.	Mot.	M	F	Both No. (%)
Total number tested		92	82	174 (100)	7	2	20	10	30 (100)
Type of ICA									
Conventional									
w+ or stronger		31	27	58 (33)	0	1	18	8	26 (87)
+ or stronger		13	8	21 (12)	0	1	10	6	16 (53)
Complement-fixing									
w+ or stronger		17	14	31 (18)	0	1	11	6	17 (57)
+ or stronger		12	6	18 (10)	0	1	5	4	9 (30)
'Second pattern'[a]		0	2	2 (1)	0	0	0	0	0

B. Other antibodies

	BARTS-WINDSOR FAMILY STUDY				NEWLY-DIAGNOSED	
	Children		Parents		(not in BWFS)	
Sex	Male No (%)	Female No (%)	Fat.	Mot.	Male No (%)	Female No (%)
Thyroid:						
Microsomal	9 (10	18 (22)	2	1	3 (15)	1 (10)
Thyroglobulin	9 (10)	16 (26)	1	0	1 (5)	1 (10)
Either or both	15 (16)	23 (28)	2	1	3 (15)	1 (10)
Gastric parietal cell	6 (7)	9 (11)	1	0	4 (20)	0
Adrenal steroidal cell[b]	2	1	0	0	1	0
Antinuclear						
Diffuse pattern	2 (2)	6 (8)	1	1	0	0
Membranous	14 (15)	12 (15)	1	1	0	0
Speckled	3 (3)	7 (9)	0	0	0	0
Any pattern(s)	16 (17)	17 (21)	1	1	0	0

NOTES a. 'Second-pattern' ICA (staining only a few peripheral cells) was specific for glucagon cells in one and somatostatin cells in another diabetic girl.

b. In general, only those subjects positive for other anti-bodies were tested for adrenal antibodies.

reactions graded as 'weakly positive' or more intense, unless all subsequent tests on the same individual proved negative (see VI-METHODS-2); and secondly, reactions graded '+' or more intense. Whichever criterion is used, the results are in line with earlier reports (LITERATURE SURVEY, III-7-C-i). The well-known association of ICA with short duration of Type 1 diabetes is present (see VIII-3-B-i-b). It is interesting that ICA appears to be at least as frequent in diabetic boys as in the girls, and possibly more so.

Variations in frequency of positive ICA results with HLA phenotypes and duration of diabetes are covered later (VIII-2, -3).

The prevalence of thyroid antibodies in the diabetic children here, as in other reports, is substantially greater than would be expected in non-diabetic children generally (III-7-B), and is significantly greater than in the non-diabetic siblings within the BWFS families ($p < 0.05$ for each sex). As expected, but unlike ICA, thyroid antibodies appear predominantly in the girls, although here the numbers are small and the sex difference is not significant ($0.05 < p < 0.1$). In the extended sample analysed for heterogeneity, however, the sex difference is significant (VIII-2-B-ii-e).

Antinuclear antibodies show no predilection for either sex in these diabetic children (Table VIII-15), nor with presence of diabetes (compare Table VIII-16).

1-C-iv. Auto-antibodies in non-diabetic subjects

At entry to the BWFS a total of 22 (3.8%) of non-diabetic subjects were positive for ICA by the conventional test if 'weakly positive' results are included, or 10 (1.7%) if not (Table VIII-16). Complement-fixing (CF-) ICA was positive in 11 non-diabetic family members overall. By both techniques the frequency of positive results appears somewhat greater in siblings than in parents, but not significantly so, nor is there any sex difference.

TABLE VIII-16. AUTO-ANTIBODIES IN NON-DIABETIC SUBJECTS AT ENTRY TO THE BARTS-WINDSOR FAMILY STUDY

A. Islet-cell antibody

	Children			Parents		
	M	F	Both No. (%)	Fat.	Mot.	Both No. (%)
Total number tested	154	125	279 (100)	146	161	307 (100)
Type of ICA						
Conventional						
w+ or stronger	9	5	14 (5)	4	4	8 (3)
+ or stronger	5	3	8 (3)	0	2	2 (0.7)
Complement-fixing						
w+ or stronger	4	4	8 (3)	1	2	3 (1)
+ or stronger	4	3	7 (3)	0	1	1 (0.3)
'Second pattern'[a]	1	0	1 (0.3)	1	2	3 (1)

B. Other antibodies

Specificity	Children		Parents	
	Male No. (%)	Female No. (%)	Fathers No. (%)	Mothers No. (%)
Thyroid:				
Microsomal	8 (5)	9 (7)	24 (16)	59 (37)
Thyroglobulin	5 (3)	15 (12)	15 (10)	27 (17)
Either or both	9 (6)	19 (15)	30 (21)	62 (39)
Gastric				
parietal cell	3 (2)	3 (2)	5 (3)	24 (15)
Adrenal				
steroidal cell[b]	0	0	2 (1)	3 (2)
Antinuclear:				
Diffuse pattern	5 (3)	12 (10)	20 (14)	30 (19)
Membranous	22 (14)	23 (18)	20 (14)	34 (21)
Speckled	9 (6)	12 (10)	28 (19)	9 (6)
Any pattern(s)	29 (19)	36 (29)	36 (25)	59 (37)

NOTES a. 'Second-pattern' ICA (staining only a few peripheral cells) was specific for glucagon cells in the child and somatostatin cells in the parents.

b. In general, only those subjects positive for other anti-bodies were tested for adrenal antibodies.

Thyroid, gastric and adrenal autoantibodies are clearly more common in the adults than in the children. This age difference is significant at the 1% level for thyroid microsomal antibody and gastric parietal-cell antibody in females, and for 'either or both' thyroid antibody in both sexes. It is significant at the 2.5% level for thyroglobulin antibody in males but not females.

The sex difference in frequency of thyroid antibodies ('either or both') is significant in both siblings ($p < 0.05$) and adults ($p < 0.001$). When all non-diabetic subjects are pooled, ANA (any pattern) is significantly more common in females than in males ($2p = 0.002$), although adults show a contrary effect of sex on speckled-pattern ANA frequency ($2p = 0.0003$)

TABLE VIII-17. ISLET-CELL ANTIBODY IN SUBJECTS AT ENTRY TO THE BWFS: CROSS-CLASSIFICATION BY TECHNIQUE USED (conventional versus complement-fixing immunofluorescence tests)

Technique	Number of results classified by intensity of fluorescence						
Conventional	0	v+	+	++,+++	No. >0	No. >v+	Grand total
Complement-fixing							
0	718	34	6	2	42	8	760
v+	1	14	1	0	15	1	16
+	0	2	9	1	12	10	12
++,+++	1	0	8	5	13	13	14
No. >0	2	16	18	6	40	24	42
Grand total	720	50	24	8	82	32	802

1-C-v. Correlation between ICA results by the two techniques

Table VIII-17 is a contingency table classifying the results of the initial tests on all informative BWFS members by both conventional and complement-fixing immunofluorescence. Overall there is a strong and highly significant association between the the results by the two methods, shown by χ^2 for mutual independence (749.9, 9 degrees of

freedom) and for linear trend (449.8, 1 degree of freedom) as well as by Goodman & Kruskal's gamma (0.981). The latter is interpreted as follows: if two individuals, A and B, are chosen at random, if the probability that their ICA results are ordered likewise by the two techniques (A stronger than B by both methods, or A weaker than B by both) is Pr(L), and if the probability that the results are ordered differently is Pr(U), then Pr(L) - Pr(U) = 0.981. Among the 42 patients positive by the conventional technique, these respective statistics remain highly significant (42.70, 6 degrees of freedom; 32.22, 1 d. freedom; gamma = 0.756). Thus there is a correlation between the results by the two tests.

It might be argued that the CF-ICA test is unnecessary if a result strongly positive by the conventional test gives the same information. However, that is not the case. If conventional ICA is considered as a predictor of a positive CF-ICA result (VII-3-C-i-a), when w+ results are counted as negative the sensitivity is only 57% but the specificity is 99% and the predictive value 75%. If w+ conventional tests are counted positive the sensitivity increases considerably to 95%, but specificity falls to 95%, with predictive value only 49%. Therefore, despite the correlation, the two tests do to some extent give different information.

1-C-vi. Discussion of basic autoimmunity results.

These findings are essentially confirmatory. Of particular interest are the excess of thyrotoxicosis in mothers of children with Type 1 diabetes (1-C-ii), the significant excess of positive thyroid antibody results in diabetic children, and the differences in occurrence of ICA and the other antibodies with respect to age and sex (1-C-iii, -iv), already mentioned in the LITERATURE SURVEY, III-7-C-i.

The lack of close correlation between conventional and CF-ICA results supports the view that the latter technique identifies a form of ICA which is not identical to (although it may be a subset of) that determined by the conventional test. The complement-fixation technique does not merely pick up the sera more strongly reactive in the conventional test. In later chapters the predictive values of the two techniques for diabetes are considered.

PART VIII — CHAPTER 2

HLA AND TYPE 1 DIABETES: MODE OF INHERITANCE

CONTENTS

A Introduction · 236
B Observations and analyses · 237
i Linkage: combined data from diabetic sib pair studies (with reference to published work) · · · · · · · · · · 237
ii Heterogeneity of Type 1 diabetes in unrelated individuals · 239
a Sources of data · 239
b Method: discrete multivariate analysis · · · · · · · · · 239
c Effects of age at entry (E) and source of data (S) · · 240
d Interactions of age at diagnosis (A,) HLA-B (B) and sex (G) in 559 patients · · · · · · · · · · · · · · · 244
e Interactions of thyrogastric antibody status (T) and season at diagnosis (Q) with other factors · · · 250
iii Further analyses relevant to mode of inheritance · · · · · 263
a Hardy-Weinberg equilibrium in diabetic probands · · · 263
b HLA association: B8,w62 versus B8 or Bw62 homozyg'y · 266
c Role of HLA genes themselves: · · · · · · · · · · · · · 270
specific HLA haplotypes and diabetes; · · · · · · · 270
specific HLA genes and linkage with diabetes · · 272
C Discussion · 275

2-A INTRODUCTION

In this chapter, diabetic sib-pair data from the literature are analysed to determine whether they fit simple dominant or recessive single-gene hypotheses, and are tested for heterogeneity between sibships with two and those with more members affected. Next, possible heterogeneity within Type 1 diabetes is further investigated by discrete multivariate analysis of data on unrelated subjects.

Another question considered is whether any of the HLA marker genotypes studied deviates significantly from Hardy-Weinberg equilibrium (IV-B). Then the strengths of linkage disequilibria in haplotypes

present in the Barts-Windsor Family Study are compared with those in 'control' haplotypes present in the same families. Finally, diabetic sib-pair sharing data from the combined sources are analysed to determine whether linkage between HLA and diabetes is influenced by the HLA-B genes present. These last two analyses are relevant to the question whether HLA genes themselves contribute to susceptibility, while a significant result in the latter might alternatively indicate genetic heterogeneity.

2-B. OBSERVATIONS AND ANALYSES

2-B-i. Diabetic sib pairs: combined data from studies worldwide.

Early in 1981, the author gathered together for analysis all data then available on haplotype sharing in Type 1 diabetic sib-pairs, and reported the results to the 2nd International Workshop on Genetics of Diabetes Mellitus (Gorsuch et al, 1982b, included here as appendix 7-4; pp 45-47). The 'UK study' referred to in that paper represents multiplex sibships from the Barts-Windsor Family Study, the BDA UK multiplex study, the Liverpool study and miscellaneous families, as defined above (VII-B, -C).

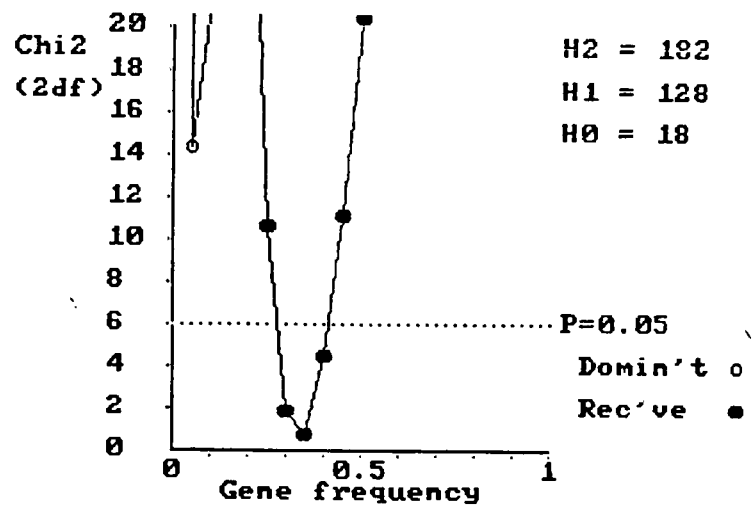
Table 2 from Appendix 7-4 is reproduced here as Table VIII-18. The method of Thomson and Bodmer (1977) is applied, with the aid of the author's BASIC program TOMBOD (METHODS, VI-3-C-iii), to determine whether simple recessive inheritance is still possible given such large numbers

TABLE VIII-18. HLA HAPLOTYPE CONCORDANCE IN PAIRS OF TYPE 1 DIABETIC SIBLINGS. Combined data, 1981.

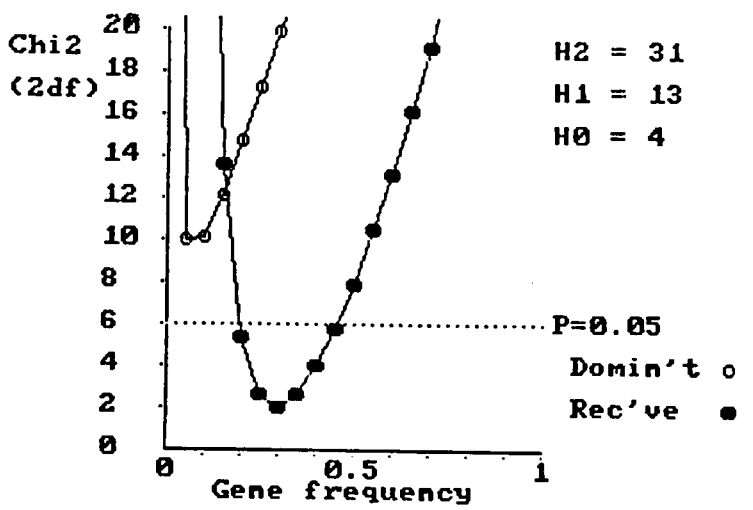
Number of diabetic sibs in family	Number of haplotyes shared			
	2	1	0	Total
2	182 (55%)	128 (39%)	18 (5%)	328 [c]
3 or 4 [a]	31 (65%)	13 (27%)	4 (8%)	48
2, 3 or 4: Observed	213 (57%)	141 (38%)	22 (6%)	376 [d]
Expected [b]	94	188	94	376

[a] One pair selected from each sibship, either at random or as first diagnosed diabetic; [b] "Expected" = numbers expected if no linkage between HLA and diabetes; [c] No significant heterogeneity: χ^2 (2 df)=2.78, p=0.25; [d] significance of difference: χ^2 (2df)=217.5: $p<<10^{-10}$.

A. Duplex



B. Triplex/tetraplex



C. All

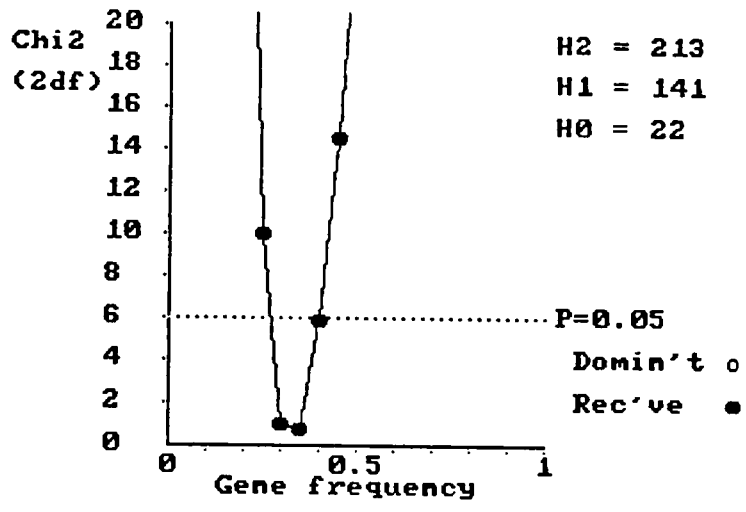


FIGURE VIII-2. HLA HAPLOTYPE CONCORDANCE IN PAIRS OF TYPE 1 DIABETIC SIBLINGS FROM 376 MULTIPLEX SIBSHIPS. Goodness-of-fit for single-gene recessive and dominant models of inheritance: variation of Chi-squared with hypothetical gene frequency. Method of Thompson and Bodmer (1977), using data of Table VIII-18. A - pairs from 328 duplex sibships; B - one pair from each of 48 triplex or tetraplex sibships; C - all 376 pairs. Graphs produced by program TOMBOD (Appendix 5f) and reduced.

of observations (see LITERATURE SURVEY III-6-E-ii). FIGURE VIII-2 shows the variation of χ^2 with putative gene frequency. The horizontal line represents the critical value of χ^2 of 5.991 for goodness of fit at the 5% level, with two degrees of freedom. Clearly, whether the sibships concerned have two only (duplex; Fig. VIII-2-A) or three to four affected members (triplex, tetraplex; Fig. VIII-2-B), the dominant hypothesis is rejected at all gene frequencies. The recessive hypothesis is acceptable in both groups of sibships, with almost identical gene frequencies for optimal fit. In the pooled data, the recessive model gives an acceptable fit at gene frequencies between 0.27 and 0.40, with best fit a little below 0.35 where $\chi^2 = 0.773$ (Figure VIII-2-C).

2-B-ii. Heterogeneity of diabetes in unrelated individuals

2-B-ii-a. Sources of data

Type 1 diabetic individuals available for this study include probands from the Barts-Windsor Family Study and the BDA multiplex study (augmented), the Liverpool cases of Professor Cudworth, and 30 additional newly-diagnosed cases ascertained by the author. (VII-B, -C). Of over 600 unrelated diabetic patients ascertained, 589 different subjects are included in these analyses, the remainder being insufficiently informative. The numbers included in individual analyses vary because of missing information - for example, neither season at diagnosis nor thyrogastric antibody results were available to the author from the Liverpool study.

2-B-ii-b. Analytical methods

Discrete multivariate analysis was performed by log-linear modelling (VI-3-C-i-b). The variables included can be considered in three broad categories (Table VIII-19): those relevant to pathogenesis of diabetes itself, those which might hypothetically discriminate between different genetic forms of the disease, and possible confounding variables.

Clearly it would be futile to subdivide the data simultaneously by all the variables and categories defined in Table VIII-19, since the corresponding multi-dimensional contingency table, with 3840 cells, would consist largely of zeroes. First, therefore, the effects of the possible

TABLE VIII-19. VARIABLES CONSIDERED IN DISCRETE MULTIVARIATE ANALYSIS FOR HETEROGENEITY WITHIN TYPE 1 DIABETES.

Type	Code	Full name	Short name	Categories
r	A	Age at diagnosis	AgeD	5 (5-year tranches: 0-4,5-9,10-14, 15-19,20+)
r	Q	Quarter (season) of year at diagnosis	Qtr	4
r,d	N	No. of siblings diabetic at entry	NDChil	2 (simplex, multiplex)
d	B	HLA-B status	B8 etc.	2 in any one analysis (+ and -)
d	T	Thyrogastric	ThG	2 (+ and -)
d,c	G	Sex (gender)	Sex	2
c	E	Age at entry to study	AgeE	4 0-9,10-14,15-19, 20+
c	S	Source of data	Source	3 1=BWFS=/-new-diagn. 2=BDA multiplex 3=Liverpool

Type: r = relevant to pathogenesis of diabetes; d = potential discrimininant between subtypes of Type 1 diabetes; c = potential confounding variable.

confounding variables E (age at entry to study) and S (source of data) were examined to assess whether they could be omitted from further consideration. Next, multidimensional contingency tables with complexity as great as could feasibly be examined were screened to guide further simplification by collapsing over variables (see METHODS VI-3-C-i-b), reduction in number of categories of A (age at diagnosis of diabetes), or partitioning where necessary. Finally the simplified tables were examined for significant interactions relevant to diabetes. HLA-B status was treated as a dichotomous variable with categories 1 (positive) and 2 (negative), the whole analysis being performed three times to seek interactions with B8, Bw62 and B8,w62 respectively.

2-B-ii-c. Effects of age at entry (E) and source of data (S)

Tables VIII-20 and VIII-21 show results of analyses screening the pooled probands of the Barts-Windsor Family Study, BDA multiplex study and newly-diagnosed patients for significant interactions involving E (age at entry to study). Common sense suggests that A (age at

diagnosis), T (thyrogastric antibody status) and perhaps N (number of affected siblings) should be considered here.

The first set of possible interactions to be examined was that involving E, A and T (Table VIII-20). Liverpool data were not included since no data on T were available. The lack of significant association with T is less surprising than might at first appear, since most of the probands were young, and means that E can be neglected so far as other interactions with T are concerned.

TABLE VIII-20. THYROGASTRIC ANTIBODY STATUS (T), AGE AT DIAGNOSIS OF DIABETES (A) AND AGE AT ENTRY TO STUDY (E) in 308 newly-diagnosed cases and probands of the Barts-Windsor and augmented BDA Multiplex family studies.

A. FREQUENCY DATA (numbers of individuals in the respective cross-classification categories):

T	0 (negative)				1 (positive)				Prop'n T-positive (%)
E code years	1 0-9	2 0-14	3 5-19	4 20-40	1 0-9	2 0-14	3 5-19	4 20-40	
A code years									
1 0- 4	19	15	10	8	4	7	5	1	25
2 5- 9	19	34	20	5	8	11	5	2	25
3 10-14	0	42	20	11	0	12	8	4	25
4 15-40	0	0	13	10	0	0	7	2	28
Proportion T-positive in E category (%)					24	25	28	25	

B. LOG-LINEAR ANALYSIS: INTERACTION EFFECTS. All models shown here include terms for grand mean and main effects.

Model ref. no.	Terms included				Overall goodness/fit			Difference				
	1st-order			2nd				From model:				
	E.A	E.T	A.T	E.A.T	X^2_L	DF	p		X^2_L	DF	p	Term
4	+	+	+	-	4.418	9	NS	Sat.	4.418	9	NS	E.A.T
5	+	-	+	-	5.344	12	NS	4	0.926	3	NS	E.T
6	+	-	-	-	5.511	15	NS	5	0.167	3	NS	A.T
2	-	-	-	-	119.6	24	<0.001	6	114.1	9	<.001	E.A *

* Note: this term is significant only if structural zeroes are included in the analysis - cf. Table VIII-21.

Later, when the second-order interaction Q.A.N was found to be significant (see 2-B-ii-e), Q, E, A and N were included in a further analysis (Table VIII-21), for which Liverpool data were again uninformative. The interaction of E with N is expected because of ascertainment bias: older probands have had more time for their siblings to develop diabetes than have younger probands.

TABLE VIII-21. SEASON AT DIAGNOSIS OF DIABETES (Q), NUMBER OF AFFECTED SIBLINGS (N), AGE AT DIAGNOSIS OF DIABETES (A) AND AGE AT ENTRY TO STUDY (E) in 308 unrelated Type 1 diabetic subjects (newly-diagnosed cases and probands of the Barts-Windsor and augmented BDA Multiplex family studies).

A. FREQUENCY DATA (numbers of individuals in the respective cross-classification categories):

N		1 (simplex)				2 (multiplex)			
E code		1	2	3	4	1	2	3	4
years		0-9	10-14	5-19	20-40	0-9	0-14	5-19	20-40
Q	A								
1	1	1	2	2	1	4	5	3	1
	2	3	8	3	0	1	7	4	3
	3	0*	22	6	2	0*	1	3	2
	4	0*	0*	7	3	0*	0*	0	1
2	1	3	4	3	0	3	3	1	2
	2	7	2	2	0	2	4	3	2
	3	0*	9	4	2	0*	3	3	1
	4	0*	0*	5	2	0*	0*	1	2
3	1	6	2	2	0	1	4	1	2
	2	1	7	1	0	2	5	2	0
	3	0*	2	6	0	0*	1	5	3
	4	0*	0*	2	0	0*	0*	2	1
4	1	5	1	3	3	2	2	1	0
	2	9	8	5	1	2	4	5	3
	3	0*	9	1	1	0*	7	1	3
	4	0*	0*	3	0	0*	0*	0	3

NOTE: * Structural zero (see text).

B. LOG-LINEAR ANALYSIS: INTERACTION EFFECTS. All models shown here include terms for grand mean and main effects.

Model ref. no.	Terms included											Overall goodness/fit			Difference				
	1st-order						2nd				3rd	X^2_L	DF	p	Model	X^2_L	DF	p	Term
	Q.A	Q.E	Q.N	A.E	A.N	E.N	Q.A.E	Q.A.N	Q.E.N	A.E.N	QAEN								
5	+	+	+	+	+	+	+	+	+	+	-	18.98	18	NS	Sat.	18.98	18	NS	Q.A.E.N
6	+	+	+	+	+	+	+	+	-	+	-	26.48	27	NS	5	7.50	9	NS	Q.E.N
7	+	+	+	+	+	+	-	+	-	+	-	56.17	45	NS	6	26.69	18	.05	Q.A.E
8	+	+	+	+	+	+	-	+	-	-	-	66.62	51	NS	7	10.45	6	NS	A.E.N
9	+	+	+	+	+	+	-	-	-	-	-	86.68	60	<0.025	8	20.08	9	<.025	Q.A.N
10	+	-	+	+	+	+	-	+	-	-	-	74.95	60	NS	8	8.33	9	NS	Q.E
11	+	-	+	-	+	+	-	+	-	-	-	80.35	66	NS	10	5.40	6	NS	A.E
12	+	-	+	-	+	-	-	+	-	-	-	100.4	69	<0.001	11	20.05	3	<.001	N.E

At first sight E appears to be strongly associated with A (age at diagnosis) (Table VIII-20). However, this association is entirely explicable by bias introduced because children entering the study at age 'e' years (where 'e' is the continuous variable coded for analysis as the categorical variable E) must by definition already be diabetic: age at diagnosis 'a' years cannot exceed 'e'. Thus, in Tables VIII-20 and VIII-21 there are structural zeroes in all cells of the data array for which E=1 and A=3 or 4, and similarly for cells where E=2 and A=4. In the log-linear analysis summarized in Table VIII-21 all such cells are zero-weighted so as to make no contribution to deviance (Bishop, Feinberg & Holland, 1975). After this correction the interaction E.A is in no way significant.

Hence E cannot explain any association between A and N or Q. This conclusion is supported by the finding that the simplest acceptable model investigated in Table VIII-21 (Model 11) includes no term in both A and E. It is true that adding the term Q.A.E to Model 6 to give Model 5 improves the fit by an amount marginally significant at the 5% level, but in the absence of significant first-order interactions Q.E and A.E (Models 9-10, 10-11) this is unlikely to reflect an interaction of any importance (Bishop, Feinberg & Holland, 1975).

Interactions between Q, A and N are described below (2-B-ii-e).

The other obvious confounding variable, S (source of data), would be expected a priori, and is later shown, to influence the distribution of diabetic subjects with respect to N, A and B: N because of varying ascertainment bias in favour of multiplex sibships (little or none in the BWFS, probably some in the Liverpool series, absolute in the BDA study); A because of differing ascertainment criteria with respect to age at diagnosis of diabetes (VII-A-iii-a, -C-i, -C-ii); and B because of both geographical variation in gene frequencies within the British Isles and improvement in HLA typing techniques between 1974-6 (Liverpool study) and 1978-80. The variable S clearly cannot be ignored, and is therefore

TABLE VIII-22. HLA-B8 (B), AGE AT DIAGNOSIS (A), SEX (G), NUMBER OF AFFECTED SIBLINGS (N) AND SOURCE OF DATA (S) in 559 unrelated Type 1 diabetic probands from the Barts-Windsor, augmented BDA multiplex and Liverpool studies.

A. FREQUENCY DATA (numbers of individuals in the respective cross-classification categories):

S Source	1 (BWFS)				2 (BDA multiplex)				3 (Liverpool)			
N NDChil	1 (simplex)		2 (multiplex)		1 (simplex)		2 (multiplex)		1 (simplex)		2 (multiplex)	
G Sex	1 (M)	2 (F)	1 (M)	2 (F)	1 (M)	2 (F)	1 (M)	2 (F)	1 (M)	2 (F)	1 (M)	2 (F)
B HLA-B8	+ -	+ -	+ -	+ -	+ -	+ -	+ -	+ -	+ -	+ -	+ -	+ -
A AgeD (yr)												
1 0- 4	11 10	8 9	1 0	3 0	- -	- -	10 7	7 7	2 3	4 9	1 2	1 2
2 5- 9	16 10	12 11	0 2	1 0	- -	- -	9 16	15 9	13 8	9 12	1 3	2 1
3 10-14	12 11	14 13	*0 *0	*0 *0	- -	- -	8 11	9 7	19 23	29 10	6 3	3 2
4 15-19	6 7	1 1	*0 *0	*1 *0	- -	- -	2 2	2 0	16 7	6 8	1 1	2 1
5 20-40	*0 *0	*0 *0	*0 *0	*0 *0	- -	- -	2 1	0 1	14 14	8 14	4 5	1 4

NOTES: "-" = structural zero; * = areas with many random zeroes.

B. LOG-LINEAR ANALYSIS: INTERACTION EFFECTS. All models shown here include terms for grand mean and main effects. Structural zeroes (indicated in the contingency table by "-") and areas with many random zeroes ("*") were zero-weighted for fitting models.

Model ref. no.	Terms included				Overall			Difference				
	1st-order	2nd	3rd	4th	X^2_L	DF	p	Model	X^2_L	DF	p	Term
	S S S S A A A B B N A B N G B N G N G G	S S S S S S A A A B A A A B B N B B N N	S S S S A A A A B B	S A								
	B N G N G G N G G G	B B N N N N G G G G	B NG									
2	- - - - - - - - - -	- - - - - - - - - -	- - - - -	-	146.6	74	<<.001					
3	+ + + + + + + + + +	+ + + + + + + + + +	+ + + + +	+	(no convergence)							
4	- - - - + + + + + +	- - - - - - + + + +	- - - - +	-	93.62	42	<.001					In S
5	+ + + - + + - + - -	+ + - + - - + - - -	+ - - - -	-	47.22	41	NS					In G
6	+ + + - + + - + - -	+ + - + - - + - - -	- - - - -	-	50.17	42	NS	5	2.95	1	NS	ABNS
7	+ + + - + + - + - -	- - - - - - - - - -	- - - - -	-	60.28	55	NS	6	10.11	13	NS	
8	+ - + - - + - - - -	- - - - - - - - - -	- - - - -	-	64.58	62	NS	7	4.30	7	NS	
10	+ - - - - - - - - -	- - - - - - - - - -	- - - - -	-	74.05	67	NS	13	5.57	1	<.05	SN
								2	72.5	7	<<.001	SA
11	+ - + - + - - - - -	- - - - - - - - - -	- - - - -	-	64.27	62	NS	13	4.21	4	NS	AB
12	+ - + - - - - - + -	- - - - - - - - - -	- - - - -	-	68.31	65	NS	13	0.17	1	NS	BG
13	+ - + - - - - - - -	- - - - - - - - - -	- - - - -	-	68.48	66	NS	8	3.90	4	NS	AN

included in all analyses except where partitioning of tables or lack of relevant interactions renders this unnecessary.

2-B-ii-d. Interactions between age at diagnosis (A,) HLA-B (B) and sex (G)

Tables VIII-22-A, VIII-24-A and VIII-27-A cross-classify 559 informative, unrelated Type 1 diabetic subjects from the BWFS, augmented BDA multiplex and Liverpool studies, according to the five variables A (5

categories), B, G, N and S defined as in Table VIII-19. All cells in the category S2 N1 are structural zeroes, since by definition there are no simplex cases in the augmented BDA multiplex study. Furthermore, in the categories S1 N2 A3, S1 N2 A4 and S1 A5, sampling zeroes predominate as a result of ascertainment criteria and the infrequency of multiplex sibships in the BWFS. The cells corresponding to all these uninformative categories are zero-weighted for the purpose of fitting linear models.

Table VIII-22-B summarizes results of analyses applied to the data of Table VIII-22-A (B = HLA-B8 status). Excessive subdivision of the data in this complex array explains the failure of GLIM to compute expected values by the saturated model (3). The unacceptable fit of Model 4, with no interaction terms in S, indicates that simplification by collapsing the table over S might be misleading. In contrast, however, removal of all interaction terms in G (sex) causes no significant deterioration in fit (Model 5). In particular, the term B.G makes very little contribution to variance (Models 12 - 13); i.e. there is no sign of any association between HLA-B8 and sex of these diabetic subjects. Terms in G are therefore excluded from the remaining models investigated.

As usual, standardized parameter estimates (not shown here) were examined as a guide to the order of removal of terms to produce successively simpler models. Only terms S.N (Models 10 - 13) and S.A (Models 10 - 2) prove to be significant at the 5% level. Both represent ascertainment effects and are irrelevant to the question of genetic heterogeneity of diabetes. These results indicate that simplification of the contingency table by collapsing over S (Table VIII-23) should permit valid interpretation concerning interactions other than those involving A.N, while the term A.N is significant in the simplified table only because of differences in ascertainment to the three studies.

There is no evidence in Table VIII-22 (Models 11 & 13) or Table VIII-23 (Models 10 - 11) for any association between HLA-B8 and either age at diagnosis or number of diabetics in sibship (standardized estimate for B.N = 0.03 in Model 7 of Table VIII-22)).

Continued on p. 248

TABLE VIII-23. HLA-B8 (B), AGE AT DIAGNOSIS (A), SEX (G), AND NUMBER OF AFFECTED SIBLINGS (N) in 559 unrelated Type 1 diabetic probands. Pooled data from three sources (cf. Table VIII-22).

A. FREQUENCY DATA (numbers of individuals in the respective cross-classification categories):

N NDChil	1 (simplex)				2 (multiplex)				Total
G Sex	1 (M)		2 (F)		1 (M)		2 (F)		
B HLA-B62/15	+	-	+	-	+	-	+	-	
A AgeD (yr)									
< 5-	13	13	12	18	12	9	11	9	97
5 -	29	18	21	23	10	21	18	10	155
10 -	31	34	43	23	14	14	12	9	181
15 -	22	14	7	9	3	3	5	1	65
20 -	14	14	8	14	6	6	1	5	70
All	109	93	91	87	45	53	47	34	559

B. LOG-LINEAR ANALYSIS: INTERACTION EFFECTS. All models shown here include terms for grand mean and main effects.

Model ref. no.	Terms included								Overall			Difference			
	First-order						2nd	3rd							
	AB	AG	BG	AN	BN	GN			X^2_L	DF	P	From	X^2_L	DF P	Term
2	-	-	-	-	-	-	-	-	51.87	32	<0.01				
4	+	+	+	+	+	+	+	-	9.002	4	NS				
5 a	+	+	+	+	+	+	-	-	24.06	17	NS	4	15.06	13 NS	
												5	0.48	2 NS	
9	+	+	+	+	-	-	-	-	24.54	19	NS	9	0.23	1 NS	BG
10	+	+	-	+	-	-	-	-	24.77	20	NS	10	4.41	4 NS	AB
11	-	+	-	+	-	-	-	-	29.18	24	NS	11	6.00	4 NS	AG
12	-	-	-	+	-	-	-	-	35.18	28	NS	2	16.69	4 <0.01	(AN)[b]

NOTES: a. Some parameter estimates for Model 5 are given in C. b. Ascertainment effect

C. LOG-LINEAR MODEL 5 (in Table B): PARAMETER ESTIMATES FOR FIRST-ORDER INTERACTIONS

Parameter	Estimate	Standard error	Standardized estimate	
			Individual	Mean (a)
A(2).B(2)	-0.1029	0.2608	-0.3946	
A(3).B(2)	-0.2428	0.2541	-0.9555	
A(4).B(2)	-0.3441	0.3287	-1.0469 .	
A(5).B(2)	0.2712	0.3202	0.8470	0.8110
A(2).G(2)	-0.1479	0.2610	-0.5667	
A(3).G(2)	-0.1527	0.2540	-0.6012	
A(4).G(2)	-0.7457	0.3356	-2.2220 *	
A(5).G(2)	-0.4333	0.3211	-1.3494 .	1.1848 .
B(2).G(2)	-0.8231E-01	0.1715	-0.4799	0.4799
A(2).N(2)	-0.1255	0.2651	-0.4734	
A(3).N(2)	-0.6747	0.2656	-2.5403 *	
A(4).N(2)	-1.175	0.3817	-3.0783 *	
A(5).N(2)	-0.7252	0.3441	-2.1075 *	2.0498 *
B(2).N(2)	0.2494E-01	0.1848	0.1350	0.1350
G(2).N(2)	-0.1258	0.1856	-0.6778	0.6778

NOTES (a) 'Mean' of absolute values of standardized estimates for all parameters involving a given pair of factors. Standardized estimates: . 1.0 to 1.96; * > 1.96

TABLE VIII-24. HLA-B15 (B), AGE AT DIAGNOSIS (A), SEX (G), NUMBER OF AFFECTED SIBLINGS (N) AND SOURCE OF DATA (S) in 559 unrelated Type 1 diabetic probands from the Barts-Windsor, multiplex and Liverpool studies.

A. FREQUENCY DATA (numbers of individuals in the respective cross-classification categories):

S Source	1 (BWFS)				2 (BDA multiplex)				3 (Liverpool)			
N NDChil	1 (simplex)		2 (multiplex)		1 (simplex)		2 (multiplex)		1 (simplex)		2 (multiplex)	
G Sex	1 (M)	2 (F)	1 (M)	2 (F)	1 (M)	2 (F)	1 (M)	2 (F)	1 (M)	2 (F)	1 (M)	2 (F)
B HLA-B15	+ -	+ -	+ -	+ -	+ -	+ -	+ -	+ -	+ -	+ -	+ -	+ -
A AgeD (yr)												
1 0- 4	9 12	4 13	0 1	1 2	- -	- -	4 13	4 10	0 5	2 11	1 2	0 3
2 5- 9	7 19	2 21	2 0	1 0	- -	- -	7 18	3 21	1 20	2 19	1 3	0 3
3 10-14	8 15	5 22	*0 *0	*0 *0	- -	- -	5 14	3 13	8 34	7 32	3 6	0 5
4 15-19	4 9	2 0	*0 *0	*1 *0	- -	- -	2 2	0 2	4 19	2 12	1 1	1 2
5 20-40	*0 *0	*0 *0	*0 *0	*0 *0	- -	- -	0 3	0 1	4 24	7 15	0 9	2 3

B. LOG-LINEAR ANALYSIS: INTERACTION EFFECTS. All models shown here include terms for grand mean and main effects. Structural zeroes (indicated in the contingency table by "-") and areas with many random zeroes ("*") were zero-weighted for fitting models.

Model ref. no.	Terms included: 1st-order (S S S S A A A B B G / A B G N B G N G N N)	2nd (S S S S S S A A A B / A A A B B G B B G G / B G N G N N G N N N)	3rd (S S S S A / A A A B B / B B G G G / G N N N N)	4th (S / A / B / GN)	Overall X^2_L	DF	p	Difference: Term	Model	X^2_L	DF	p
4	+ + + + + + + + + +	+ + + + + + + + + +	- - - - -	-	22.45	15	NS	Sat.	22.45	15	NS	3rd+4th
5	- - - - + + + + + +	- - - - - - - + + + +	- - - - +	-	109.3	42	<<.001	Sat.	109.3	42	<.001	All.S
6	+ + - + + - + + + -	+ - + - + - - + - -	- + - - -	-	47.89	40	NS	Sat.	47.89	40	NS	Most.G
9	+ + - + + - + + + -	- - - - - - - + - -	- - - - -	-	57.34	50	NS	7	9.45	10	NS	
8	+ + - + + - + + + -	- - - - - - - - - -	- - - - -	-	63.36	54	NS	9	6.02	4	NS	A.B.N
10	+ + - + - - + + + -	- - - - - - - - - -	- - - - -	-	67.79	58	NS	8	4.43	4	NS	A.B
11	+ + - + - - - + + -	- - - - - - - - - -	- - - - -	-	71.69	62	NS	10	3.90	4	NS	A.N
12	+ + - + - - - + - -	- - - - - - - - - -	- - - - -	-	73.21	63	NS	11	1.52	1	NS	B.N
13	+ + - + - - - - - -	- - - - - - - - - -	- - - - -	-	75.35	64	NS	12	2.14	1	NS	B.G
14	+ + - - - - - - - -	- - - - - - - - - -	- - - - -	-	80.92	65	NS	13	5.57	1	<.025	N.S
15	+ - - + - - - - - -	- - - - - - - - - -	- - - - -	-	83.29	66	NS	13	7.94	1	<.025	B.S
16	+ - - - - - - - - -	- - - - - - - - - -	- - - - -	-	88.86	67	<.05	14	"	"	"	"
17	- - - - - - - - - -	- - - - - - - - - -	- - - - -	-	161.4	74	<<.001	16	72.5	7	<<0.001	A.S
18	+ + - + + - + + + -	+ - - - - - - - - -	- - - - -	-	58.36	47	NS	8	5.00	7	NS	A.B.S
19	+ + + + + + + + + +	+ - - - - - - - - -	- - - - -	-	55.54	47	NS	8	7.82	7	NS	1st .G
20	+ + + + - - + + - -	- - - + - - - - - -	- - - - -	-	65.65	55	NS					
21	+ + + + - - + + - -	- - - - - - - - - -	- - - - -	-	69.29	57	NS	20	3.64	2	NS	B.G.S

C. LOG-LINEAR MODEL 8 (in Table B): PARAMETER ESTIMATES FOR FIRST-ORDER INTERACTIONS

Parameter	Estimate	Standard error	Stand'zd estimate		Parameter	Estimate	Standard error	Stand'zd estimate	
A(2).B(2)	0.4309	0.3204	1.3449	·	A(3).S(2)	0.4603	0.5972	0.7708	
A(3).B(2)	0.0216	0.3043	0.0711		A(3).S(3)	1.2130	0.3306	3.6691	*
A(4).B(2)	0.2827	0.3896	0.7256		A(4).S(2)	0.1865	0.8285	0.2251	
A(5).B(2)	0.0992	0.4267	0.2324		A(4).S(3)	1.6810	0.4159	4.0418	*
A(2).N(2)	0.6032	0.4932	1.2230	·	A(5).S(2)	2.9250	0.7047	4.1507	*
A(3).N(2)	0.6188	0.5107	1.2117	·	A(5).S(3)	Zero-aliased			
A(4).N(2)	0.8882	0.6363	1.3959	·	B(2).S(2)	0.7183	0.4401	1.6321	.
A(5).N(2)	0.1451	0.5186	0.2798		B(2).S(3)	0.8439	0.2712	3.1117	*
B(2).N(2)	0.4616	0.3597	1.2833	·	N(2).S(2)	Zero-aliased			
A(2).S(2)	0.7794	0.5535	1.4081	·	N(2).S(3)	1.1470	0.5106	2.2464	*
A(2).S(3)	0.5175	0.3353	1.5434	·	B(2).G(2)	0.3049	0.2099	1.4526	.

Key: · - possibly significant (stand. est. >1.0); * - probably significant (stand. est. >1.96).

Table VIII-24 shows the corresponding analysis for HLA-B15 (Liverpool) and Bw62 (Barts-Windsor and BDA studies). Since Bw63, the other 'split' of B15, was not detected in any proband of either of the latter two studies, Bw62 and B15 are here considered equivalent and are termed B15. The terms A.S and N.S make contributions to deviance (X^2_L) identical to those in Table VIII-22, as expected in view of the identity of the data sets so far as A, N and S are concerned.

An interesting result in Table VIII-24 is the significant interaction between HLA-B15 and S, the major difference being between S(3) (Liverpool) and S(1) (BWFS), where the respective B15 antigen frequencies obtained from Table VIII-24-A are 16.8% and 28.8%, with an intermediate figure for the BDA study (22.4%).

Like HLA-B8, B15 shows no significant association with A, N or G here. However, analyses described later (2-B-ii-e) indicate that B.G is significant at the 1% level among pooled probands of the BWFS (with newly-diagnosed cases) and BDA study, the HLA-B15 frequencies being 32% and 19% in male and female probands respectively (Table VIII-33). It is surprising that a term so highly significant in a near-subset of the data in Table VIII-24-A should not be significant when the whole table is analysed, since the interaction between B15 and sex is not significantly affected by source (term B.G.S, models 20 & 21 in Table VIII-24-B). In fact, when the data of Table VIII-24-A are partitioned by source to form two separate tables, B.G is significant at the 5% level in the pooled data for S(1) (BWFS) with S(2) (BDA), where 31% of male and 20% of female probands are B15-positive (Table VIII-25-A). In contrast, among the Liverpool probands the frequency of B15 is the same in males (16%) as in females (18%) (Table VIII-25-B).

The apparent discrepancies between Tables VIII-33 and VIII-25-A in strength of association and level of significance of B.G reflect minor differences in the samples concerned. The former, unlike the latter, includes the thirty newly-diagnosed subjects (6/20 male and 1/10 female being B15-positive) not entering the BWFS proper, but excludes 7

TABLE VIII-25. TYPE 1 DIABETIC PROBANDS CATEGORIZED BY SEX, NUMBER OF SIBLINGS AFFECTED, HLA-B15 STATUS AND TRANCHE OF AGE AT DIAGNOSIS. Log-linear analysis of frequency data, separately according to source (data of Table VIII-24-A). All models shown here include terms for grand mean and main effects.

A. POOLED DATA FROM BARTS-WINDSOR AND BDA STUDIES, S(1) + S(2) of Table VIII-24-A: Interaction effects.

Model ref. no.	Terms included: First-order AB	AG	BG	AN	BN	GN	2nd	3rd	Overall X^2_L	DF	P	Difference: From	X^2_L	DF	P	Term
2	-	-	-	-	-	-	-	-	40.71	32	NS					
4	+	+	+	+	+	+	+	-	4.647	4	NS					
5	+	+	+	+	+	+	-	-	13.23	17	NS	4	8.56	13	NS	2nd
6	+	+	+	+	-	-	-	-	13.26	19	NS	5	0.03	2	NS	BN,GN
7	+	+	-	+	-	-	-	-	17.49	20	NS	6	4.23	1	<.05	BG
8	-	+	-	+	-	-	-	-	23.88	24	NS	7	6.39	4	NS	AB
9	-	-	-	+	-	-	-	-	30.58	28	NS	8	6.70	4	NS	AG
												2	9.87	4	<0.05	AN

B. DATA FROM LIVERPOOL STUDY, S(3) of Table VIII-24-A: Interaction effects.

Model ref. no.	Terms included: First-order AB	AG	BG	AN	BN	GN	2nd	3rd	Overall X2L	DF	P	Difference: From	X2L	DF	P	Term
2	-	-	-	-	-	-	-	-	33.29	32	NS					
4	+	+	+	+	+	+	+	-	8.252	4	NS					
5	+	+	+	+	+	+	-	-	18.59	17	NS	4	10.34	13	NS	2nd
6	+	+	-	+	-	+	-	-	19.34	19	NS	5	0.75	2	NS	BG,BN
7	+	+	-	+	-	-	-	-	20.22	20	NS	6	0.88	1	NS	GN
8	+	+	-	-	-	-	-	-	23.67	24	NS	7	3.45	4	NS	AN
9	+	-	-	-	-	-	-	-	28.85	28	NS	8	5.18	4	NS	AG
												2	4.44	4	NS	AB

multiplex probands uninformative for season at diagnosis (0/4 males and 1/3 females being B15-positive).

Common sense, and inspection of parameter estimates in Table VIII-24-C, suggest that the data structure of Table VIII-24-A may reasonably be simplified by amalgamation of categories A2-A5 and of S1 with S2, without risk of misinterpretation of interactions involving A, B and G. This simplification (Table VIII-26) does not materially affect the conclusion that there is no evidence in these 559 individuals for any interaction between age at diagnosis of diabetes, sex, number of affected siblings and HLA-B15.

The analyses for heterogeneity with respect to the HLA-B8,15 heterozygous state (B), A, G and N are summarized in Table VIII-27. No interaction term involving B is significant. The full data array (Table

TABLE VIII-26. HLA-B15 (B), AGE AT DIAGNOSIS (A), SEX (G), NUMBER OF AFFECTED SIBLINGS (N) AND SOURCE OF DATA (S) in 559 unrelated Type 1 diabetic probands from the Barts-Windsor, augmented BDA multiplex and (HLA-B15) Liverpool studies. Data of Table VIII-24, simplified by amalgamation of multiple categories of A and S.

A. FREQUENCY DATA (numbers of individuals in the respective cross-classification categories):

S Source	1 (BWFS + BDA multiplex)				2 (Liverpool)			
N NDChil	1 (simplex)		2 (multiplex)		1 (simplex)		2 (multiplex)	
G Sex	1 (M)	2 (F)	1 (M)	2 (F)	1 (M)	2 (F)	1 (M)	2 (F)
B HLA-B15	+ -	+ -	+ -	+ -	+ -	+ -	+ -	+ -
A AgeD (yr)								
1 0- 4	9 12	4 13	4 14	5 12	0 5	2 11	1 2	0 3
2 5-40	19 43	9 43	16 37	8 37	17 97	18 78	5 19	3 13

B. LOG-LINEAR ANALYSIS: INTERACTION EFFECTS. All models shown here include terms for grand mean and main effects.

Model ref. no.	Terms included: 1st-order (S S S S A A A B B G / A B G N B G N G N N)	2nd (S S S S S S A A A B / A A A B B G B B G G / B G N G N N G N N N)	3rd (S S S S A / A A A B B / B B G G G / G N N N N)	4th (S / A / B / GN)	Overall X^2_L	DF	p	Difference: Term	Model X^2_L	DF	p
2	- - - - - - - - - -	- - - - - - - - - -	- - - - -	-	114.6	26	<<.001				
8	+ + + + + + + + + +	- - - + - - - - - -	- - - - -	-	11.93	15	NS				
9	+ + + + + + + + + +	- - - - - - - - - -	- - - - -	-	15.16	16	NS	BGS	3.23	1	NS
12	+ + + + + - + + - -	- - - - - - - - - -	- - - - -	-	16.82	19	NS				
13	+ + + + - - + + - -	- - - - - - - - - -	- - - - -	-	17.12	20	NS	AB	0.30	1	NS
18	+ + - + - - - + - -	- - - - - - - - - -	- - - - -	-	17.63	22	NS				
19	+ + - + - - - - - -	- - - - - - - - - -	- - - - -	-	19.50	23	NS	BG	1.87	1	NS
20	+ - - + - - - - - -	- - - - - - - - - -	- - - - -	-	26.53	24	NS	BS	7.03	1	<.01
								AS+NS	78.07	2	<<.001

VIII-27-A) is sparse in that only 53 of the 559 probands are B8,15-positive, so weak interactions cannot be confidently excluded. However, it is noteworthy that only 21 (9.3%) of the 247 probands aged less than 10 years at diagnosis are B8,15-positive - much the same as (and certainly not more than) the 32 (11.4%) of 312 probands aged 10 or more years at diagnosis.

2-B-ii-e. Interactions of thyrogastric antibody status (T) and season at diagnosis (Q) with other factors.

For this analysis the Liverpool sample could not be included because no detailed information was available concerning Q (season) or T (thyrogastric antibody). The BDA Multiplex study and BWFS probands were

TABLE VIII-27. HLA-B8,15 (B), AGE AT DIAGNOSIS (A), SEX (G), NUMBER OF AFFECTED SIBLINGS (N) AND SOURCE OF DATA (S) in 559 unrelated Type 1 diabetic probands from the Barts-Windsor, augmented BDA multiplex and (HLA-B15) Liverpool studies.

A. FREQUENCY DATA (numbers of individuals in the respective cross-classification categories):

S Source	1 (BWFS)				2 (BDA multiplex)				3 (Liverpool)			
N NDChil	1 (simplex)		2 (multiplex)		1 (simplex)		2 (multiplex)		1 (simplex)		2 (multiplex)	
G Sex	1 (M)	2 (F)	1 (M)	2 (F)	1 (M)	2 (F)	1 (M)	2 (F)	1 (M)	2 (F)	1 (M)	2 (F)
B HLA-B8,15	+ -	+ -	+ -	+ -	+ -	+ -	+ -	+ -	+ -	+ -	+ -	+ -
A AgeD (yr)												
1 0- 4	3 18	1 16	0 1	1 2	- -	- -	3 14	2 12	0 5	1 12	0 3	0 3
2 5- 9	3 23	1 22	0 2	1 0	- -	- -	2 23	1 23	1 20	1 20	0 4	0 3
3 10-14	4 19	2 25	*0 *0	*0 *0	- -	- -	2 17	2 14	3 39	5 34	1 8	0 5
4 15-19	2 11	1 1	*0 *0	*1 *0	- -	- -	1 3	0 2	3 20	1 13	1 1	0 3
5 20-40	*0 *0	*0 *0	*0 *0	*0 *0	- -	- -	0 3	0 1	1 27	2 20	0 9	0 5

B. LOG-LINEAR ANALYSIS: INTERACTION EFFECTS. All models shown here include terms for grand mean and main effects. Structural zeroes (indicated in the contingency table by "-") and categories with many random zeroes ("*") were zero-weighted for fitting models.

Model ref. no.	Terms included				Overall			Difference				
	1st-order	2nd	3rd	4th	X^2_L	DF	p	Model	X^2_L	DF	p	Term
	S S S S A A A B B G A B G N B G N G N N	S S S S S S A A A B A A A B B G B B G G B G N G N N G N N N	S S S S A A A A B B B B G G G	S A B G N								
1	+ + + + + + + + + +	+ + + + + + + + + +	- - - - -	-	14.69	15	NS					
2	+ + + + + + + + + +	- - - - - - - - - -	- - - - -	-	34.10	47	NS	1	19.41	32	NS	All 2nd.
8	+ + - + + + + - - -	- - - - - - - - - -	- - - - -	-	35.95	52	NS	2	1.85	5	NS	AB
9	+ - - + + + + - - -	- - - - - - - - - -	- - - - -	-	38.38	54	NS	8	2.43	2	NS	BS
10	+ - - + - + + - - -	- - - - - - - - - -	- - - - -	-	44.48	58	NS	9	6.10	4	NS	AB
11	+ - - + - - + - - -	- - - - - - - - - -	- - - - -	-	51.11	62	NS	10	6.63	4	NS	AG
12	+ - - + - - - - - -	- - - - - - - - - -	- - - - -	-	55.01	66	NS	11	3.90	4	NS	AN

therefore supplemented by the 30 newly-diagnosed cases, giving a total of 315. Of these, 308 were informative for Q, 302 for T and 300 for both Q and T.

To avoid excessive subdivision of data in the multivariate analysis, the source of data (S) is omitted. This is justified as follows. The newly-diagnosed group can be included with the BWFS, being ascertained in the same way and the same geographical area. The BDA probands are all multiplex, while very few of the BWFS probands are multiplex; thus classification by N (number affected in sibship) is, in the context of this analysis, almost exactly equivalent to classification by S.

TABLE VIII-2B. HLA-B8 (B), AGE AT DIAGNOSIS (A; 4 categories), QUARTER AT DIAGNOSIS (Q), SEX (G) AND THYROGASTRIC ANTIBODY STATUS (T) in 300 unrelated Type 1 diabetic probands from the pooled BDA multiplex and Barts-Windsor studies (with newly-diagnosed cases).

A. FREQUENCY DATA (numbers of individuals in the respective cross-classification categories)

G Sex		1 (male)								2 (female)							
A AgeD (y)		1 (0- 4)		2 (5- 9)		3 (10-14)		4 (15-40)		1 (0- 4)		2 (5- 9)		3 (10-14)		4 (15-40)	
B HLA-B8		+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-
Q Qtr	T Th6																
1	-	4	6	3	7	6	10	3	3	2	3	5	4	5	6	2	0
	+	0	0	4	1	4	2	2	0	2	2	4	0	1	2	0	1
2	-	2	2	5	5	7	2	1	5	4	3	5	2	3	1	1	0
	+	3	0	0	1	3	0	1	1	1	2	3	0	5	1	1	0
3	-	9	1	3	2	1	4	2	1	1	1	3	4	3	5	0	0
	+	1	1	1	1	0	1	0	0	3	0	2	2	1	1	1	1
4	-	3	6	8	9	4	7	2	2	2	3	3	10	4	4	0	1
	+	0	1	2	1	1	0	1	0	1	0	2	1	0	2	0	0

B. LOG-LINEAR ANALYSIS: INTERACTION EFFECTS. All models shown here include terms for grand mean and main effects. GLIM zero-aliased 64 terms.

Model ref. no.	Terms included				Overall			Comments
	1st-order	2nd	3rd	4th	X^2_L	DF	p	
	Q Q Q Q A A A B B G A B G T B G T G T T	Q Q Q Q Q Q A A A B A A A B B G B B G G B G T G T T G T T T	Q Q Q Q A A A A B B B B G G G G T T T T	Q A B GT				
2	- - - - - - - - - -	- - - - - - - - - -	- - - - -	-	181.7	119	<0.001	
3	+ + + + + + + + + +	+ + + + + + + + + +	+ + + + +	+	114.4	64		Zero-aliased: all in A4 and all in A.T
4	+ + + + + + + + + +	+ + + + + + + + + +	+ + + + +	-	114.4	64		" "
5	+ + + + + + + + + +	+ + + + + + + + + +	- - - - -	-	131.6	73		" "
6	+ + + + + + + + + +	- - - - - - - - - -	- - - - -	-	145.3	97		" "

C. LOG-LINEAR ANALYSIS AFTER AMALGATION OF CATEGORIES 3 AND 4 OF VARIABLE A (ageD): INTERACTION EFFECTS. All models shown here include terms for grand mean and main effects.

Model ref. no.	Terms included				Overall			Difference; Comments				
	1st-order	2nd	3rd	4th	X^2_L	DF	p	Model	X^2_L	DF	P	Term
	Q Q Q Q A A A B B T A B T G B T G T G G	Q Q Q Q Q Q A A A B A A A B B T B B T T B T G T G G T G G G	Q Q Q Q A A A A B B B B T T T T G G G G	Q A B TG								
2	- - - - - - - - - -	- - - - - - - - - -	- - - - -	-	107.4	87	NS	(critical value = 110)				
5	+ + + + + + + + + +	+ + + + + + + + + +	- - - - -	-	40.71	29	NS	Sat.	26.37	23	NS	
6	+ + + + + + + + + +	+ - + - - - - - - -	- - - - -	-	52.07	51	NS	5	11.36	22	NS	
7	+ + + + + + + + + +	+ - - - - - - - - -	- - - - -	-	58.59	57	NS	6	6.52	6	NS	QAG
8	+ + + + + + + + + +	- - - - - - - - - -	- - - - -	-	68.29	63	NS	7	9.70	6	NS	QAB
9	- + - - - - - + - +	- - - - - - - - - -	- - - - -	-	84.44	82	NS	8	16.15	19	NS	QT etc
10	- - - - - - - + - +	- - - - - - - - - -	- - - - -	-	92.87	85	NS	9	8.43	3	<0.05	QB
11	- - - - - - - + - -	- - - - - - - - - -	- - - - -	-	98.68	86	NS	10	5.81	13	<0.025	TG
								2	8.72	1	<0.01	BT
13	+ + + + + - + + + +	+ - + + + + - + - +	- + - + -	-	35.58	32	NS	Sat.	35.58	32	NS	In AT

Since very few of the 300 probands were aged >19 years at diagnosis, the number of categories of A was reduced from five to four, the fourth including all aged 15 or more at diagnosis.

Tables VIII-28-A and -B show the 5-dimensional contingency table and a preliminary analysis for interactions between A, B(HLA-B8), G (sex), Q and T. Variable N is omitted here (see below). GLIM found the category A4 and the interaction A.T to be uninformative and so zero-aliased all related terms. The first step in simplification of this sparse data array was therefore to reduce further the number of categories of A by amalgamating A3 and A4.

Log-linear analysis then gave the results shown in Table VIII-28-C. No 2nd- or higher-order term is significant here. Thyrogastric antibody status is significantly associated with sex (models 11 & 10): of 169 male and 133 female probands informative for T, 34 (20%) and 42 (32%) respectively are positive. Thyrogastric antibody status is also significantly associated with HLA-B8 status (models 11 & 2). Overall, 51 (32%) of the 158 B8-positive probands are positive for T, compared with only 25 (17%) of the 144 B8-negative probands.

Interestingly, there is no evidence here for any interaction between T and either of the variables directly relevant to diabetes, i.e. A or Q (Models 8 - 9; Model 13). Furthermore, the standardized parameter estimate for the interaction A.B.T in Model 5 is only 0.995, indicating that this term is not significant. Thus 'Type 1B diabetes' does not form a significant subgroup of these probands (see Discussion, 2-C).

The interaction B.Q is significant at the 5% level (Models 9-10), meriting further examination after cautious simplification of the data structure. This was first attempted by collapsing Table VIII-28-A over T (with A in three categories, as above). In the resulting analysis by four factors (A(3 categories), B(2), G(2), Q(4)), GLIM zero-aliased all terms in A3 (details not shown). Therefore in subsequent analyses, A is considered in two categories only: A1 (age at diagnosis 0-4 years) and A2 (>4 years).

(Continued on P. 255)

TABLE VIII-29. HLA-B8 (B), AGE AT DIAGNOSIS (A; 2 categories), QUARTER AT DIAGNOSIS (Q), SEX (G) AND NUMBER OF AFFECTED SIBLINGS (N) in 308 unrelated Type 1 diabetic probands from the pooled BDA multiplex and Barts-Windsor studies (with newly-diagnosed cases).

A. FREQUENCY DATA (numbers of individuals in the respective cross-classification categories):

N	NDChil	1 (simplex)				2 (multiplex)			
G	Sex	1 (male)		1 (female)		1 (male)		1 (female)	
A	AgeD (y)	1 (0- 4)	2 (>4)	1 (0- 4)	2 (>4)	1 (0- 4)	2 (>4)	1 (0- 4)	2 (>4)
B	HLA-B8	+ -	+ -	+ -	+ -	+ -	+ -	+ -	+ -
Q	Qtr								
	1	2 2	19 17	0 2	10 8	2 4	3 7	4 3	7 5
	2	2 1	11 7	4 3	12 3	3 1	6 7	2 3	7 1
	3	5 2	3 4	2 1	3 9	5 0	4 5	3 0	7 5
	4	2 5	11 9	2 3	5 12	1 2	7 11	1 1	4 6

B. LOG-LINEAR ANALYSIS: INTERACTION EFFECTS. Terms for grand mean and main effects are included.

Model ref. no.	Terms included				Overall			Comments			
	1st-order	2nd	3rd	4th	X^2_L	DF	p	Model	X^2_L	DF	p	Term
	Q Q Q Q A A A B B N A B N G B N G N G G	Q Q Q Q Q Q A A A B A A A B B N B B N N B N G N G G N G G G	Q Q Q Q A A A A B B B B N N N N G G G G	Q A B NG							
2	- - - - - - - - - -	- - - - - - - - - -	- - - - -	-	185.6	63	<<0.001				
4	+ + + + + + + + + +	+ + + + + + + + + +	+ + + + +	-	2.298	3	NS				
5	+ + + + + + + + + +	+ + + + + + + + + +	- - - - -	-	10.54	16	NS	4	8.24	13	NS	3rd
6	+ + + + + + + + + +	- - - - - - - - - -	- - - - -	-	53.91	38	<0.05	5	43.37	22	<0.01	2nd
7	+ + + + + + + + + +	+ + - - - + + - - +	- - - - -	-	24.44	27	NS	5	13.90	11	NS	
14	+ + + + + + + + + +	+ + - - - + - - - +	- - - - -	-	25.46	28	NS	7	1.02	NS	NS	ABN
15	+ + + + + + + + + +	+ + - - - - - - - +	- - - - -	-	30.68	31	NS	14	5.22	3	NS	QNG
20	+ + + + + + + + + +	+ + - - - - - - - -	- - - - -	-	33.53	32	NS	15	2.85	1	NS	BNG
21	+ + + + + + + + + +	- + - - - - - - - -	- - - - -	-	43.63	35	NS	20	10.10	3	<0.025	QAB
								6	10.28	3	<0.025	QAN

C. LOG-LINEAR MODEL 5 (in Table B): PARAMETER ESTIMATES FOR SECOND-ORDER INTERACTIONS

Parameter	Estimate	Standard error	Standardized estimate: Individual		Mean (a)		Parameter	Estimate	Standard error	Standardized estimate: Individual		Mean (a)	
B2.Q2.G2	-0.3241	0.6816	-0.47550				B2.Q2.N2	0.0274	0.6988	0.03914			
B2.Q3.G2	0.5044	0.7450	0.67705				B2.Q3.N2	-1.363	0.7814	-1.74431	.		
B2.Q4.G2	0.7376	0.6365	1.15884	.	0.77046		B2.Q4.N2	-0.2070	0.6639	-0.31179		0.69841	
B2.Q2.A2	-0.0697	0.8032	-0.08677				B2.G2.N2	-0.8410	0.4947	-1.70002	.	1.70002	.
B2.Q3.A2	2.585	0.9800	2.63776	*			Q2.G2.N2	-1.400	0.6950	-2.01439	*		
B2.Q4.A2	0.1658	0.8202	0.20215		0.97556		Q3.G2.N2	-1.028	0.7366	-1.39560	.		
B2.G2.A2	-0.4220	0.5991	-0.70439		0.70439		Q4.G2.N2	-1.193	0.6636	-1.79777	.	1.73592	.
Q2.G2.A2	-1.023	0.8104	-1.26234	.			B2.A2.N2	0.6642	0.6227	1.06665	.	1.06665	.
Q3.G2.A2	0.9123	0.8832	1.03295	.			Q2.A2.N2	1.385	0.8121	1.70546	.		
Q4.G2.A2	0.1182	0.8217	0.14385		0.81305		Q3.A2.N2	2.714	0.9270	2.92772	*		
							Q4.A2.N2	2.239	0.8227	2.72153	*	2.45157	*
							G2.A2.N2	-0.0980	0.1591	-0.16573		0.16573	

NOTES (a) Mean of absolute values of standardized estimates for all parameters involving a given combination of factors. KEY: . = 1.0 < standardized estimate < 1.96; * = standardized estimate > 1.96

TABLE VIII-30. HLA-B8 (B), QUARTER AT DIAGNOSIS (Q), SEX (G) AND NUMBER OF AFFECTED SIBLINGS (N): Log-linear analysis, separately for 73 probands aged <5 years and for 235 probands aged 5 or more at diagnosis. The data analysed are shown in Table VIII-29-A, categories A1 and A2 respectively.

A. Age at diagnosis <5 years (category A1 of Table VIII-29-A).

Model ref. no.	Terms included: 1st-order (Q B, Q N, Q G, B N, B G, N G)	2nd (Q B, Q N, Q G, B N, B G, N G)	3rd (Q, B, N, G)	Overall X^2_L	DF	p	Difference: Model	X^2_L	DF	p	Term
2	- - - - - -	- - - -	-	29.90	25	NS					
4	+ + + + + +	+ + + +	-	2.77	3	NS					
5	+ + + + + +	- - - -	-	8.961	13	NS	4	6.19	10	NS	2nd
6	+ + - + - -	- - - -	-	12.64	18	NS	5	3.68	5	NS	
7	+ + - - - -	- - - -	-	13.64	19	NS	6	1.00	1	NS	BN
8	+ - - - - -	- - - -	-	19.40	22	NS	7	5.76	3	NS	QN
							2	10.50	3	<0.025	QB

B. Age at diagnosis 5 or more years (category A2).

Model ref. no.	Terms included: 1st-order (Q B, Q N, Q G, B N, B G, N G)	2nd (Q B, Q N, Q G, B N, B G, N G)	3rd (Q, B, N, G)	Overall X^2_L	DF	p	Difference: Model	X^2_L	DF	p	Term
2	- - - - - -	- - - -	-	38.07	25	<0.05					
4	+ + + + + +	+ + + +	-	0.032	3	NS					
5	+ + + + + +	- - - -	-	16.76	13	NS	4	16.73	10	NS	2nd
8	+ + + + + +	- - - +	-	13.02	12	NS	5	3.74	1	NS	BGN
9	+ + - - - -	- - - -	-	22.37	19	NS	5	5.61	6	NS	
10	- + - - - -	- - - -	-	31.28	22	NS	9	8.91	3	<0.05	QB
							2	6.79	3	NS	QN

C. Frequency data: Q versus HLA-B8 in the two categories of A. From Table VIII-29-A, collapsed over N and G.

A AgeD		1 (<5 y)		2 (>=5 y)	
B HLA-B8		1	2	1	2
Q Quarter	1	8	11	39	37
	2	11	8	36	18
	3	15	3	17	23
	4	6	11	27	38

Table VIII-29-A (previous page) shows the 308 probands informative for Q, categorized by Q, A, B(HLA-B8), G and N. Log-linear analysis now indicates that the second-order interactions Q.A.N and Q.A.B are significant at the 2.5% level (Table VIII-29-B, Models 15, 20, 21, 6; analysis guided by parameter estimates for Model 5, shown in Table C). Varying the order of elimination of terms makes no difference to the results (models not shown).

Model 6 of Table VIII-29-B, with all first-order but no second-order interactions, does not give an acceptable fit. Therefore the interaction Q.B should be examined separately in the two contingency tables derived from Table VIII-29-A by partitioning according to the two categories of A. In the 73 probands aged <5 years at diagnosis, Q.B is significant at

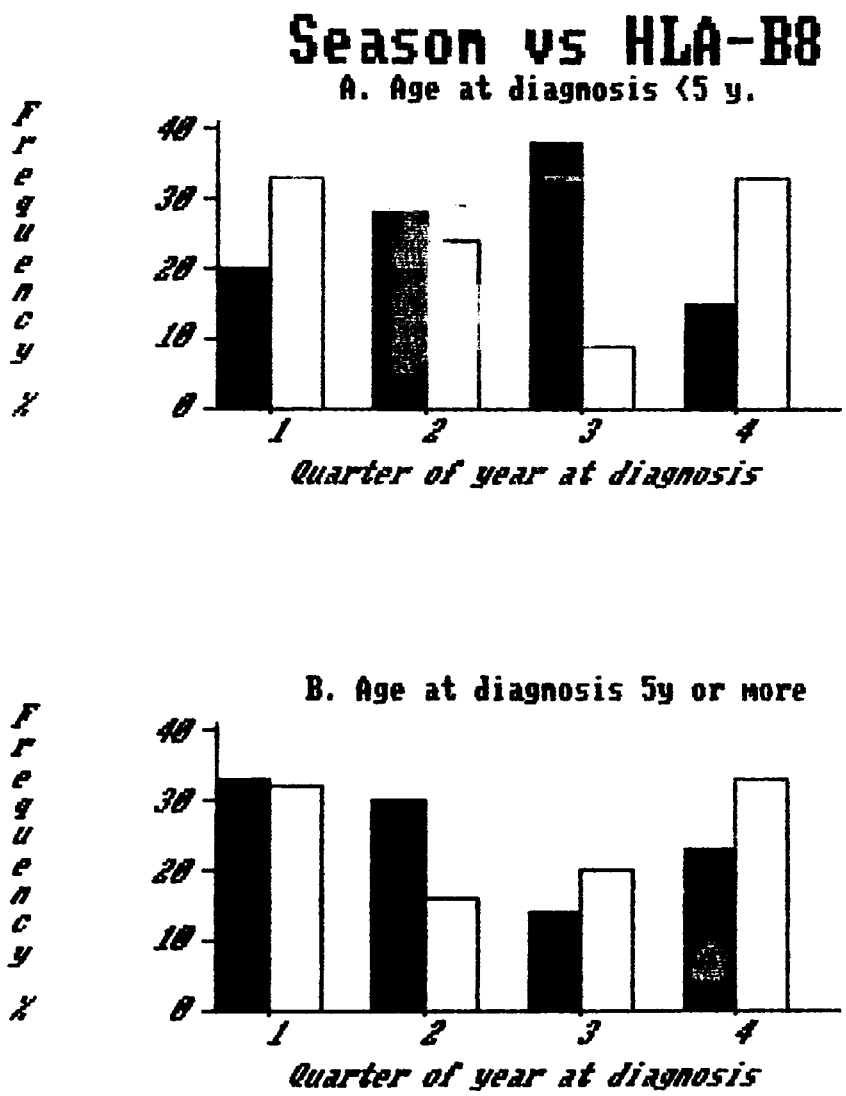


FIGURE VIII-3. SEASON AT DIAGNOSIS OF DIABETES VERSUS HLA-B8 AND AGE AT DIAGNOSIS.
A - Age at diagnosis <5 years (n=73). B - Age at diagnosis >5 years (n=235)
Solid bars = HLA-B8 positive, open bars = B8 negative.

The bars represent the relative proportions (%) of diagnoses in the four quarters of the year within each group of probands. Data from Table VIII-30-C.

the 2.5% level (Table VIII-30-A), while in the much larger number (235) of probands aged 5 or more years at diagnosis Q.B attains only 5% significance (Table VIII-30-B). No term in G is significant, nor is B.N; hence the interaction Q.B may be examined further within categories of A after collapsing over G and N.

Figure VIII-3 illustrates graphically the effects of HLA-B8 status and age on seasonal variation in diagnosis of diabetes. The very young, B8-positive children, in contrast to the other groups, show a marked excess of cases diagnosed in late summer, while the B8-negative young probands show a contrary trend so that, if HLA status is ignored, no seasonal variation is detectable in pre-school children. In older children who are B8-positive the seasonal minimum in diagnosis rate occurs in the second quarter of the year, whereas in those who are B8-negative it is in the third quarter.

Even though Q.A.N is significant in Table VIII-29, after partitioning by A the interaction Q.N fails to reach the 5% level of significance in either the younger or the older children (Table VIII-30). This is presumably because the total number in the A1 category is relatively small. However, after partitioning by N, Q.A is found to be significant in the 181 simplex probands, only those diagnosed aged 5 or more showing the well-recognized seasonal variation (Tables VIII-31-A and -C; Figure VIII-4-A). In contrast, there is no sign of any such variation in the 127 multiplex children (Table VIII-31-B and -C; Figure VIII-4-B).

The final analysis performed with respect to HLA-B8 included the five variables Q, A (2 categories), B8 status B, N and thyrogastric antibody status T. The aim was to assess whether T might be involved in a complex interaction with Q.A.B or Q.A.N. The results (not shown) gave no evidence for any such complex interaction, only B.T, Q.A.N, Q.A.B and simpler, related terms being significant, as before.

TABLE VIII-31. HLA-B8 (B), QUARTER AT DIAGNOSIS (Q), SEX (G) AND AGE AT DIAGNOSIS (A): Log-linear analysis, separately for 181 simplex and 127 multiplex probands. The data analysed are shown in Table VIII-29-A, categories N1 and N2 respectively.

A. Simplex (category N1 of Table VIII-29-A).

Model ref. no.	Terms included: 1st-order (Q A, Q B, Q G, A B, A G, B G)	2nd (Q A, Q B, Q G, A B, A G, B G)	3rd (Q, A, B, G)	Overall X^2_L	DF	p	Difference: Model	X^2_L	DF	p	Term
2	- - - - - -	- - - -	-	37.58	25	<0.05					
4	+ + + + + +	+ + + +	-	4.10	3	NS					
5	+ + + + + +	- - - -	-	16.57	13	NS	4	12.47	10	NS	2nd
6	+ + + + + +	+ - - -	-	11.00	10	NS	5	4.10	3	NS	BQA
7	+ + + - - -	- - - -	-	17.44	16	NS	5	0.87	6	NS	
8	+ - + - - -	- - - -	-	24.69	19	NS	7	7.25	3	NS	BQ
9	+ - - - - -	- - - -	-	29.12	22	NS	8	4.43	3	NS	QG
							2	8.46	3	<0.05	QA

B. Multiplex (category N2).

Model ref. no.	Terms included: 1st-order (Q A, Q B, Q G, A B, A G, B G)	2nd (Q A, Q B, Q G, A B, A G, B G)	3rd (Q, A, B, G)	Overall X^2_L	DF	p	Difference: Model	X^2_L	DF	p	Term
2	- - - - - -	- - - -	-	31.03	25	<0.05					
4	+ + + + + +	+ + + +	-	3.04	3	NS					
5	+ + + + + +	- - - -	-	15.11	13	NS	4	12.07	10	NS	2nd
6	+ + - - - +	- - - -	-	18.56	18	NS	5	3.34	5	NS	
7	+ - - - - +	- - - -	-	24.16	21	NS	6	5.60	3	NS	BQ
8	- - - - - +	- - - -	-	28.64	24	NS	7	4.48	3	NS	QA
							2	2.39	1	NS	BG

C. Frequency data: Q versus A in the two categories of N. From Table VIII-29-A, collapsed over B and G.

N NDChil		1 simplex		2 multiplex	
A AgeD		1	2	1	2
Q Quarter	1	6	54	13	22
	2	10	33	9	21
	3	10	19	8	21
	4	12	37	5	28

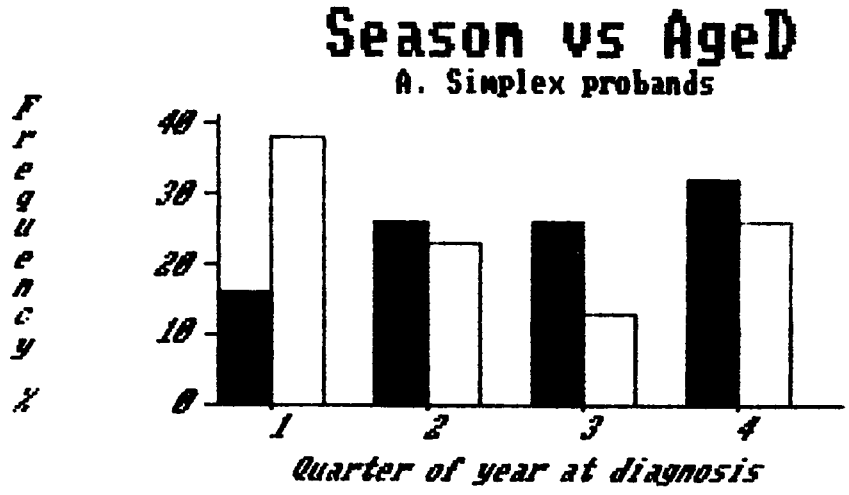


FIGURE VIII-4. SEASON AT DIAGNOSIS OF DIABETES VERSUS AGE AT DIAGNOSIS AND NUMBER OF AFFECTED SIBLINGS.
A - probands from simplex sibships (n=181). B - probands (first-diagnosed diabetic child) from multiplex sibships (n=127) Solid bars = age at diagnosis <5 years, open bars = age >= 5 y.

The bars represent the relative proportions (%) of diagnoses in the four quarters of the year within each group of probands. Data from Table VIII-31-C.

Heterogeneity involving HLA-B15 (B) and the other variables considered above, including Q and T, was investigated next. Table VIII-32 summarizes the analysis including Q, T, A (3 categories), B and G (sex). The association T.G between thyrogastric antibody T and female sex is confirmed (see above; Table VIII-28), while that between HLA-B15 and male sex (B.G) is particularly marked in the sample of diabetic probands studied here (see 2-B-ii-d above, and Table VIII-25). However,

TABLE VIII-32. HLA-B15 (B), AGE AT DIAGNOSIS (A; 3 categories), QUARTER AT DIAGNOSIS (Q), SEX (G) AND THYROGASTRIC ANTIBODY STATUS (T) in 300 unrelated Type 1 diabetic probands from the pooled BDA multiplex and Barts-Windsor studies (with newly-diagnosed cases).

A. FREQUENCY DATA (numbers of individuals in the respective cross-classification categories):

G Sex		1 (male)						2 (female)					
A AgeD (y)		1 (0- 4)		2 (5- 9)		3 (10-40)		1 (0- 4)		2 (5- 9)		3 (10-40)	
B HLA-B15		+	-	+	-	+	-	+	-	+	-	+	-
Q Qtr	T Th6												
1	-	2	8	3	7	5	17	0	5	4	5	3	10
	+	0	0	1	4	5	3	2	2	0	4	0	4
2	-	2	2	4	6	3	12	3	4	0	7	1	4
	+	1	2	1	0	3	2	1	2	0	3	0	7
3	-	3	7	2	3	2	6	1	1	0	7	4	4
	+	0	2	0	2	0	1	0	3	1	3	2	2
4	-	5	4	5	12	5	10	1	4	1	12	1	8
	+	0	1	1	2	1	1	0	1	0	3	0	2

B. LOG-LINEAR ANALYSIS: INTERACTION EFFECTS. Terms for grand mean and main effects are included.

Model ref. no.	Terms included				Overall			Difference				
	1st-order	2nd	3rd	4th	X^2_L	DF	p	Model X^2_L	DF	p	Term	
	Q Q Q Q A A A B B G A B G T B G T G T T	Q Q Q Q Q Q A A A B A A A B B G B B G G	Q Q Q Q A A A A B B	Q A								
		B G T G T T G T T T	B B G G G G T T T T	B GT								
2	- - - - - - - - - -	- - - - - - - - - -	- - - - -	-	106.3	87	NS					
4	+ + + + + + + + + +	+ + + + + + + + + +	+ + + + +	-	No convergence							
5	+ + + + + + + + + +	+ + + + + + + + + +	- - - - -	-	41.45	29	NS	Sat. 41.45	29	NS	3rd+4th	
8	+ + + + + + + + + +	- - - - - - - - - -	- - - - -	-	74.72	63	NS	7 5.34	6	NS	QAB	
9	+ + + + + + + + + +	+ - - + - - - - - +	- - - - -	-	60.57	53	NS	5 19.12	24	NS	2nd	
10	+ + + + + + + + + +	- - - + - - - - - +	- - - - -	-	66.16	59	NS	9 5.59	6	NS	QAB	
11	+ + + + + + + + + +	- - - - - - - - - +	- - - - -	-	72.86	62	NS	10 6.70	3	NS	QBG	
								8 1.86	1	NS	BTG	
12	- - - - - - - + - +	- - - - - - - - - -	- - - - -	-	93.68	85	NS	8 18.96	21	NS		
13	- - - - - - - + - -	- - - - - - - - - -	- - - - -	-	99.49	86	NS	12 5.81	1	<0.025	GT	
								2 6.8	1	<0.01	BG	

C. LOG-LINEAR MODEL 8 (in Table B): PARAMETER ESTIMATES FOR FIRST-ORDER INTERACTIONS

Parameter	Estimate	Standard error	Standardized estimate Individual	Mean (a)	Parameter	Estimate	Standard error	Standardized estimate Individual	Mean (a)
Q(2).A(2)	-0.1846	0.4430	-0.41670		Q(2).T(2)	0.2154	0.3516	0.61263	
Q(2).A(3)	-0.2711	0.4057	-0.66823		Q(3).T(2)	0.0542	0.3832	0.14157	
Q(3).A(2)	-0.3469	0.4524	-0.76680		Q(4).T(2)	-0.7235	0.3966	-1.82426 .	0.85948
Q(3).A(3)	-0.6860	0.4256	-1.61184 .		A(2).T(2)	0.0387	0.3720	0.10401	
Q(4).A(2)	0.4271	0.4259	1.00282 .		A(3).T(2)	0.0786	0.3533	0.22256	0.16328
Q(4).A(3)	-0.3404	0.4170	-0.81631	0.88045	B(2).T(2)	-0.0261	0.3156	-0.08276	0.08276
Q(2).B(2)	-0.0498	0.3614	-0.13774		Q(2).G(2)	0.1467	0.3265	0.44931	
Q(3).B(2)	-0.0555	0.3893	-0.14249		Q(3).G(2)	0.3289	0.3493	0.94160	
Q(4).B(2)	0.0529	0.3572	0.14812	0.14278	Q(4).G(2)	0.0076	0.3200	0.02372	0.47154
A(2).B(2)	0.3816	0.3589	1.06325 .		A(2).G(2)	0.1885	0.3227	0.58413	
A(3).B(2)	0.1705	0.3339	0.51063	0.78694	A(3).G(2)	-0.1247	0.3110	-0.40096	0.49254
					B(2).G(2)	0.7043	0.2811	2.50551 *	2.50551 *
					-T(2).G(2)	0.6403	0.2767	- 2.31406 *	2.31406 *

NOTES (a) Mean of absolute values of standardized estimates for all parameters involving a given combination of factors. KEY . = 1.0 < standardized estimate < 1.96. * = standardized estimate > 1.96

TABLE VIII-33. HLA-B15 (B), AGE AT DIAGNOSIS (A; 2 categories), QUARTER AT DIAGNOSIS (Q), SEX (G) AND NUMBER OF AFFECTED SIBLINGS (N) in 308 unrelated Type 1 diabetic probands from the pooled BDA multiplex and Barts-Windsor studies (with newly-diagnosed cases).

A. FREQUENCY DATA (numbers of individuals in the respective cross-classification categories):

N NDChil	1 (simplex)				2 (multiplex)			
G Sex	1 (male)		1 (female)		1 (male)		1 (female)	
A AgeD (y)	1 (0- 4)	2 (>4)	1 (0- 4)	2 (>4)	1 (0- 4)	2 (>4)	1 (0- 4)	2 (>4)
B HLA-B15	+ -	+ -	+ -	+ -	+ -	+ -	+ -	+ -
Q Qtr								
1	0 4	13 23	0 2	4 14	2 4	1 9	2 5	3 9
2	1 2	6 12	2 5	0 15	2 2	5 8	2 3	1 7
3	3 4	1 6	1 2	5 7	0 5	3 6	0 3	2 10
4	5 2	5 15	1 4	1 16	0 3	7 11	1 1	1 9

B. LOG-LINEAR ANALYSIS: INTERACTION EFFECTS. Terms for grand mean and main effects are included.

Model ref. no.	Terms included				Overall			Comments				
	1st-order	2nd	3rd	4th	X^2_L	DF	p	Model	X^2_L	DF	p	Term
	Q Q Q Q A A A B B G A B G N B G N G N N	Q Q Q Q Q Q A A A B A A A B B G B B G G B G N G N N G N N N	Q Q Q Q A A A A B B B B G G G G N N N N	Q A B GN								
4	+ + + + + + + + + +	+ + + + + + + + + +	- - - - -	-	22.32	16	NS					
8	+ + + + + + + + + +	- + + + - - - - - -	- - - - -	-	37.49	29	NS	4	15.17	13	NS	
9	+ + + + + + + + + +	- + + - - - - - - -	- - - - -	-	42.87	32	NS	8	5.38	3	NS	QBG
10	+ + + + + + + + + +	- - + - - - - - - -	- - - - -	-	49.09	35	NS	9	6.22	3	NS	QGA
11	+ + + + + + + + + +	- - - - - - - - - -	- - - - -	-	59.38	38	<0.025	10	10.19	3	<0.025	QAN
13	+ - - - - - + + - -	- - - - - - - - - -	- - - - -	-	65.19	51	NS	11	5.81	13	NS	
14	- - - - - - + + - -	- - - - - - - - - -	- - - - -	-	68.18	54	NS	13	2.99	3	NS	QA
15	- - - - - - - + - -	- - - - - - - - - -	- - - - -	-	69.95	55	NS	14	1.77	1	NS	AN
16	- - - - - - - - - -	- - - - - - - - - -	- - - - -	-	76.70	56	<0.05	15	6.75	1	<0.01	BG

no interaction involving HLA-B15 and Q or T comes to light here, nor any directly relevant to diabetes. Similarly, the analysis described in Table VIII-33 gives no indication of interactions between HLA-B15, Q, A, G and N, except for those already mentioned.

Among the probands informative for Q and/or T, only 33 are HLA-B8,15 heterozygous. Multivariate analysis for heterogeneity involving these variables is therefore of limited value. Results are shown in Tables VIII-34 and VIII-35. As already shown, T is significantly associated with female sex (T.G significant at 2.5% level). T is also associated with HLA-B8,15, which is not surprising in view of the association with B8 already described. No new interaction relevant to diabetes itself is demonstrated here.

TABLE VIII-34. HLA-B8,15 (B), AGE AT DIAGNOSIS (A; 3 categories), QUARTER AT DIAGNOSIS (Q), SEX (G) AND THYROGASTRIC ANTIBODY STATUS (T) in 300 unrelated Type 1 diabetic probands from the pooled BDA multiplex and Barts-Windsor studies (with newly-diagnosed cases).

A. FREQUENCY DATA (numbers of individuals in the respective cross-classification categories):

G Sex		1 (male)						1 (female)					
A AgeD (y)		1 (0- 4)		2 (5- 9)		3 (10-40)		1 (0- 4)		2 (5- 9)		3 (10-40)	
B HLA-B8,15		+	-	+	-	+	-	+	-	+	-	+	-
Q Qtr	T ThG												
1	-	1	9	1	9	1	21	0	5	1	8	2	11
	+	0	0	1	4	4	4	2	2	0	4	0	4
2	-	1	3	3	7	1	14	1	6	0	7	1	4
	+	1	2	0	1	3	2	0	3	0	3	0	7
3	-	2	8	1	4	1	7	0	2	0	7	1	7
	+	0	2	0	2	0	1	0	3	1	3	1	3
4	-	1	8	0	17	1	14	1	4	0	13	1	8
	+	0	1	0	3	1	1	0	1	0	3	0	2

B. LOG-LINEAR ANALYSIS: INTERACTION EFFECTS. Terms for grand mean and main effects are included.

Model ref. no.	Terms included: 1st-order	2nd	3rd	4th	Overall X^2_L	DF	p	Difference: Model X^2_L	DF	p	Term
	Q Q Q Q A A A B B G A B G T B G T G T T	Q Q Q Q Q Q A A A B A A A B B G B B G G B G T G T T G T T T	Q Q Q Q A A A A B B B B G G G G T T T T	Q A B GT							
1	+ + + + + + + + + +	+ + + + + + + + + +	- - - - -	-	32.70	29	NS				
2	+ + + + + + + + + +	+ + - - - - - - - -	- - - - -	-	49.94	51	NS				
3	+ + + + + + + + + +	- + - - - - - - - -	- - - - -	-	56.34	57	NS	2	6.40	6	NS	Most 2e
4	+ + + + + + + + + +	- - - - - - - - - -	- - - - -	-	61.03	63	NS	3	4.69	6	NS	
5	- + - - - - - - + +	- - - - - - - - - -	- - - - -	-	82.54	82	NS					
6	- - - - - - - - + +	- - - - - - - - - -	- - - - -	-	86.65	85	NS	5	4.11	3	NS	QB
7	- - - - - - - - + -	- - - - - - - - - -	- - - - -	-	92.47	86	NS	6	5.82	1	<0.025	GT
8	- - - - - - - - - -	- - - - - - - - - -	- - - - -	-	96.34	87	NS	7	3.87	1	<0.05	BT

C. LOG-LINEAR MODEL 4 (in Table B): PARAMETER ESTIMATES FOR FIRST-ORDER INTERACTIONS

Parameter	Estimate	Standard error	Standardized estimate: Individual	Mean (a)
Q(2).A(2)	-0.1786	0.4434	-0.40280	
Q(2).A(3)	-0.2714	0.4056	-0.66913	
Q(3).A(2)	-0.3563	0.4527	-0.78706	
Q(3).A(3)	-0.6887	0.4255	-1.61857 .	
Q(4).A(2)	0.3963	0.4269	0.92832	
Q(4).A(3)	-0.3439	0.4180	-0.82273	0.87144
Q(2).B(2)	-0.1376	0.4520	-0.30442	
Q(3).B(2)	0.0822	0.5133	0.16012	
Q(4).B(2)	0.7230	0.5594	1.29246 .	0.58567
A(2).B(2)	0.6147	0.5115	1.20176 .	
A(3).B(2)	0.0792	0.4359	0.18178	0.69177

Parameter	Estimate	Standard error	Standardized estimate: Individual	Mean (a)
Q(2).T(2)	0.2019	0.3545	0.56953	
Q(3).T(2)	0.0624	0.3861	0.16151	
Q(4).T(2)	-0.6721	0.3994	-1.68277 .	0.80460
A(2).T(2)	0.0870	0.3750	0.23213	
A(3).T(2)	0.0855	0.3560	0.24003	0.23608
B(2).T(2)	-0.7693	0.3858	-1.99404 *	1.99404 *
Q(2).G(2)	0.1505	0.3247	0.46350	
Q(3).G(2)	0.3160	0.3470	0.91066	
Q(4).G(2)	-0.0199	0.3184	-0.06250	0.47889
A(2).G(2)	0.2041	0.3206	0.63662	
A(3).G(2)	-0.1064	0.3088	-0.34456	0.49059
B(2).G(2)	0.6223	0.3875	1.60594 .	1.60594 .
T(2).G(2)	0.6923	0.2778	2.49208 *	2.49208 *

NOTES: (a) Mean of absolute values of standardized estimates for all parameters involving a given combination of factors. KEY: . = 1.0 < standardized estimate < 1.96. * = standardized estimate > 1.96.

TABLE VIII-35. HLA-B8,15 (B), AGE AT DIAGNOSIS (A; 2 categories), QUARTER AT DIAGNOSIS (Q), SEX (G) AND NUMBER OF AFFECTED SIBLINGS (N) in 308 unrelated Type 1 diabetic probands from the pooled BDA multiplex and Barts-Windsor studies (with newly-diagnosed cases).

A. FREQUENCY DATA (numbers of individuals in the respective cross-classification categories):

N	NDChil	1 (simplex)				2 (multiplex)			
G	Sex	1 (male)		1 (female)		1 (male)		1 (female)	
A	AgeD (y)	1 (0- 4)	2 (>4)	1 (0- 4)	2 (>4)	1 (0- 4)	2 (>4)	1 (0- 4)	2 (>4)
B	HLA-B8,15	+ -	+ -	+ -	+ -	+ -	+ -	+ -	+ -
Q	Qtr								
	1	0 4	6 30	0 2	2 16	1 5	1 9	2 5	1 11
	2	0 3	5 13	0 7	0 15	2 2	2 11	1 4	1 7
	3	2 5	1 6	0 3	2 10	0 5	1 8	0 3	1 11
	4	1 6	1 19	1 4	0 17	0 3	1 17	0 2	1 9

B. LOG-LINEAR ANALYSIS: INTERACTION EFFECTS. Terms for grand mean and main effects are included.

Model ref. no.	Terms included: 1st-order	2nd	3rd	4th	Overall X^2_L	DF	p	Comments: Model X^2_L	DF	p	Term
	Q Q Q Q A A A B B G / A B G N B G N G N N	Q Q Q Q Q Q A A A B / A A A B B G B B G G / B G N G N N G N N N	Q Q Q Q A / A A A B B / B B G G G / G N N N N	Q / A / B / GN							
2	- - - - - - - - - -	- - - - - - - - - -	- - - - -	-	61.26	56	NS				
5	+ + + + + + + + + +	+ + + + + + + + + +	- - - - -	-	14.21	16	NS				
6	+ + + + + + + + + +	- - - - - - - - - -	- - - - -	-	44.92	38	NS				
11	+ + + + + + + + + +	- - + - - - - - - -	- - - - -	-	34.63	35	NS	6	10.29	3	<0.025 QAN
13	- + - - - - + + - -	- - - - - - - - - -	- - - - -	-	53.23	51	NS				
14	- - - - - - + + - -	- - - - - - - - - -	- - - - -	-	57.28	54	NS	13	4.05	3	NS QB
15	- - - - - - + - - -	- - - - - - - - - -	- - - - -	-	59.50	55	NS	14	2.22	1	NS BG
								2	1.76	1	NS AN

2-B-iii. Further analyses relevant to mode of inheritance

2-B-iii-a. Hardy-Weinberg equilibrium in diabetic probands

Table VIII-36-A gives results of tests for Hardy-Weinberg equilibrium of those HLA-A, -B and -C alleles which were sufficiently frequent (7 or more occurrences) for analysis in unrelated probands of the Barts-Windsor Family Study. HLA-Cw6-Cw8 are not included because typing was incomplete (see 1-B-i-a). The method of computation has been described (VI-3-C-iii).

In most cases the observed distribution of genotypes closely matches that predicted by the Hardy-Weinberg equation. For HLA-B8, however, there is an excess of heterozygotes which appears to be significant at

the 5% level. This requires further investigation. In the 125 probands (first-diagnosed diabetic children) in the augmented BDA Multiplex Family Study, the gene frequency of HLA-B8 by gene-counting is 0.2840. Assuming Hardy-Weinberg equilibrium, the expected numbers of probands homozygous B8-positive, heterozygous and homozygous B8-negative should be 10.1, 50.8

TABLE VIII-36. HARDY-WEINBERG EQUILIBRIUM OF HLA-A, -B AND -C ALLELES

A. In 165 unrelated probands of the Barts-Windsor Family Study. All alleles occurring 7 or more times are included, except Cw6-Cw8 for which typing was incomplete. The equilibrium of each allele with its complement is examined individually.

HLA specif'ty	No. positive		Gene frequency		Genotype distribution							
	Phenotps (n=165)	Alleles (n=330)	Observed	Expected (Hardy-W)	Homoz.+ve Obs.	Exp.	Heterozgs Obs.	Exp.	Homoz.-ve Obs.	Exp.	Diff. Obs-Exp: $X^2_{(2)}$	p
HLA-A:												
1	81	89	0.2865	0.2697	8	12.0	73	65.0	84	88.0	2.5015	NS
2	92	102	0.3349	0.3091	10	15.8	82	70.5	73	78.8	4.4146	NS
3	34	35	0.1091	0.1061	1	1.9	33	31.3	131	131.9	0.4941	NS
11	11	12	0.0339	0.0364	1	0.2	10	11.6	154	153.2	3.0169	NS
24	32	32	0.1022	0.0970	0	1.6	32	28.9	133	134.6	1.9026	NS
28	12	12	0.0370	0.0364	0	0.2	12	11.6	153	153.2	0.2350	NS
29	9	9	0.0277	0.0273	0	0.1	9	8.8	156	156.1	0.1297	NS
31	8	8	0.0245	0.0242	0	0.1	8	7.8	157	157.1	0.1018	NS
HLA-B:												
7	21	21	0.0658	0.0636	0	0.7	21	19.7	144	144.7	0.7621	NS
8	89	96	0.3213	0.2909	7	14.0	82	68.1	76	83.0	6.9067	<0.05
14	7	7	0.0214	0.0212	0	0.1	7	6.9	158	158.1	0.0775	NS
18	15	16	0.0645	0.0485	1	0.4	14	15.2	150	149.4	1.0670	NS
27	11	11	0.0339	0.0333	0	0.2	11	10.6	154	154.2	0.1962	NS
35	14	15	0.0434	0.0455	1	0.3	13	14.3	151	150.3	1.3985	NS
39	18	18	0.0561	0.0545	0	0.5	18	17.0	147	147.5	0.5492	NS
44	33	35	0.1056	0.1061	2	1.9	31	31.3	132	131.9	0.0140	NS
60	26	27	0.0822	0.0818	1	1.1	25	24.8	139	139.1	0.0117	NS
62	47	51	0.1543	0.1545	4	3.9	43	43.1	118	117.9	0.0012	NS
HLA-Cw:												
1	7	7	0.0212	0.0214	0	0.1	7	6.9	158	158.1	0.0775	NS
2	12	12	0.0364	0.0370	0	0.2	12	11.6	153	153.2	0.2350	NS
3	81	88	0.2667	0.2865	7	11.7	74	64.5	84	88.7	3.5507	NS
4	16	17	0.0515	0.0497	1	0.4	15	16.1	149	148.4	0.8021	NS
5	32	34	0.1030	0.1022	2	1.8	30	30.5	133	132.8	0.0438	NS

B. HLA-B8: POOLED DATA. BWFS = Barts-Windsor Family Study probands, NDD = newly-diagnosed cases ascertained during BWFS, BDA = BDA Multiplex family study (augmented).

Source of data	Total number of subjects	No. positive for HLA-B8		Gene frequency observed	Genotype distribution						Diff. Obs-Exp:	
		Phenotps	Alleles		Homoz.+ve Obs.	Exp.	Heterozgs Obs.	Exp.	Homoz.-ve Obs.	Exp.	$X^2_{(2)}$	p
BWFS	165	89	96	0.2909	7		82		76			
NDD	30	14	14	0.2333	0		14		16			
BDA	125	63	71	0.2840	8		55		62			
Pooled	320	166	181	0.2828	15	25.6	151	129.8	154	164.6	8.5342	<0.025

and 64.1 respectively. The observed numbers are 8, 55 and 62 respectively. Again there appears to be a small excess of heterozygotes, but here it is not significant (X^2 = 0.8527, 2 df).

Since there is no evidence of heterogeneity between the BDA study and the BWFS in respect of HLA-B8 genotype distribution (X^2 for heterogeneity = 1.3259, 2df), it is reasonable to pool the data. Table VIII-36-B shows the data from these sources and also from the 30 newly-diagnosed cases ascertained during the BWFS. After pooling, this excess over Hardy-Weinberg expectation is significant at the 2.5% level.

TABLE VIII-37. HARDY-WEINBERG EQUILIBRIUM OF COMPLEMENT FACTOR B (Bf) ALLOTYPES IN 53 UNRELATED PROBANDS OF THE BARTS-WINDSOR FAMILY STUDY, partially selected for presence of HLA-B18.

Bf allo-type	Genotype distribution						Diff. Obs-Exp:	
	Homoz.+ve Obs.	Exp.	Heterozgs Obs.	Exp.	Homoz.-ve Obs.	Exp.	$X^2_{(2)}$	p
F	2	0.7	8	10.6	43	41.7	3.2657	NS
F1	1	0.8	11	11.4	41	40.8	0.0670	NS
S	32	30.9	17	19.1	4	2.9	0.6427	NS
S1	0	0	0	0	0	0		

TABLE VIII-38. HARDY-WEINBERG EQUILIBRIUM OF COMPLEMENT FACTOR C2 ALLOTYPES B and C IN 106 UNRELATED PROBANDS OF THE BARTS-WINDSOR FAMILY STUDY. C2-A was not detected in this sample.

C2 genotype	Genotype distribution		Diff. Obs-Exp:	
	Observed	Hardy-W'bg expectation	$X^2_{(2)}$	p
B-B	0	0.2		
B-C	10	9.5	0.2598	NS
C-C	96	96.2		

TABLE VIII-39. HARDY-WEINBERG EQUILIBRIUM OF ERYTHROCYTE GLYOXALASE-I (GLO) ALLOTYPES 1 and 2 IN 155 UNRELATED PROBANDS OF THE BARTS-WINDSOR FAMILY STUDY.

GLO genotype	Genotype distribution		Diff. Obs-Exp:	
	Observed	Hardy-W'bg expectation	$X^2_{(2)}$	p
1-1	36	30.7		
1-2	66	76.6	2.9526	NS
2-2	53	47.7		

Tables VIII-37, VIII-38 and VIII-39 give corresponding results for Bf, C2 and GLO genotype distributions. No significant deviation from Hardy-Weinberg distribution is found.

2-B-iii-b. HLA-association: B8,15 versus B8,8 and B15,15.

The finding of an excess of HLA-B8 heterozygotes over homozygotes in probands (see above) is compatible with the notion that B8,15 is more closely associated with Type 1 diabetes than are the homozygous genotypes B8,8 and B15,15 (III-6-C-i; III-6-E-iii). One way of testing this further is to determine the frequencies of the various possible genotypic combinations of alleles 8, 15 and X (the remainder) observed in diabetic probands, and to compare these observed frequencies with those to be expected if inheritance of each of these alleles from a parent were random, independent of that from the other parent, and had a probability proportional to its gene frequency (the principle underlying Hardy-Weinberg equilibrium). This may simply be done by testing for random distribution of diabetic subjects with respect to these three alleles, both paternal and maternal, in a 3x3 contingency table.

Table VIII-40 summarizes such an analysis of the genotype frequencies observed in 310 probands from the Barts-Windsor study, the newly-diagnosed series ascertained during the latter, and the augmented BDA multiplex study. In five cases HLA-B homozygosity is assumed since only one specificity was detected; two of these typed B8, and one each Bw60, Bw50 and B7. In all other cases genotypes were confirmed.

Because of the possibility of genetic heterogeneity between familial and sporadic Type 1 diabetes, and also because multiplex families are more likely to possess multiple diabetes susceptibility genes than are simplex, probands from simplex and multiplex sibships are categorized separately in Table VIII-40-A. In fact, log-linear analysis (Table VIII-40-B) shows that no term in N (no. of diabetic members in sibship) is significant, so the two 3x3 tables may be amalgamated (Table VIII-40-D). Here, distribution differs significantly from random ($p < 0.025$ for term P.M in Table VIII-40-B; similarly, Table VIII-40-D gives $X^2 = 12.8764$,

TABLE VIII-40. GENOTYPIC COMBINATIONS OF THE HLA-B 'ALLELES' 8, 15 AND X: FREQUENCIES IN 310 TYPE-1 DIABETIC PROBANDS. X represents any allele other than 8 or 15.

A. Frequency data in full.

N NDChil	1 (simplex)				2 (multiplex)			
P Pat. HLA-B	X	15	8	All	X	15	8	All
M Mat. HLA-B								
X	61	11	35	107	44	10	21	75
15	12	4	12	28	7	0	5	12
8	30	9	8	47	24	10	7	41
All	103	24	55	182	75	20	33	128

B. Log-linear analysis: interaction effects. All models shown here include terms for grand mean and main effects.

Model ref. no.	Terms included				Overall goodness/fit			Difference				
	1st-order			2nd	X^2_L	DF	p	Model	X^2_L	DF	p	Term
	P.M	P.N	M.N	P.M.N								
2	+	+	+	-	3.920	4	NS	Sat.	3.920	4	NS	P.M.N
3	+	-	-	-	7.529	8	NS	2	3.609	4	NS	P.M, M.N
4	-	-	-	-	20.38	12	NS *	3	12.85	4	<.025	P.M

Note: * critical value of X^2 = 21.03 for p = 0.05, 12 degrees of freedom.

C. Parameter estimates for model 3 in table B.

Genotype	Parameter	Estimate	SE	St.Est
15,15	P2.M2	0.0513	0.5997	0.0855
15,8	P2.M3	0.5649	0.3582	1.5771 .
8,15	P3.M2	0.5174	0.3726	1.3886 .
8,8	P3.M3	-0.6523	0.3315	-1.9677 *

Key: . = 1.0 < standardized estimate < 1.96; * = st. est. > 1.96

D. Table A, collapsed over N.

P Pat. HLA-B	X	15	8	All	Gene frequency (mean pat & mat)	
M Mat. HLA-B						
X	105	21	56	182	X	0.5806
15	19	4	17	40	15	0.1355
8	54	19	15	88	8	0.2839
All	178	44	88	310	All	1.0000

E. Genotype frequencies: observed, and expected by Hardy-Weinberg equilibrium.

Genotype	X,X	15,X	8,X	15,15	8,8	8,15	All
Frequency Observed	105	40	110	4	15	36	310
Expected	104.5	48.8	102.2	5.7	25.0	23.9	310.1

X^2 (5 degrees of freedom) = 12.8367 (p < 0.025)

4 degrees of freedom, $p < 0.025$), while its simplified derivative, Table VIII-40-E, gives a similar result.

The question particularly at issue now is whether these genotype frequencies are compatible with the hypothesis that diabetes susceptibility is determined by a single gene *s* in recessive mode, this single gene being in linkage disequilibrium with both B8 and B15. The hypothesis predicts that in diabetic subjects the frequencies of genotypes 8,15, 8,8 and 15,15 will all be similarly increased above Hardy-Weinberg expectations. Inspection of the Table VIII-40-D suggests that, while there is indeed an excess frequency of B8,15, there is on the contrary a deficit of 8,8 (the expected and observed frequencies of 15,15 are too small to make any important contribution). Is this difference between 8,8 and 8,15 significant?

The parameter estimates for Model 2 (Table VIII-40-B and C) are informative here. It should be noted that GLIM sets to zero the parameter estimate relating to the first category of any variable involved in a given interaction, and computes the estimates pertaining to other categories relative to this. Here, the contingency table is so arranged that the first category of both P and M represents allele X, enabling comparison of parameter estimates for terms in 8 and/or 15 only. For parameter P3.M3, representing the 8,8 homozygotes, the (absolute) standardized estimate exceeds 1.96 and hence the estimate is significant. Since this parameter has negative sign (i.e. there is a deficit of 8,8 genotypes), while the estimates for 15,8 and 8,15 (P2.M3 and P3.M2) are positive and moderately large, it follows that the respective frequencies of 8,8 and 15,8/8,15 probands deviate from Hardy-Weinberg expectation in significantly different ways.

A more rigorous treatment follows.

Let the probability of genotype a under the recessive hypothesis

$$P_r(a) = R(a).P_h(a),$$

where $P_h(a)$ is the probability of genotype a under Hardy-Weinberg equilibrium, and R(a) is a linking function which is itself a function of

the strengths of association between the recessive susceptibility gene and the HLA-B genes constituting genotype a. Now R(a) will also be a function of other variables such as the frequency of the recessive gene s in the population, and perhaps the penetrance of the susceptible genotype ss. Direct calculation of its actual value is therefore not feasible. However, HLA-B8 and B15 (or 62) have similar strengths of association with diabetes, as shown for the BWFS by the 'rel. risk' figures in Table VIII-1. Hence, within the constraints of the hypothesis and as a reasonable approximation, R(a) can be considered to be a constant (R) for the three genotypes under consideration.

Thus, a maximum-likelihood estimate for the expected number of subjects having genotype 8,15 among N type 1 diabetics, under the recessive hypothesis, is given by

$$E_r(8,15) = N.R.P_h(8,15) = R.E_h(8,15) \qquad [1]$$

(where E_h is the number expected assuming Hardy-Weinberg equilibrium), and similarly for genotypes 8,8 and 15,15. Now, if the hypothesis is true, the observed number should not differ significantly from those expected, so

$$O_r(8,15) + O_r(8,8) + O_r(15,15) = R.(E_h(8,15) + E_h(8,8) + E_h(15,15)),$$

whence $$R = \frac{O_r(8,15) + O_r(8,8) + O_r(15,15)}{E_h(8,15) + E_h(8,8) + E_h(15,15)} = 1.0073$$

(using the values given in Table VIII-40-E). The small value of R results from the fact that the joint deficit of 8,8 and 15,15 almost exactly balances the excess of 8,15 when observed are compared with expected frequencies.

Using this value for R in Equation [1], the frequencies of the three HLA-B genotypes expected under the recessive hypothesis are:

$$E_r(15,15) = 5.7; \quad E_r(8,8) = 25.2; \quad E_r(8,15) = 24.9.$$

Since both $E_r(15,15)$ and $O_r(15,15)$ are small, the 15,15 genotype is combined with the 8,8 to form the 'homozygous' group. Thus the observed frequencies are ... (homozygous and 8,15, respectively)

19 and 36

while those expected are 30.9 and 24.9,

the difference being significant at the 1% level (χ^2 = 8.7932, 1 df). Therefore the hypothesis is rejected. Susceptibility to Type 1 diabetes is not determined by a single, recessive gene.

2-B-iii-c. Role of HLA genes themselves in susceptibility.

This question is approached here in two ways: first by comparing 330 'diabetic' and 307 'control' haplotypes from the BWFS, and secondly by examining the effects of HLA-B alleles on haplotype concordance in diabetic sib pairs.

In the first approach, linkage disequilibrium between pairs of alleles at different HLA loci is calculated by gene counting (see III-6-B-iii) in probands, and is compared with the value obtained in 'control' haplotypes (i.e. those which do not appear in any diabetic family member). Any difference is tested in tridimensional contingency tables: the three variables included are (B) HLA-B (dichotomized, e.g. B8+ and B8-); (A) HLA-A (similarly); and (D) 'diabetes' (presence or absence). Thus each layer includes all four mutually exclusive and exhaustive classes of haplotype in one set of haplotypes.

If overall χ^2 indicates presence of significant interactions, the table is tested, using hierarchical log-linear models, against the hypothesis of no second-order interaction. If second-order interaction is present this implies that the strength of linkage disequilibrium is related to genetic susceptibility to Type 1 diabetes, and that the presence of the HLA-A allele concerned influences the association between the B allele and diabetes (epistasis).

For this analysis 'probable homozygotes' at one or both of the the loci concerned are again considered positive on both haplotypes for the single specificity detected. This is acceptable for HLA-A and -B since no confirmed blanks have been found in the families. For Cw the probability of error is small - only four probands type Cw'x'.

In Table VIII-41-A, the two-locus linkage disequilibria investigated are those between alleles which are both sufficiently frequent for

TABLE VIII-41. LINKAGE DISEQUILIBRIA WITHIN THE HLA REGION: COMPARISON BETWEEN 'DIABETIC' AND 'CONTROL' HAPLOTYPES IN THE BARTS-WINDSOR FAMILY STUDY.

A. Two-locus linkage disequilibria (Allele 1 at one locus, Allele 2 at the other).

Variable		Haplotype frequencies and relative l.d. parameters									Log-linear analysis	
Diabetes		+					-				Model with main effects only	Model lacking only 2nd-order term D.1.2
Allele 1		+		-		Δ_r	+		-		Δ_r	
	Allele 2	+	-	+	-		+	-	+	-		X^2_L(4df) p	X^2_L(1df) p
B8	A1	68	28	21	213	0.667	28	9	20	250	0.712	251.2 <0.001	0.6975 NS
B7	A3	5	16	30	279	0.148	18	32	45	231	0.294	40.00 <0.001	0.7062 NS
B39	A24	10	8	22	290	0.508	1	3	25	278	0.181	36.51 <0.001	1.522 NS
Bw62	Cw3	49	2	32	245	0.948	11	2	16	277	0.831	252 <0.001	0.375 NS
Bw60	Cw2	1	26	11	290	0.001	0	11	14	281	0.048	7.636 NS	0.848 NS
B8	BfS	26	0	55	25	0.313	9	0	69	12	0.148	29.89 <.001	No convergence

B. The three-locus linkage disequilibrium between B18, Cw5 and BfF1.

Variable		Haplotype frequencies								Log-linear analysis	
D Diabetes		+				-				Model with main effects only	Model lacking B.C.F.D and all 2nd-order terms in D
F BfF1		+		-		+		-			
C Cw5		+	-	+	-	+	-	+	-	X^2_L(11df) p	X^2_L(1df) p
B B18	+	11	0	0	5	3	0	2	6	119.3 <.001	2.256 NS
	-	0	2	8	80	0	0	8	71		

analysis and associated with Type 1 diabetes in the population. Each row of the table relates to one particular pair of alleles, one (allele 1) at one locus and the other (allele 2) at another on the same haplotype. The frequency data given are the numbers of haplotypes observed to fall into the four cross-classification categories, positive or negative for each of the two alleles concerned, first in diabetic and then in control haplotypes. In most cases the relative linkage disequilibrium parameter Δr, calculated as described (III-6-B-iii, Eqn. 6-7), is obviously similar in the two groups. No second-order interaction term (Diabetes.Allele1.Allele2) is significant. However, the term Diabetes.B8.BfS might prove to be significant given larger numbers of subjects; the table concerned could not be collapsed over the Diabetes variable, which was significantly associated with both BfS ($p < 0.025$) and B8 ($p < 0.01$) (see also VIII-1-B-ii-a and Table VIII-5).

Table VIII-41-B is similar, but examines the three-locus linkage disequilibrium between B18, Cw5 and BfF1 (see VIII-1-B-ii). Again there is no evidence for any interaction between diabetes and this linkage disequilibrium, since no second- or third-order term in D is significant. The data array is sparse, but the frequency distributions in diabetic and control groups are very similar and it would be surprising if study of much larger numbers of subjects were to alter the conclusion.

Thus there is no evidence here for any direct effect of HLA-A, B or Cw alleles on susceptibility (or liability) to Type 1 diabetes.

The second approach is to examine whether HLA haplotype concordance in pairs of diabetic sibling is influenced by the HLA alleles present. Table VIII-42 shows separately for various HLA-B specificities the segregation pattern of parental HLA haplotypes to pairs of diabetic siblings in 142 duplex families from the augmented BDA multiplex family study, the BWFS and the Liverpool series (VII-C). Triplex and tetraplex sibships are omitted for simplicity of analysis and to avoid bias. Where a parent's two haplotypes cannot be distinguished because of homozygosity or non-availability for typing, that parent's haplotypes are omitted from the analysis (27 cases). Thus haplotypes from 257 parents are available for study.

In Table VIII-42 the 514 parental haplotypes are classified according to the HLA-B allele present and the number of occurrences in the diabetic children. Each haplotype (a, b from the father, c, d from the mother) is scored 2 if it occurs in both diabetic diabetic children (a 'repeat'), 1 if it occurs in only one diabetic child, and 0 if it occurs in neither.

There are two possible segregation patterns for the two haplotypes from a given parent: 2-0 (a 'repeat') and 1-1 (a 'single'). In the absence of linkage between HLA and diabetes these would have equal probabilities, and the proportions of haplotypes scoring 2, 1 and 0 respectively would then be 0.25, 0.50 and 0.25. Overall - irrespective

TABLE VIII-42. HAPLOTYPE SHARING AND HLA-B SPECIFICITIES IN DUPLEX SIBSHIPS: SEGREGATION TO DIABETIC CHILDREN OF PARENTAL HAPLOTYPES CATEGORIZED BY HLA-B ALLELE.

Score (no. of occurrences in diabetic sib pair)	Number of parental haplotypes, categorized by HLA-B						
	'High-risk' (BS)				BX		Total
	B8	B18	B15	Bw40	B7	Other (BY)	
2	52	17	22	20	16	61	188
1	31	5	12	10	12	68	136
0	17	6	13	12	26	114	188
Total	100	28	47	42	54	243	514

Note: Each haplotype (a, b from the father, c, d from the mother) is scored 2 if it occurs in both diabetic diabetic children (a 'repeat'), 1 if it occurs in only one diabetic child, and 0 if it occurs in neither.

TABLE VIII-43. HAPLOTYPE SHARING AND HLA-B SPECIFICITIES IN DUPLEX SIBSHIPS: SEGREGATION TO DIABETIC CHILDREN OF BS-POSITIVE HAPLOTYPES FROM PARENTS WITH GENOTYPE BS,X, COMPARED WITH THAT OF BX-POSITIVE HAPLOTYPES FROM PARENTS WITH GENOTYPES BS,S and BX,X, AND SEGREGATION OF B7-POSITIVE HAPLOTYPES FROM PARENTS WITH GENOTYPE BY,Y COMPARED WITH THAT OF BY-POSITIVE HAPLOTYPES FROM PARENTS WITH GENOTYPE BY,Y where BS is any of the 'high-risk' HLA-B alleles (8, 18, 15, w40), BX is any other HLA-B allele (including B7), and BY is any allele other than BS or B7.

A. Haplotype distribution data.

Number of occurrences of haplotype in diabetic sib-pair	HLA-BS versus BX					HLA-B7 versus BY			
	Parent BS,X			Parent BS,S	Parent BX,X	Parent B7,Y			Parent BY,Y
	Haplotypes			All haplotypes	All haplotypes	Haplotype			All haplotypes
	BS	BX	All			B7	BY	All	
2	82	19	101	29	59	8	4	12	44
1	32	32	64	26	46	8	8	16	30
0	19	82	101	29	59	4	8	12	44
Total	133	133	266	84	164	20	20	40	118

B. Comparison of segregation patterns to diabetic sib-pairs in the three groups.

Segregation pattern	HLA-BS versus BX				HLA-B7 versus BY		
	Frequency of segregation pattern				Frequency of pattern		
	Parent BS,X	Parent BS,S	Parent BX,X	X^2 p (2 df)	Parent B7,Y	Parent BY,Y	p (Fisher exact test)
2-0	101 (76%)	29 (69%)	58 (70%)		12 (60%)	44 (75%)	
1-1	32	13	24	1.1312 NS	8	15	0.1692 NS
Total	133	42	82		20	67	

of HLA-B - the excess of 'repeats' observed is well-recognized and due to linkage between diabetes and HLA (Green & Woodrow, 1977). The number of haplotypes scoring 0 is constrained to equal the number scoring 2 because all haplotypes from the 257 parents are included.

Within this distribution, the haplotypes having a 'high-risk' B allele BS (i.e. B8, B18, B15 or Bw40, known to be associated with diabetes in the population) show a marked excess scoring 2 and a deficit scoring 0, with a contrary trend among the remaining haplotypes (BX). The difference between the distributions of BS and BX haplotypes is highly significant (X^2 = 43.38, 2 df, $p << 0.001$; BS considered as a single class), and is to be expected because of the association between BS and diabetes in the population: in a parent with one BS and one BX haplotype, the BS haplotype is the more likely to carry a susceptibility gene and therefore the more likely to be present in diabetic offspring.

The question now is whether this effect of BS on segregation to diabetic children is entirely explicable by linkage disequilibrium with a susceptibility gene (the null hypothesis), or alternatively whether, given that a haplotype occurs in a diabetic child and therefore presumably carries a susceptibility gene(s), the presence of BS on that haplotype further increases the probability that it occurs also in the other member of the diabetic sib pair. This may be investigated by subdividing the parents according to genotype (BS,S; BS,X; BX,X). Table VIII-43-A gives the numbers of haplotypes occurring twice, once or not at all in the diabetic sib pairs.

We are interested in the overall segregation patterns in the three groups of parents, summarized in Table VIII-43-B. If the null hypothesis is true, the overall proportions of repeats (pattern 2-0) should be similar. If BS does have a direct effect on liability to diabetes, the proportion of repeats is likely to be greatest when one and only one haplotype is BS-positive, while when the two haplotypes of a parent are evenly matched in terms of BS, a 1-1 segregation pattern is somewhat more likely and hence the proportion of repeats should be lower. In fact, no significant difference between the three groups is present. (It might be

argued that the correct test here is a one-tailed test for an excess of repeats in the BS,X group over the combined BS,S and BX,X groups. In this case we obtain corrected X^2 = 0.8166 for 1 df; p = 0.1831.) Thus we cannot reject the hypothesis that there is no direct contribution of BS to liability to diabetes.

A similar analysis may be used to investigate whether HLA-B7 has a protective effect against diabetes. In Table VIII-42, B7-positive haplotypes show a deficit of 'singles' and, given a 'repeat' pattern in a sibship, a deficit in the diabetic siblings, as expected from the known linkage of HLA with susceptibility and the negative association between B7 and Type 1 diabetes. However, the distribution of B7 haplotypes among the groups scoring 2, 1 and 0 does not differ significantly from that of BY (i.e. alleles other than B7 and BS; X^2 = 0.91 for 2 df, figures of Table VIII-42), nor does the segregation pattern to diabetic siblings from parents with genotype B7,Y differ significantly from that when parents have genotype BY,Y (Table VIII-43-B). This is compatible with the view that the deficit of B7 in Type 1 diabetes is secondary to excess of BS alleles, rather than reflecting a protective gene.

A further implication of these results is that they provide evidence against genetic heterogeneity between sib pairs positive and those negative for 'high risk' HLA-B specificities: the proportion of repeats among haplotypes from parents with genotype BS,S is identical to that in the case of parents with genotype BX,X.

2-C. DISCUSSION

The first part of the Results section (2-B-i) confirms that the pattern of HLA haplotype concordance between diabetic siblings is incompatible with Mendelian dominant inheritance but quite compatible with recessive inheritance. However, this does not prove that inheritance is in fact recessive. Indeed, the minimum gene frequency for which the recessive model fits is for other reasons unacceptably high (III-6-E-ii). Further and compelling evidence against recessive

inheritance is the excess of B8,15 heterozygotes over that expected under Hardy-Weinberg equilibrium documented above (VIII-2-B-iii-a and -b).

The lack of any evidence (in the haplotype concordance analysis) for heterogeneity between duplex and triplex/tetraplex sibships is itself interesting. It might be supposed that multiple susceptibility genes would be particularly likely in the latter group, but the gene frequency for minimum χ^2 in Figure VIII-2-B is virtually identical to that for duplex sibships. One interpretation is that highly multiplex families are, as a group, genetically no different from duplex families, non-genetic factors being responsible for an increase in penetrance of the susceptible genotype.

An important question is whether there is genetic heterogeneity between familial and sporadic Type 1 diabetes (III-6-E; III-10-C-ii-a). In the discrete multivariate analyses of section VIII-2-B-ii-d, 380 probands from simplex and 179 from multiplex sibships show no significant differences in frequency of HLA B8, B15 or B8,15, in sex ratio, or in age at diagnosis (provided ascertainment effects are taken into account). Likewise, the frequency of thyrogastric antibodies is similar in 181 simplex and 127 multiplex probands (VIII-2-B-ii-e, final analysis reported for B8).

Further and perhaps more telling evidence against this form of heterogeneity is that deviation from Hardy-Weinberg equilibrium between HLA-B 8, 15 and other alleles is identical in simplex to that in multiplex probands (VIII-2-B-iii-b; Table VIII-40). This is incompatible with the view that susceptibility in simplex families is predominantly recessive whereas in multiplex families it is not (Greenberg & Anderson 1983; III-10-C-iii-b), since in that case the simplex group should show an excess of homozygotes rather than the observed deficit.

However, the analyses reported here do detect one interaction involving the number of diabetic siblings of the probands: only the simplex probands aged > 5 years show seasonal variation in time of diagnosis, thereby differing significantly from the multiplex probands

(VIII-2-B-ii-e, Table VIII-31 and Figure VIII-4). This cannot be explained by earlier diagnosis because of increased awareness in a family where one child is already diabetic, since for the purpose of this study the proband in a multiplex sibship is always the first to have been diagnosed (except in five cases where insufficient information is available on that member). Possibly it represents a chance aberration in the sample studied (a fluke). If not, taken together with the lack of any other evidence for genetic heterogenity in this respect, it suggests that there is some common factor directly or indirectly influencing both the probability of and any seasonal effects on development of clinical diabetes in a child with HLA-linked susceptibility to the disease. This common factor may or may not be genetic. If it is, then it is likely to be independent of HLA.

Another interaction of season and age at diagnosis is with HLA-B8 (VIII-2-B-ii-e; Table VIII-30; Figure VIII-3). So far as can be determined, this is independent of number of diabetic siblings. Since the interval between initiation of the disease process and clinical diabetes is probably much longer than a season (Chapter VIII-C), such marked seasonal variation of incidence as is shown here is unlikely to be relevant to that disease process itself. It is quite conceivable that HLA-B8 or associated gene(s) could alter host responses to different environmental agents (viruses, for example), which could then precipitate metabolic decompensation to frank diabetes in subjects with critically reduced islet B-cell mass. Since susceptibility to different viruses varies with age, such a mechanism would readily explain the observations summarized in Figure VIII-3, without any implications for heterogeneity of genetic susceptibility to diabetes.

None of the analyses performed supports some earlier reports that HLA-B15 or B8,15 is particularly associated with onset in early childhood rather than later (VIII-2-B-ii-d). Nor is there any evidence here in favour of the concept of Type 1B diabetes: among the diabetic probands, presence of thyrogastric antibody is indeed associated with HLA-B8 and with female sex, just as would be expected in a non-diabetic population, but it is independent of age at diagnosis and of the other variables

directly relevant to the diabetes (season at diagnosis; number of affected siblings) (VIII-2-B-ii-e). However, it should be remembered that clinical Type 1B diabetes typically begins after the third decade of life (III-10-C-iii-c), well beyond the upper limit of age at diagnosis of the probands informative for thyrogastric antibody in this study.

The observation that HLA-B15 frequency differs quite markedly between the Windsor area and Liverpool probably results from genetic variation in the parent populations. This explanation, rather than technical differences in HLA typing, is supported by the intermediate frequency found in the BDA study covering the whole of Great Britain, for which HLA typing was performed at the same time and in the same laboratory as that for the Barts-Windsor study. It is not relevant to diabetes, but because of differences in ascertainment it might easily have led to spurious conclusions concerning HLA and age at diagnosis or number of affected siblings, if the source of data had not been included as one variable in the analyses.

The association of HLA-B15 (in fact Bw62) with male sex in the BWFS and BDA studies is curious, especially as it is absent in Liverpool, but is difficult to disregard with a significance level of 1% (VIII-2-B-ii-d, Table VIII-25; VIII-2-B-ii-e, Table VIII-33). Further confirmation will need to be sought in new samples.

In conclusion, there is some support for genetic heterogeneity within Type 1 diabetes, expressed not as two or three distinct forms but as minor variation in clinical, epidemiological and immunological features. A number of different genes on several chromosomes contribute to susceptibility, including at least two in or near the HLA complex (see above), probably a non-HLA-linked gene influencing expression of autoimmunity (see VIII-5-C), and others mentioned in the Literature Survey (III-6-D).

These features suggest multifactorial (polygenic) inheritance, which would be confirmed (irrespective of heterogeneity) if it could be shown that incidence in identical twins exceeded by a factor of more than four

that in other siblings of affected individuals (III-5-C). In UK some 3 - 5% of siblings overall are themselves affected by age 16 years (III-6-F; VIII-4-d; Table VIII-5). The rate in identical twins is not known, since ascertainment bias undoubtedly contributes to the figure of 55% observed in UK (Barnett et all, 1981), but it is hardly likely to be as low as 20%.

It is therefore highly likely that Type 1 diabetes is genetically multifactorial. In that case, the fact that the significant relative risk associated with a particular gene is low, as in the case of HLA genes, does not exclude a direct contribution to susceptibility. Two novel approaches are used here (VIII-2-B-iii-c) to seek any direct effect of HLA-A or -B alleles on susceptibility to Type 1 diabetes: discrete multivariate analysis to determine whether 'diabetic' and 'control' haplotypes differ in linkage disequilibria between alleles at loci in the HLA region, and assessment of differences in haplotype concordance in diabetic sib pairs. No such direct effect is found, nor do haplotype concordance analyses support a 'protective' effect associated with HLA-B7. This is hardly surprising, since any such direct effects are likely to involve HLA-DR3 and DR4 (susceptibility) or DR2 (protection). These methods will, it is hoped, be used again with HLA-DR data.

PART VIII — CHAPTER 3

THE PROSPECTIVE STUDY: THE PRE-DIABETIC PERIOD.

ICA AS A MARKER. ENVIRONMENTAL FACTORS

CONTENTS

A Introduction · 280
B Observations and analyses · 281
i Prospective follow-up: numbers and duration · · · · · · · · · · 281
ii Islet-cell antibodies after diagnosis of diabetes · · · · · · · 282
iii Islet-cell antibodies in non-diabetic subjects · · · · · · · · 284
iv Environmental events · 287
v New cases of diabetes developing during follow-up · · · · · · · 289
vi Glucose intolerance short of diabetes · · · · · · · · · · · · · 290
a Urinalysis · 290
b Glycated haemoglobin · 291
vii Islet-cell antibody as a marker of the pre-diabetic state · 294
C Discussion · 296

3-A. INTRODUCTION

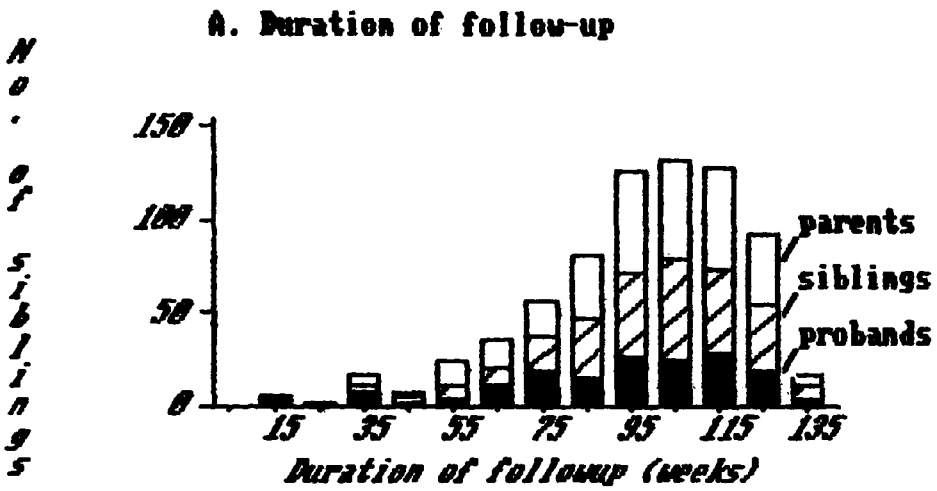
The results reported here relate to families taking part in the Barts-Windsor Family Study and followed up personally by the author during the first 2.5 years from April 1978. Since mid-1980, colleagues have extended and continued the study. Certain early observations are not reported here in any detail since they have been covered more appropriately and more completely by those colleagues.

A paper written by the author and published in December 1981 forms the basis for part of this chapter, and is bound here as Appendix 7-2 (Gorsuch et al, 1981).

3-B. OSERVATIONS AND ANALYSES

3-B-i. Prospective follow-up: numbers and duration

Among the 155 probands, 271 siblings, 145 fathers and 152 mothers followed up serologically in the BWFS (VII-A-ii), the median duration of follow-up by 31st October 1980 was 101 (range 13-133) weeks, and the median number of follow-up blood samples per subject was 8 (range 2-19) (Figure VIII-5-B). The total number of blood samples obtained from these subjects during the period concerned was 5436.



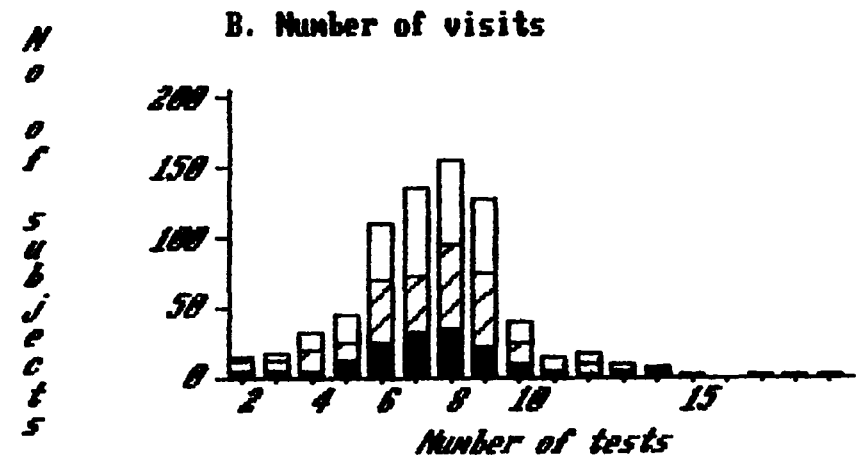


FIGURE VIII-5. SEROLOGICAL FOLLOW-UP OF 155 PROBANDS (solid bars), 271 SIBLINGS (diagonal-striped) AND 297 PARENTS (open bars) by 31st October 1980.
A. Duration of follow-up. Frequency distribution diagram giving the numbers of subjects followed up for periods between x-5 and x+5 weeks, where x is the position of a bar on the horizontal axis.
B. Number of visits yielding serological samples. Frequency distribution.

Some form of contact was maintained with all those who withdrew from full participation, sufficient at least to confirm that clinical diabetes did not develop by 31st October 1980, 1.4-2.5 years after entry to the study. One non-diabetic father was killed in a road accident in March 1980.

3-B-ii. Islet-cell antibodies after diagnosis of diabetes

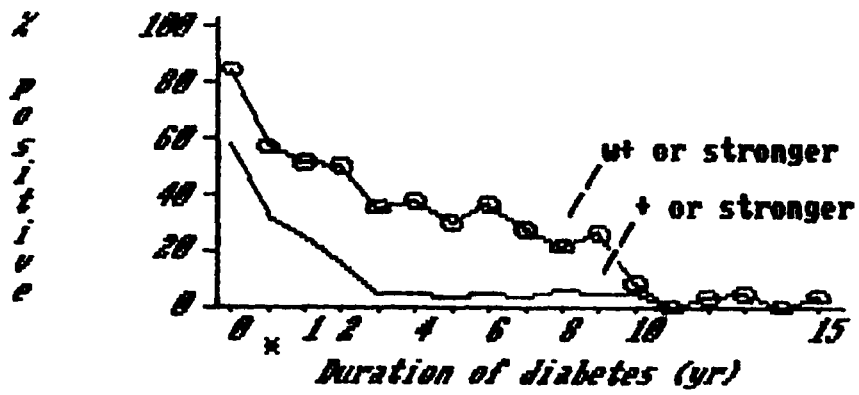
Tests for ICA were performed on 184 diabetic BWFS participants, including those developing diabetes during prospective study, and on 30 newly-diagnosed cases studied at the same time but not entering the prospective study (VII-B). Excluding repeats within two months the total number of such tests by 31st October 1980 was 1230, of which 1138 were within 15.75 years of diagnosis. Among these the proportion of positive results falls with increasing duration of disease (Figure VIII-6).

It is interesting that, after the initial falling curve over the first few years, a low frequency of positive tests persists to the 10-year mark. This applies to both the curve for CF-ICA (Figure VIII-6-B), and that for conv-ICA if '+' is taken as the criterion for a positive result (Figure VIII-6-A). It supports the concept of a subgroup in which ICA persists.

In Table VIII-44, the first available result five or more years after diagnosis is examined for each of the 81 individuals with sufficient duration of diabetes. The frequency of positive results ('persistent ICA') is identical in HLA-B8 positive and B8-negative individuals, and, even though the number of such results is very small (11), no strong association between persistent ICA and B8 can seriously be entertained in this sample.

When sequential ICA results for the same individual are examined, in a few cases disappearance of the antibody is seen to have been followed by its reappearance. In some families, this occurred simultaneously in both the diabetic proband and non-diabetic members. Whether such fluctuation of ICA is genuine or artefactual is not clear. It has been

A. Conventional ICA



Total no. of tests	58 45	117 112	111 113	118 121	70 107	23 39	29 18	16 23	26

B. Complement-fixing ICA

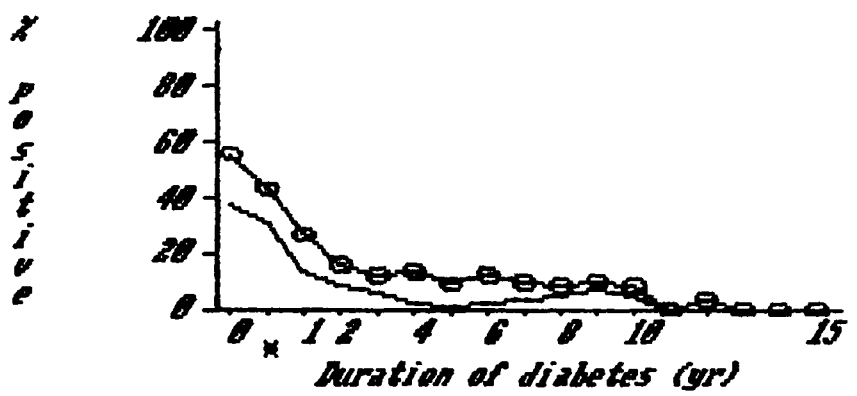


FIGURE VIII-6. ISLET-CELL ANTIBODY: FREQUENCY OF POSITIVE RESULTS VERSUS DURATION OF DIABETES, among 1138 tests performed within 15.75 years after diagnosis of Type 1 diabetes. A: By conventional technique; B: complement fixation technique. Two criteria for 'positive' results are used here: w+ or stronger (circles) and + or stronger (line). * The first point on the X-axis represents tests performed within 1 month of diagnosis, the next represents tests in the period 1 to 9 months, and thereafter the scale is linear. 'n year' represents the period (n-1) yr 9 months to n year 9 months from diagnosis.

TABLE VIII-44. PERSISTENT ISLET-CELL ANTIBODY VERSUS HLA-B8 IN TYPE 1 DIABETES IN THE BARTS-WINDSOR FAMILY STUDY. The table shows frequencies of positive (w+ or stronger reaction) and negative results for the first ICA test performed five or more years after diagnosis of diabetes in each of 81 individuals, categorized by HLA-B8 status.

HLA-B8	Total number of subjects	Conv-ICA positive Number	(%)	CF-ICA positive Number	(%)
Positive	47	6	(13)	3	(6)
Negative	34	5	(15)	2	(6)
Total	81	11	(14)	5	(6)

examined by Dr. K M M Spencer, using data from the extended BWFS (Spencer et al, 1984), and is not considered further here.

3-B-iii. Islet cell antibodies in non-diabetic subjects

The frequency of islet-cell antibodies at entry to the study has been described in an earlier Results chapter (VIII-1-C-iv; Table VIII-16). Further details, and changes during follow-up to 31st October 1980, are summarized in Tables VIII-45-A to -C, which give the numbers of subjects positive for Conv-ICA (by two alternative criteria) and for CF-ICA.

In Table VIII-45-A, a 'weakly-positive' (w+) or stronger immunofluorescence reaction in the conventional test for ICA is counted as positive. At entry to the study, the proportion of siblings positive is significantly associated with increasing haplotype concordance with the proband, being 13% in the pooled group of siblings classed H2 or H2/1, 5% in those classed H1 or H1/0, and 0% in the H0 group (X^2 = 11.0186, 2 degrees of freedom, $p < 0.01$; linear trend X^2 = 10.4976, 1 df, $p < 0.01$). (If the groups with ambiguous haplotype concordance are considered as separate classes, X^2 = 20.7921 with 4 df, $p < 0.001$).

During prospective follow-up ICA was newly detected in a further 20 siblings and 12 parents (positive conversion). The overall frequency of ICA positive at any time during the study is 12% in siblings and 7% in parents. It should be noted that no test for ICA is here counted positive unless confirmed in the same individual by at least a borderline result on another occasion, except where the first positive result was at the last test.

The ICA appearing by positive conversion shows major differences from that present at entry to the study. First, the distribution in siblings favours those sharing no HLA haplotypes with their probands, and is significantly different from that of ICA present at entry (X^2 = 14.7333, 2 df, $p < 0.01$). Secondly, even though all but 2 of the 20 siblings concerned were followed up serologically, the intensity of the 'new' ICA never exceeded w+; in contrast, in 9/14 siblings positive at

TABLE VIII-45. ISLET-CELL ANTIBODIES IN NON-DIABETIC PARTICIPANTS IN THE BARTS-WINDSOR FAMILY STUDY: PROSPECTIVE FOLLOW-UP TO 31ST OCTOBER 1980, OR TO DEVELOPMENT OF DIABETES IF SOONER. Siblings are grouped by haplotype concordance with proband. NOTES: a: 1 converted to negative when treated with carbimazole for thyrotoxicosis; b: 'conversion' from w+; c: 1 from w+; d: 2 from w+; e: to w+

NON-DIABETIC RELATVS	SIBLINGS							PARENTS	TOTAL
Haplotypes shared with proband	MZ twin	2	2/1	1	1/0	0	All sibs		
Total no. tested	1	45	9	145	8	72	280	304	584
Total followed up	1	43	8	139	7	66	264	288	552
A. CONVENTIONAL ICA w+ OR STRONGER:									
At entry	0	4	3	7	0	0	14	8	22
Persisting		4	3	6			13	4	17
Disappearing		0	0	0			0	2 a	2
Not re-tested		0	0	1			1	2	3
Pos.conversion	0	0	0	14	0	6	20	12	32
Persisting				6		4	10	4	14
Disappearing				7		1	8	7	15
Not re-tested				1		1	2	1	3
At any time	0	4	3	21	0	6	34	20	54
At last test	0	4	3	14	0	5	26	11	37

B. CONVENTIONAL ICA + OR STRONGER:	SIBLINGS MZ	2	2/1	1	1/0	0	All	PARENTS	TOTAL
At entry	0	3	1	4	0	0	8	2	10
Persisting		3	1	3			7	1	8
Disappearing		0	0	0			0	1 a	1 a
Not re-tested		0	0	1			1	0	1
Pos.conversion	0	0	1 b	0	0	0	1 b	3 c	4 d
Persisting			0				0	1	1
Disappearing			1 e				1 e	2 e	3 e
Not re-tested			0				0	0	0
At any time	0	3	2	4	0	0	9	5	14
At last test	0	3	1	4	0	0	8	2	10

C. COMPLEMENT-FIXING ICA w+ OR STRONGER:	SIBLINGS MZ	2	2/1	1	1/0	0	All	PARENTS	TOTAL
At entry	0	4	1	3	0	0	8	3	11
Persisting		4	1	2			7	1	8
Disappearing		0	0	0			0	2 a	2 a
Not re-tested		0	0	1			1	0	1
Pos.conversion	0	0	1	3	0	0	4	3	7
Persisting			0	0			0	1	0
Disappearing			1	3			4	2	4
Not re-tested			0	0			0	0	0
At any time	0	4	2	6	0	0	12	6	18
At last test	0	4	1	3	0	0	8	2	10

entry the intensity was graded as +, either from the outset (8 cases) or later (1 case). Thirdly, 8 of the 18 siblings who were re-tested after positive conversion became ICA-negative, unlike any of the 13 siblings positive at entry and followed up serologically.

A likely explanation of these differences is that the sensitivity of the conventional ICA tests performed for this study increased towards the end of 1978 (when the 'positive conversions' began to occur), that ICA detectable by sensitive assays occurs fairly frequently in first-degree relatives of Type 1 diabetic children irrespective of HLA haplotype concordance, but that higher titres (or rather, more intense reactions) are predominantly found in siblings having haplotype concordance with diabetic probands. In support of this, the distribution of siblings found to be ICA positive at any time (Table VIII-45-A, near bottom) is not significantly associated with HLA haplotype concordance with probands (χ^2 = 2.03, 2 df; 'ambigous' groups amalgamated with unambiguous as above). Furthermore, if only reactions classed '+' or stronger are counted as positive and the remainder as negative, there is again a significant association with haplotype concordance of siblings with probands (Table VIII-45-B; χ^2 = 8.8802, 2 df, $p < 0.025$; for linear trend, χ^2 = 7.9915, 1 df, $p < 0.01$).

Table VIII-45-C summarizes complement-fixing ICA observations over the period of prospective study. Once again, results positive at any time show a significant association with haplotype concordance of siblings with probands; with the 'ambigous' groups amalgamated with unambiguous as above, goodness-of-fit χ^2 = 9.3737 with 2 degrees of freedom ($p < 0.01$), while for linear trend χ^2 = 8.9305 with 1 degree of freedom ($p < 0.01$). There is no significant difference in distribution between siblings positive for CF-ICA at entry and positive conversions.

Several other observations are interesting here. Although Table B shows positive conversion to conv-ICA graded '+' or more in four cases, only in two of these was the initial grade definitely negative rather than 'w+'. Similarly, two of the three cases of 'negative conversion' recorded in this table entailed a change of only one grade ('+' to 'w+').

The remaining case is particularly interesting. This mother aged 47 years was positive for conventional ICA (+), CF-ICA (w+) and thyroid microsomal antibodies when first tested. Clinically she was then thyrotoxic, and after confirmation of the diagnosis treatment was started with carbimazole. 7 weeks later CF-ICA was negative, but conv-ICA was still 'w+'. By the third test, 3 months after entry, she was (and has since remained) ICA-negative.

All CF-ICA positive subjects were also positive for conv-ICA, although in one case the latter was first detected two months after the former. In two cases initially ICA-negative, appearance of CF-ICA followed that of conv-ICA by 0.2-0.3 year, in four cases positive for conv-ICA at entry the CF-ICA appeared at tests performed 0.2, 0.3, 0.5 and 1.5 year later, and in 11 cases both conv-ICA and CF-ICA were present when first tested. All four HLA-identical non-diabetic siblings who were positive for conv-ICA were also positive for CF-ICA. CF-ICA reaction strength increased by two categories (e.g w+ to ++) in three siblings, and decreased by two categories in these three as well as in three other cases.

Overall, at some time during the study 25 (8.9%) of the 280 siblings had ICA w+ but not stronger, while in a further 9 (3.2%) the maximum ICA result was + or stronger. Of the non-diabetic parents, 4.9% of parents at some time had detectable ICA never exceeding w+, and 1.6% had results graded as + or stronger.

The relationship between changes in ICA and development of diabetes during prospective study is covered below (3-B-v).

3-B-iv. Environmental events

Families and individual participants clearly varied widely in the reliability of their record diaries (METHODS, VI-1-C-ii), ranging from some who completed them in obsessional detail to a few who ignored them altogether. Diabetic family members and parents commonly did not keep good records, and this analysis is confined to non-diabetic siblings.

In general those siblings who were more assiduous in routine 'monthly' urine testing had the more complete records of minor infections and other events. Among the 241 HLA-genotyped siblings who were followed up for at least 26 weeks and provided at least three blood samples, 37 averaged only four or fewer urine tests per year, and are excluded from further consideration here as uninformative. Figure VIII-7 shows the distribution of urine-testing frequency (tests/year) in the remaining 204.

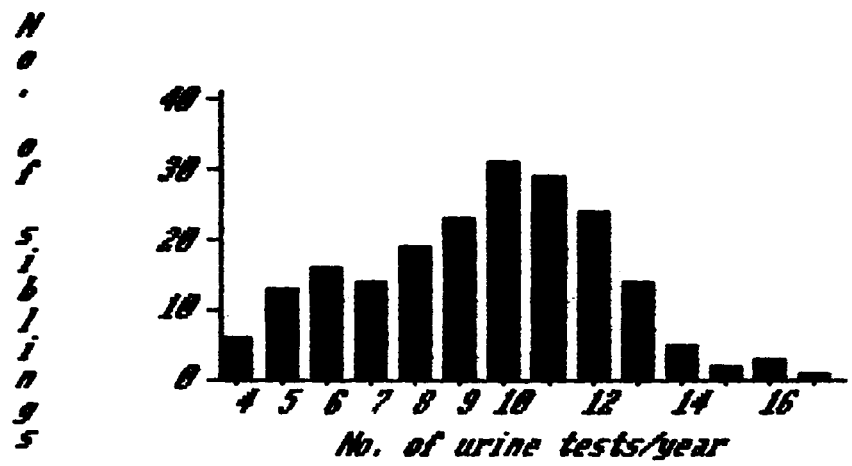


FIGURE VIII-7. FREQUENCY OF URINE-TESTING BY 204 NON-DIABETIC SIBLINGS IN THE BARTS-WINDSOR FAMILY STUDY. Only HLA-genotyped siblings recording four or more urine tests/year and followed up for at least six months with at least three blood samples are included.

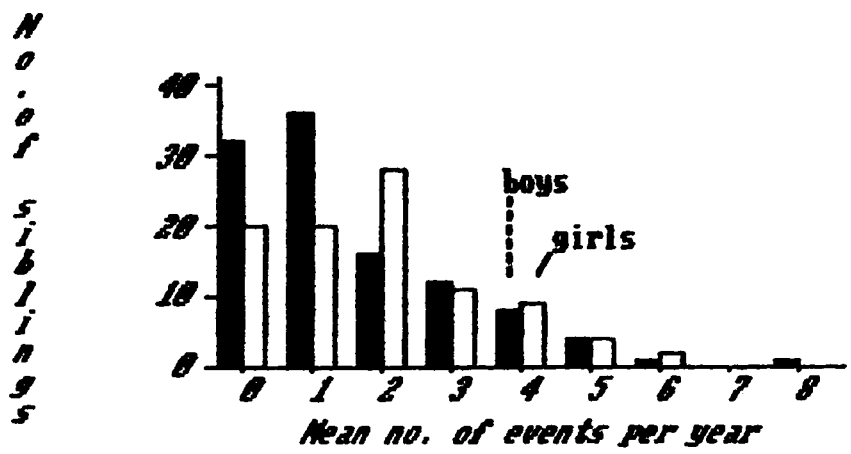


FIGURE VIII-8. RATE OF OCCURRENCE OF ILLNESSES ATTRIBUTABLE TO VIRUSES IN 110 MALE AND 94 FEMALE NON-DIABETIC SIBLINGS IN THE BARTS-WINDSOR FAMILY STUDY. Only HLA-genotyped siblings recording four or more urine tests/year and followed up for at least six months with at least three blood samples are included. Key: solid bars = boys, open = girls.

Events noted in the diaries were recorded individually in the GENOBASE follow-up records. For analysis, the following were automatically counted as 'probable viral infections': upper respiratory tract infections, influenza, measles, glandular fever, herpes labialis, posterior cervical lymphadenopathy, mumps, rubella, chickenpox, herpes zoster (shingles), warts, and a few other illnesses at the discretion of the author.

Probable viral infections were reported significantly more frequently by girls (median 2.3, range 0-6.7 events/year) than by boys (median 1.6, range 0-8.7; $p = 0.0322$ by Mann-Whitney 'U' test), with the distribution shown in Figure VIII-8. Overall there is a weak but significant negative correlation between frequency of viral infections per year and age at entry to the study (Spearman rank correlation coefficient $rs = -0.325$, $p < 0.001$). Four or more infections per year were reported by 39% of siblings aged <8 yr, by 15% of those aged 8 to 15 yr and by 9% of those aged 16 yr or more.

Ideally, information on viral infection in all participants in the BWFS would have been obtained by estimating viral antibody titres in serial blood samples. In practice, however, resources did not permit this because of the large number of tests needed (of the order of 150,000 titres to be determined if 30 viruses were to be studied). Selected sibships were investigated, but during the period up to 31st October 1980 these produced insufficient data for statistical analysis. There was no obvious relationship between changes in viral titre and new appearance of islet-cell antibodies in the few informative subjects studied then.

Probable viral infection rates, as reported by families, are considered further below (3-B-v; 5-B-i-e).

3-B-v. New cases of diabetes developing during follow-up

Five initially non-diabetic family members developed diabetes during the first 2.5 years of the prospective study. All were positive for conv-ICA, and all but one for CF-ICA, by the time of the first test. For

details please see Appendix 7-2, Cases 1, 2, 4, 5 and 6. Case 3 and the Addendum case developed diabetes later.

(Although strictly outside the scope of the present work, it is worth noting subsequent developments in the two cases not requiring insulin before publication in late 1981. Case 3 eventually became insulin-dependent in November 1982, after a total of 4.2 years under observation. Case 6 was treated with glibenclamide; his diabetes remained under satisfactory control until he died of myocardial infarction in October 1982.)

3-B-vi. Glucose intolerance short of diabetes

For logistic reasons it was impossible to perform formal glucose tolerance studies on any but a few, selected, non-diabetic participants in the BWFS during the period up to 31st October 1979. Instead, all non-diabetic subjects were asked to test monthly for glycosuria, and glycosylated haemoglobin (HbA_{1a+b+c}) was assayed on at least two occasions in all available subjects in 1979 - 1980.

3-B-vi-a. Urinalysis

Urine tests were performed by 265 non-diabetic siblings. Out of a total of 4316 tests, 22 were reported as positive by 12 siblings. In these, random blood glucose was checked as soon as possible. In two cases this showed diabetes, confirming clinical symptoms (Cases 1 and 4 above). In another case, an HLA-identical, CF-ICA-positive brother, post-prandial blood glucose was 7.7 mmol/l a few days after the first positive urine test; four weeks later, weekly urine tests were reported as positive on two occasions, but blood glucose remained normal then and later (4.3 - 5.3 mmol/l), while HbA_1 was normal before, during and after the first detected episode of glycosuria. In fact, although this child did not become diabetic duting the follow-up period covered here, he did so later (Addendum case above).

Overall, of the 12 siblings reporting glycosuria, two became diabetic by 31st October 1980, while one did so a year later. Of the remaining 253 siblings whose tests remained negative up to 31st October

1980, one developed diabetes by then and one the following year (Cases 2 and 3 above). Thus, as a predictor of development of diabetes within 1.4-2.5 years in these siblings, the finding of glycosuria has a sensitivity of approximately 67%, a specificity of 96%, and a predictive value of approximately 14%, while persistently negative urine tests have a predictive value of 99.6% for non-diabetes ('approximately' because of the very small numbers becoming diabetic; see METHODS VI-3-C-i-a).

In general, parents were unreliable at urine testing and their results are not analysed.

3-B-vi-b. Glycosylated haemoglobin

Glycosylated haemoglobin was less useful than anticipated. Batch mean results from non-diabetic members of the BWFS varied, tending to be higher than expected from laboratory controls (see also METHODS, VI-2-D-iii). Therefore, individual results obtained between February and September 1979 were expressed in terms of standard-deviation units above or below batch mean for samples from non-diabetic, ICA-negative subjects.

Of the 132 non-diabetic fathers and 141 non-diabetic mothers tested, 7 fathers and 6 mothers had a single 'high' result (i.e. > mean +2SD) (Table VIII-46). Two results exceeded mean +3SD. The higher was 13.3%, 3.85 SD units above the mean of 9.40%, and was obtained at the time of diagnosis of diabetes in Case 6 above, being the first HbA_1 test performed in that ICA-positive individual. The other was 12.1%, 3.13 SD units above mean 9.50, after an influenza-like illness in an ICA-negative father whose HbA_1 had been unremarkable (1.08 SD units above mean) three months before. Random blood glucose concentration at the next visit, three months later, was normal (4.0 mmol/l); no glycosuria was detected in 19 tests over two years.

The remaining 11 parents with HbA_1 results between 2 and 3 SD units above batch mean were all ICA-negative, and none showed glycosuria or developed clinical diabetes during the followup period. In all but one, HbA_1 was tested on two or more occasions, but in no case was another result high. Random blood glucose estimations in four were normal.

TABLE VIII-46. GLYCOSYLATED HAEMOGLOBIN IN NON-DIABETIC PARTICIPANTS IN THE BARTS-WINDSOR FAMILY STUDY. Distribution of 'high' results with respect to HLA haplotype concordance with proband.

Subject	PARENTS	SIBLINGS								
		Haplotypes shared with proband:						All sibs	Distribution between groups 2, 1 and 0:	
		MZT	2	2/1	1	1/0	0			
Total no. tested	279	1	41	8	128	8	56	243		
Total no. of tests	491	1	79	19	211	15	89	414	$X^2_{(2)}$	p
Results > mean + 2SD										
Number	13	0	7	1	6	1	0	13	10.4141	<0.01
(%)	(3)		(9)		(3)		(0)	(3)		
Results > mean + SD										
Number	90	0	17	2	16	2	6	39	13.6795	<0.001
(%)	(18)		(21)		(8)		(7)	(9)		

KEY: MZT - monozygotic twin

Among the 243 non-diabetic siblings tested for HbA_1, 13 (3%) of 414 results exceeded batch mean + 2SD (Table VIII-46). In one HLA-identical, CF-ICA-positive sister, HbA_1 was tested on seven occasions between March 1979 and June 1980. The first result was 9.8%, high by comparison with those from ICA-negative subjects in the batch concerned (mean 8.2769, SD 0.7116, n = 13). The second (May 1979) was similar (10.1%), but within normal limits for that batch (mean 9.3077, SD 0.7868, n=26). Thereafter HbA_1 remained normal (7.8 - 9.0%). Random blood glucose and tests for glycosuria also remained normal during follow-up to 31st October 1980. (Interestingly, this subject later became diabetic - Case 3 above.)

One haplo-identical brother had a normal HbA_1 (9.8%) in February, 1979, but three months later, after an upper respiratory infection and a minor accident, the result was 12.2%, more than 3 SD above batch mean. At this time he was ICA-negative. Later, HbA_1 was normal (7.7 - 8.9%, 4 tests), but the conventional ICA test gave weakly positive reactions on five occasions from November 1979 to June 1980. Blood glucose was normal in 1980 (2.6 - 5.2 mmol/l, 4 tests); it was not tested in 1979. A total of 25 urine tests were negative for glucose by October 1980.

The remaining 11 siblings, each with a single high HbA_1 result, were ICA-negative and detected no glycosuria throughout the follow-up period covered here.

The HbA_1 results in non-diabetic subjects in the BWFS show considerable intra-individual variation over periods of several months. Thus, of the 25 parents and siblings with high results, 20 had two or more tests, and yet in no case was more than one result high. This may represent random technical variation and regression to the mean, but it also raises the possibility of temporary, mild hyperglycaemia, perhaps related to environmental influences such as viruses.

For eighteen of the 20 subjects with both high and normal HbA_1 results, at least three months' records of environmental events (colds, etc.) were also available. The numbers of probable viral infections recorded at the time of high HbA_1, and the numbers recorded at the time of one normal HbA_1 result in each of the same group of individuals, were compared. This method is used to eliminate the effects of inter-individual differences in reliability of the record, age, sex, and other constitutional factors. Table VIII-47 summarizes the results. Twelve compared with only 6 at the time of normal HbA_1. This difference is of borderline significance (one-tailed p = 0.0943, Fisher's exact test).

TABLE VIII-47. FREQUENCY OF PROBABLE VIRAL INFECTIONS SHORTLY BEFORE 'HIGH' AND 'NORMAL' HbA_1 RESULTS IN NON-DIABETIC BWFS PARTICIPANTS. Paired data from individuals with at least 3 months' recording of infections before both a 'high' and a 'normal' HbA_1 result. The contingency table gives the numbers of observations categorized by HbA_1 level and number of probable infections recorded since the previous visit (normally three months beforehand).

No. of 'viral events' preceding HbA_1 test	0	1	2	3	4	All
No. of visits:						
High HbA_1	6	10	0	1	1	18
Normal HbA_1	12	4	1	0	1	18

The most striking feature of the HbA_1 observations in non-diabetic siblings is that the frequency of high results shows a highly significant association with HLA-haplotype concordance with probands (Table VIII-46). This applies whether the criterion for a high result is 2 SD or 1 SD above the mean. Even so, when absolute values of HbA_1 in HLA-identical siblings are compared with those in non-identical siblings there is no significant difference overall (mean 9.23 ± SEM 0.12 % versus 9.17 ± 0.10 % respectively), presumably because of both the batch-to-batch variation in mean HbA_1 and the fact that the high results represent only a small minority.

3-B-vii. Islet-cell antibody as a marker of pre-diabetes.

Two questions are considered here: how useful is ICA as a marker of the Type 1 diabetic disease process itself (whether or not diabetes is to develop later); and how valuable is it as a predictor of future clinical diabetes?

A quantitative approach to the first question is not possible, since definite diagnosis of the Type 1 diabetic disease process in metabolically normal individual is not ethically feasible. However, it is reasonable to assume that the underlying pathological process has the same genetic characteristics as does clinically manifest diabetes. In that case, the distribution of ICA in HLA-genotyped non-diabetic BWFS siblings (3-B-iii and Table VIII-45 above) strongly supports the supposition that it does indeed have such a marker role, but with a high 'false-positive' rate (i.e. a low specificity) if the weakest immunofluorescence reactions are counted as positive.

The second question - the value of ICA in predicting future diabetes - can be approached quantitatively since the metabolic abnormality provides an accessible endpoint. Of particular interest is the predictive value of a positive (or negative) ICA result, that is the probability that, given such a result, a non-diabetic sibling or parent will (or will not, respectively) develop diabetes within the 1.4 to 2.5 years of follow-up here. Table VIII-48 shows the results obtained using different criteria for positive ICA. Because of the small number of

TABLE VIII-48. ISLET-CELL ANTIBODIES AND DEVELOPMENT OF DIABETES IN 584 INITIALLY NON-DIABETIC RELATIVES OF TYPE 1 DIABETIC PROBANDS FOLLOWED UP FOR UP TO 2.5 YEARS.

A. Using only the initial test results.

Islet-cell Ab		Diabetes by	Number of subjects		Sensit'y %	Specif'y %	PVpos % (95% cl)	PVneg % (95% cl)
Method	Criterion	30/10/80	ICA +	ICA -				
Conv	w+	+	5	0	100	97.1	23 (16-32)	(100.0) n/a
		-	17	562				
Conv	+	+	3	2	60	98.8	30 (18-46)	99.7 (99.3-99.8)
		-	7	572				
Conv	++	+	3	2	60	99.7	60 (35-81)	99.7 (99.3-99.8)
		-	2	577				
CF	w+	+	4	1	80	98.8	36 (24-51)	99.8 (99.4-99.9)
		-	7	572				
CF	+	+	3	2	60	98.8	30 (18-46)	99.7 (99.3-99.8)
		-	7	572				
CF	++	+	3	2	60	99.5	50 (30-70)	99.7 (99.3-99.8)
		-	3	576				

B. Using the maximum result during (pre-diabetic) follow-up.

Islet-cell Ab		Diabetes by	Number of subjects		Sensit'y %	Specif'y %	PVpos % (95% cl)	PVneg % (95% cl)
Method	Criterion	30/10/80	ICA +	ICA -				
Conv	w+	+	5	0	100	91.5	9 (6-14)	(100.0) n/a
		-	49	530				
Conv	+	+	4	1	80	98.5	31 (20-44)	99.8 (99.4-99.9)
		-	9	574				
Conv	++	+	3	2	60	99.3	43 (26-62)	99.7 (99.3-99.8)
		-	4	575				

NOTES

Statistical significance: in all cases the excess of ICA-positive results in subjects who later develop diabetes is highly significant ($p < 0.0001$ by Fisher's exact test).

PVpos (predictive value of positive ICA result) = probability that an initially non-diabetic family member with a positive ICA result is diabetic by end of follow-up period.

PVneg = probability that an initially non-diabetic family member with a negative ICA result remains non-diabetic during follow-up period.

For method of calculation of sensitivity, specificity and predictive values please see METHODS, VI-3-C-i-a.

n/a = method for confidence limits not applicable when one category is empty.

relatives developing diabetes the figures for sensitivity and PVpos have wide confidence limits, but should give some indication of the best ICA criterion to use for different purposes.

3-C. DISCUSSION

The most striking feature of these prospective observations is the strong evidence they provide for a long latency of Type 1 diabetes. This evidence, and its implications, are discussed further in the published paper (Gorsuch et al, 1981; Appendix 7-2).

Since publication of these findings, further reports from the BWFS (or its successor, the Barts-Oxford Family Study) and from elsewhere in Europe and the USA have provided confirmation, documenting presence of ICA throughout observation periods of up to 8 years before onset of diabetes (Srikanta et al, 1983; Bruining et al, 1985; Tarn et al, 1987).

In no one case have both the time of first appearance of ICA and that of subsequent progression to diabetes been defined. It is quite conceivable that the disease process could begin early in life in all cases, with the net rate of islet B-cell loss determining age at onset of clinical diabetes. Hence age at onset of diabetes could be influenced by any - and probably many - endogenous (genetic, immunological) or exogenous (infective, toxic) factors altering the balance between B-cell destruction and regeneration. This suggestion is in keeping with the minor and subtle interactions between HLA, number affected in sibship, age and season at diagnosis reported in the previous chapter (VIII-2-B-ii-e; see also the Discussions to that chapter and to Appendix 7-2). In the search for prophylaxis against Type 1 diabetes, even factors operative in utero and in early infancy need to be considered.

The use of ICA as a marker of pre-diabetes in first-degree relatives of Type 1 diabetic children is vindicated both by the antibody's observed distribution in non-diabetic HLA-genotyped siblings (3-B-iii) and by its association with diabetes developing during prospective follow-up (3-B-vii). Quantitative estimates of predictive value with respect to future

diabetes are potentially important, since such probability estimates may be combined by Bayesian methods with estimates of genetic risk to give an estimate of overall risk of diabetes.

The predictive values obtained here (3-B-vii) are of limited usefulness, for several reasons. First, their confidence intervals are wide because of the small numbers of individuals becoming diabetic. Secondly, they relate only to development of diabetes within 1.4 to 2.5 years after the first test. With longer follow-up, not only are the confidence intervals shortened as more subjects develop diabetes, but also the estimates for sensitivity and PVpos increase. Thus, extending the period by another year brings in two new cases (Case 3 and the Addendum case of Appendix 7-2), both of whom were strongly positive for ICA, giving the following PVpos values (and 95% confidence limits) for initial ICA results: Conv-ICA, threshold criterion +, 50% (37 - 63%); Conv-ICA ++, 80% (50 - 94%); CF-ICA ++, 83% (57 - 95%). It is conceivable that the great majority of family members having Conv-ICA ++ is destined to develop diabetes at some time in the future.

Thirdly, and perhaps most importantly, classification of ICA immunofluorescence results (w+, +, ++) is subjective and only semi-quantitative; although every effort was made to ensure consistency within the laboratory, there was undoubtedly some variation during the course of this study (VIII-3-B-iii). The figures obtained here are certainly not applicable to other laboratories until there is general standardization of the assay.

Notwithstanding these provisos, some conclusions are still possible from Table VIII-48. The most sensitive screening test to predict diabetes within the next 2.5 years is Conv-ICA with 'w+' as threshold criterion, but there are many false positives and the predictive value of a positive result is low. Serial investigation increases sensitivity and decreases specificity and predictive value still further. Not surprisingly, more strongly positive results are associated with higher probability of diabetes developing. In this context, CF-ICA seems to offer no particular advantage over Conv-ICA, provided the strength of

fluorescence is quantified. Finally, whatever the method used, less than 0.7% of ICA-negative first-degree relatives of Type 1 diabetic children is likely to develop clinical diabetes within 2.5 years.

The findings reported here point to a need for further prospective observation of more numerous families over extended periods, when the number of new cases of diabetes should become sufficient for a life-table risk analysis. Such an extension of the present study is under way. Furthermore, there is a need for a well-standardized, quantitative ICA assay method which is sufficiently economical in human and substrate resources for large-scale use. Only then is reasonably accurate long-term prediction of Type 1 diabetes in such families as these (and possibly in the general population) likely to be achieved.

Despite reservations about HbA_1 measurement under the conditions of this study, the results raise the possibility that a proportion of family members, especially the HLA-identical (and thus genetically susceptible) siblings, may experience transient hyperglycaemia without becoming overtly diabetic at the time (3-B-vi). This has since been confirmed by retrospective glucose assay on stored serum samples (Tarn et al, 1987).

It is interesting that, before the time of high HbA_1 results, the reported frequency of probable viral infections showed a slight tendency to be greater than before normal HbA_1 results in the same individuals (VIII-3-B-vi). The fact that the majority of the individuals concerned were ICA-negative has implications regarding the mechanism of action of susceptibility genes (VIII-5-C).

Finally, the probands of the BWFS fail to support the concept of Type 1B diabetes, in that ICA present more than 5 years after diagnosis of diabetes showed no association with HLA-B8. Perhaps this is because cases with diagnosis of diabetes in adult life, a feature of Type 1B diabetes, are hardly represented at all among these probands.

PART VIII — CHAPTER 4

WHAT IS THE RISK OF TYPE 1 DIABETES IN SIBLINGS OF AN AFFECTED CHILD?

CONTENTS

A Introduction · 299
B Methods: further details · 299
i Potential prevalence of Type 1 diabetes in control pop'n · · 299
a Relative cumulative frequency versus age · · · · · · · · · 299
b Absolute cumulative frequency at reference age (14) · · · 300
c Adjustment to age distribution of BWFS siblings · · · · · 301
ii Relative risk calculation. · · · · · · · · · · · · · · · · · · · 301
C Observations and analyses using local control data · · · · · · · 301
i Relative risk of Type 1 diabetes in HLA-genotyped siblings · 301
D Discussion · 302

4-A. INTRODUCTION

The paper published in 1982 and included here as Appendix 7-3 forms the majority of this chapter. What follows below explains details of methods used, and gives a new risk analysis using local, contemporary control data which were not available at the time the paper was written.

4-B. METHODS: FURTHER DETAILS

4-B-i. Potential prevalence of Type 1 diabetes in control pop'n

4-B-i-a. Relative cumulative frequency versus age

This is covered in detail in the Literature Survey (III-4-A-v). Relative cumulative frequency r (relative to that at age 14 yr), is determined by taking weighted means of the values obtained in several different surveys. The resulting data are plotted graphically in Figure III-2, and tabulated in Appendix 6-2 (third column).

4-B-i-b. Absolute cumulative frequency at reference age (14)

The method used for the published analysis (Appendix 7-3) is described in the Literature Survey (III-4-A-v). The same method is here applied to the new incidence data obtained in East Berkshire by the author (VII-B-i).

The 1981 Census report (Office of Population Censuses and Surveys, 1982) gives total population figures, grouped by age, in the three districts concerned. It also states the proportions of heads of households born in the 'New Commonwealth'. Subtraction of corresponding proportions from the totals gives approximate estimates of the numbers of white Caucasoids. This is necessary since only white Caucasoid diabetics were ascertained to the present study. From the figures given, age correction factors were calculated; most are close to unity, but they do show a 19% deficit of pre-school-age children in Windsor & Maidenhead and a 15% excess of 16-to-19-year-olds in Slough (see Table A in Appendix 6-1).

These age correction factors (f) were then applied to the crude incidence rates of Type 1 diabetes observed in the same population groups, to give an overall age-adjusted rate for each district (see Appendix 6-1, Table B). It is obvious that the result for Bracknell is much too low (see Discussion).

If Bracknell is ignored, the age-adjusted estimate of incidence rate (i) for Type 1 diabetes among children and adolescents in the remainder of East Berkshire comes to 17.75 per 100,000 per year, with 95% confidence limits 11.51 - 23.98 (calculated assuming binomial distribution). Multiplication by 14 gives a cumulative frequency (potential prevalence) of 0.28% at the reference age of 14 years. The expected cumulative frequency, c, at each age a, is calculated as r.i.14, where r is the relative cumulative frequency for age a (see above, and Appendix 6-2, Column 4). The expected number of diabetics of age a in the population concerned is e = c.p/20, where p is the total population aged 0-19 years and the age distribution is temporarily assumed to be uniform (fifth column of Appendix 6-2).

4-B-i-c. Adjustment to age distribution of BWFS siblings

For comparison with the number of BWFS siblings having diabetes, the number in the control population needs to be adjusted for comparable age distribution. If there are n siblings of age a among the 288 in the BWFS (excluding an identical twin), and there are 34 age classes (i.e. ages 0-33 years), then the BWFS-adjusted number of diabetics expected in the control population of age a is d = e.n.34/288. The sum over ages 0 - 33 gives a notional number of diabetic and non-diabetic individuals in the control population.

This is the method used in the published paper (Appendix 7-3) to give control figures based on the Northampton and the Danish data. Applied to the East Berkshire data, it gives a notional control frequency of 251 diabetics among 57060 subjects with the age distribution of the BWFS siblings (Appendix 6-2, last column), or 0.45% with SD 0.03% assuming binomial distribution.

4-B-ii. Relative risk calculation.

Please see the published paper (Appendix 7-3), and elsewhere in this Dissertation (III-6-B-v; VI-3-C-i-a). For calculation of RR, to use the control figures (251/57060) obtained above would give spuriously narrow confidence limits, since they are based on only 21 new cases incident over 2 years. Therefore these figures are reduced proportionately to 21/4914.5.

4-C. OBSERVATIONS AND ANALYSES USING LOCAL CONTROL DATA
(Please see also Results in Appendix 7-3)

4-C-i. Relative risk of Type 1 diabetes in HLA-genotyped siblings.

Table VIII-49 is analogous to Table 3 of the published paper, but gives new, lower relative risk figures since the new control figure for age-adjusted cumulative frequency is higher.

TABLE VIII-49. PREVALENCE OF TYPE 1 DIABETES ON 30:10:1980 IN HLA-TYPED SIBLINGS OF PROBANDS IN THE BWFS, GROUPED BY HAPLOTYPE CONCORDANCE WITH PROBAND, WITH RELATIVE RISK BY COMPARISON WITH CONTROL FREQUENCY.

Haplotypes shared	Monozyg. twin	2	2/1[a]	1	1/0[a]	0	TOTAL
Number of siblings	1 [b]	49	9	149	8	73	289
Number diabetic	0	7	0	6	0	1	14
Prevalence (%)	-	14	-	4	-	1	5
Compared with W:[c] Relative risk (95% conf.lim.)	-	40 (13-121)	-	10 (2-40)	-	NS	12 (7-21)
p (Fisher)		<0.000001		0.000102		0.28	<0.000001

NOTES:

a. In 10 families the two haplotypes in one parent were indistinguishable,so that haplotype concordance is ambiguous. 2/1 = either 2 or one, 1/0 = 1 or 0 haplotype in common with the proband.

b. One identical twin sister of a proband is diabetic, but lives abroad and is not included here.

c. W = 21/4914.5, the potential prevalence of Type 1 diabetes in the general population of Windsor & Maidenhead and Slough districts of East Berkshire (excluding the proportion of 'New Commonwealth' origin), corrected for age distribution of that population and adjusted to age distribution of the BWFS siblings on 30:10:1980. Calculated from the number of new cases (21) observed in these districts over 2 years, and from the known distribution of age at onset of diabetes in published epidemiological surveys (see text).

4-D. DISCUSSION (Please see also Discussion to Appendix 7-3)

The figures obtained here for relative risk of diabetes in HLA concordance groups of siblings of diabetic probands are likely to be more accurate than those published in 1982 (Appendix 7-3), since the control data are more closely applicable. It is true that the age-adjusted mean annual incidence figure for the general population in 1978-80 is higher than anticipated, at 17.8 per 100,000, but in fact this is quite in line with other observations in Northern Europe (see III-4-A-v) and is very similar to that in Scotland at the same time (Grisso et al, 1985).

If this incidence rate is maintained over many years - which need not be the case (III-4-B-iii) - then the potential prevalence in the general population at any age within the range 0 - 30 may be obtained

from Appendix 6-2, column 4. Multiplication by the appropriate relative risk value (from Table VIII-49) then gives an estimate of the probability of diabetes at that age in an HLA-typed sibling of an affected proband.

Some selected results are given in Table VIII-50. It should be emphasized that these estimates of absolute risk are not generally applicable, unlike the relative risks (Table VIII-49). However, the latter should permit similar calculations in different populations of white Caucasoid siblings, given good estimates of age-specific cumulative frequency in the corresponding general populations.

TABLE VIII-50. ABSOLUTE RISK OF TYPE 1 DIABETES IN SIBLINGS GROUPED BY HAPLOTYPE CONCORDANCE WITH PROBAND, IN EAST BERKSHIRE, 1978-80.

AGE (y)	ABSOLUTE RISK (%) Estimate, lower and upper 95% conf. limits		
	H2 SIBLINGS	H1 SIBLINGS	ALL SIBLINGS
	Est. (lcl-ucl)	Est. (lcl-ucl)	Est. (lcl-ucl)
5	2 (1 - 7)	1 (0 - 2)	1 (0 - 1)
10	6 (2 - 18)	1 (0 - 6)	2 (1 - 3)
16	11 (4 - 34)	3 (1 - 11)	3 (2 - 6)
20	14 (4 - 41)	3 (1 - 14)	4 (2 - 7)
30	19 (6 - 57)	5 (1 - 19)	6 (3 - 10)

Several possible sources of error need to be considered. First is that, where ascertainment is incomplete, multiplex families are in general more likely to be ascertained than are simplex families. If this applied to the BWFS, over-estimation of prevalence of diabetes in siblings, and hence of relative risk in siblings, would result. However, there does not appear to be any such ascertainment bias here (see METHODS: VI-1-A-i-e).

A second problem concerns the level of underlying incidence in the general population. Care has been taken to ensure full ascertainment of

new cases; even so, the figure obtained for Bracknell is clearly out of line (Appendix 6-1). Of the 85 diabetic children aged <20 ascertained to the original BWFS while resident in East Berks (see VI-1-A-i), only 16, or 19%, were from Bracknell, which contributes 44% of the white Caucasoid East Berks population in this age-group. One of these Bracknell families was multiplex (in fact triplex) at entry. Possibly there is indeed a low incidence of diabetes in Bracknell, which would itself be interesting. More likely is under-ascertainment, because of referral to hospitals in neighbouring health districts (Frimley, and the Reading group; see map, Figure VI-1). In either case it seems appropriate to exclude Bracknell from the population incidence calculations. Since only two Bracknell siblings were diabetic at entry, and no more became diabetic during follow-up, their inclusion in relative risk calculations should not bias results appreciably.

A third possible source of error is that Type 1 diabetes does not necessarily have an acute clinical onset (see previous chapter), and one of the siblings classified as developing the disease, Case 3 of Appendix 7-2, did not require insulin until much later. If she is excluded then the relative risks obtained are reduced by a factor of about 1/14 (7%). This is a small error in comparison with the confidence limits.

It is reassuring that the estimates of absolute risk (Table VIII-50) are very similar to those produced indirectly for age 20 years in Denmark by Platz et al (1981) (12% in H2 sibs, 4% in H1 and 1% in H0), and those observed empirically for siblings irrespective of HLA at age 20 in Denmark (5.1%; Degnbol & Green) and Canada (Wagener et al, 1982), while the corresponding value of 5.6% derived by Gamble (1980a) for UK siblings aged 16 is well within the 95% confidence limits of our estimate.

It should be noted that the estimates given here relate to haplotype concordance of siblings and probands as defined at the HLA-B locus. Since one 'H0' diabetic sibling in the BWFS had a B-DR recombinant haplotype (VIII-1-B-v-b), use of DR-locus concordance would give a higher RR value for H1 siblings.

It is not possible to combine the available ICA and HLA genotype results in calculating risk of diabetes in siblings, since the number of individuals developing diabetes while under prospective observation by the author was too small. However, with longer follow-up of more families a life-table method of risk calculation becomes feasible, and is preferable to that used here. This is being undertaken by Dr. Anne Tarn. Meanwhile, the results described here carry obvious implications for genetic counselling.

PART VIII - CHAPTER 5

AETIOLOGICAL MECHANISMS: THE NATURE OF HLA-LINKED SUSCEPTIBILITY AND ITS INTERACTION WITH OTHER FACTORS IN PATHOGENESIS

CONTENTS

A Introduction · 306
B Observations and analyses · · · · · · · · · · · · · · · · · · · 307
i Immune response and HLA-linked susceptibility to diabetes 307
a Poliovirus antibody titres after previous vaccination. 307
b Rubella antibody titres · · · · · · · · · · · · · · · · · 312
c Organ-specific and antinuclear antibodies · · · · · · · 313
d Frequency of probable viral infections · · · · · · · · · 314
ii Islet B-cell function and HLA-linked susceptibility to diabetes · 314
C Discussion · 314

5-A. INTRODUCTION

This chapter explores the nature of HLA-linked susceptibility to Type 1 diabetes, by investigating HLA-genotyped siblings of affected probands for evidence of immune response genes linked to susceptibility (Literature Survey, III-11). The particular measures of immune response considered are antibody levels after previous polio vaccination (poliovirus being a picornavirus - see III-8-B), rubella antibody titres, frequency of symptomatic infections likely to be viral in origin, and frequency of antinuclear antibodies and of thyroid/gastric parietal-cell antibodies as indicators of organ-specific immune propensity in general. The latter results have been published in full (Appendix 7-1) and the remainder in abstract form (Gorsuch, Gamble & Ingram, 1985; Gorsuch et al, 1982a).

In addition, the question of genetic impairment of islet B-cell function is addressed superficially. More detailed work on this has been performed by colleagues.

5-B. OBSERVATIONS AND ANALYSES

5-B-i. Immune response and HLA-linked susceptibility to diabetes

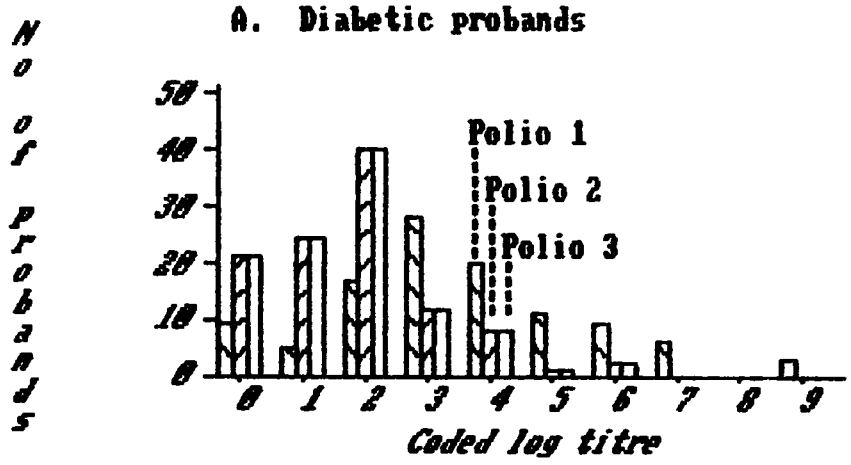
5-B-i-a. Poliovirus antibody titres after previous vaccination.

Neutralizing antibody titres to Poliovirus types 1, 2 and 3 (see Methods, VI-2-C) were estimated in the first 132 families to enter the BWFS. Results are available for 108 probands and 197 non-diabetic siblings reported by parents to have received polio vaccine. (A further four probands and 9 siblings were reported not to have been vaccinated, while vaccination status was unknown in 14 and 23 respectively, and four families are excluded because of expaternity of the proband). Unequivocal HLA genotypes could be determined for 184 of the siblings. Titres weaker than 1/8 are by convention counted as negative. The samples tested are those obtained on entry to the study.

Among the vaccinated probands, log-transformed reciprocal titres show skew distributions (Figure VIII-9-A), with no significant effect of sex (Mann-Whitney 'U' test), nor any significant correlation with either age at the time of testing or age at diagnosis of diabetes (Spearman rank correlation). No significant association was detected between polio titres and the HLA-B phenotypes known to be associated with Type 1 diabetes (B8, Bw62, B8,w62) (Mann-Whitney 'U' test; Fisher exact test, treating polio titre as a dichotomous variable).

Polio titres in vaccinated, non-diabetic siblings did not differ significantly from those in probands (Figure VIII-9-B), and again were independent of sex. Polio 1 and Polio 2 titres in the siblings showed weak negative correlations with age (respectively, $r_s = -0.159$, $p < 0.05$; $r_s = -0.181$, $p < 0.001$).

An interesting and potentially important finding is that all of the HLA-identical (i.e. probably susceptible) siblings were positive for Polio 1 and 2 antibodies. In contrast, a proportion of the H1 and H0 siblings were negative ('non-responders'). Thus, 17% ofH0 siblings were



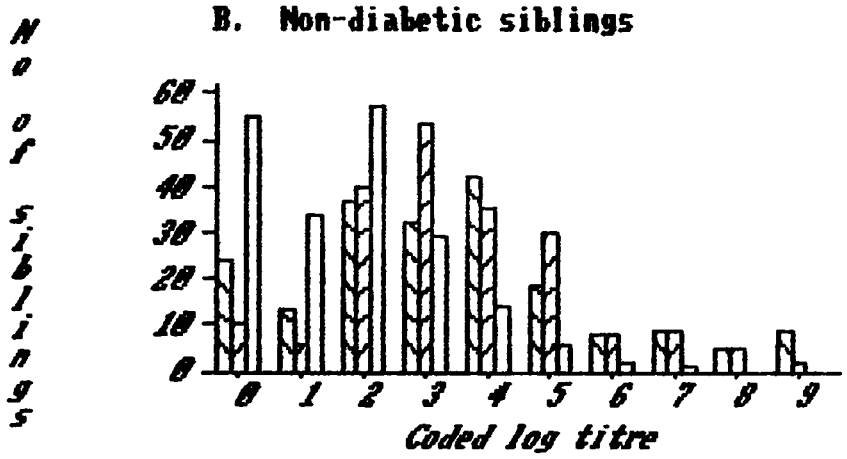
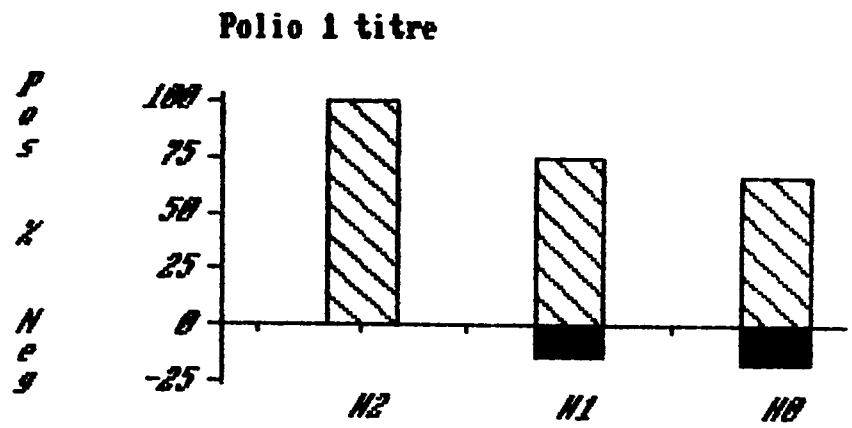


FIGURE VIII-9. POLIO TITRES IN PREVIOUSLY-VACCINATED CHILDREN IN THE BWFS. Coded titres: 0 = weaker than 1/8, 1 = 1/8, 2 = 1/16, 3 = 1/32 and so on.

non-responders to Polio 1 ($p = 0.016$, 2x2 comparison with H2 siblings), and the proportions of non-responders in the three groups of siblings show a significant linear trend ($\chi^2 = 5.29$, 1 d.f., $p < 0.025$) (Figure VIII-10). Similar although individually non-significant trends are apparent for Polio 2 and 3 titres.



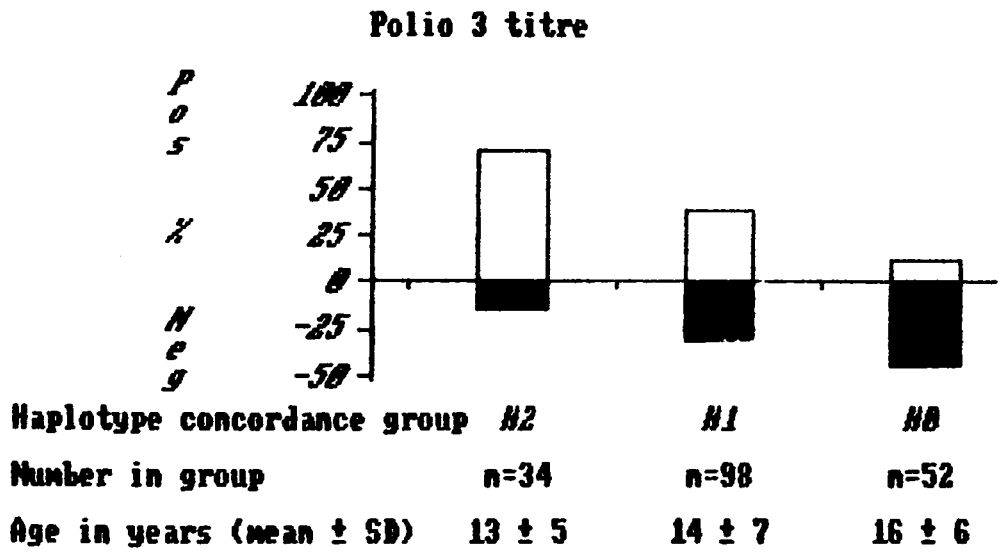


FIGURE VIII-10. PROPORTIONS (%) OF "RESPONDERS" AND "NON-RESPONDERS" TO POLIOVIRUS TYPES 1 (top), 2 (middle) AND 3 (bottom) AMONG VACCINATED BWFS SIBLINGS GROUPED BY HLA HAPLOTYPE CONCORDANCE WITH PROBANDS. The bars represent the proportions (%) of siblings positive (above zero line) or negative (below line) for the antibody concerned. A positive result is defined as a neutralization titre of 1/8 or stronger.

As an index of overall responsiveness to polio vaccine, the mean of the log reciprocal titres to Polio 1, 2 and 3 was computed for each individual. Here, too, the proportion of non-responders shows a significant linear trend, increasing as haplotype concordance with probands decreases ($p < 0.05$) (Figure VIII-11).

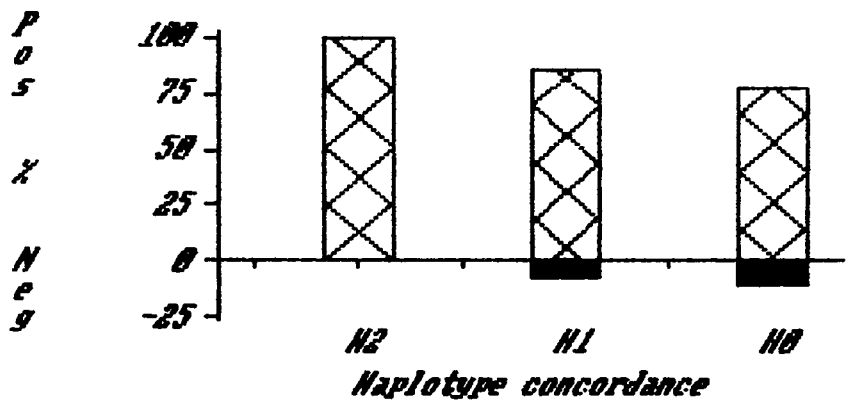


FIGURE VIII-11. MEAN OF LOG TITRES TO POLIOVIRUS TYPES 1, 2 AND 3, IN VACCINATED, NON-DIABETIC BWFS SIBLINGS GROUPED BY HLA-HAPLOTYPE CONCORDANCE WITH PROBANDS. The bars represent the proportions (%) of siblings positive (above zero line) or negative (below line). A positive result is defined as a mean coded log titre of 1.0 or more, representing a geometric mean neutralization titre of 1/8 or stronger. Total numbers are given in Figure VIII-11.

Figure VIII-12 shows the frequency distributions of titres to Polio 1 among the three groups of siblings. Non-parametric tests show no significant quantitative differences in titres between the groups. Non-responders do, however, appear as a partly-distinct subgroup, which is absent from the H2 siblings.

The deficit of non-responders among H2 siblings may be partly explicable by effects of age, since this group was significantly younger at entry to the study (when samples were taken for viral serology) than

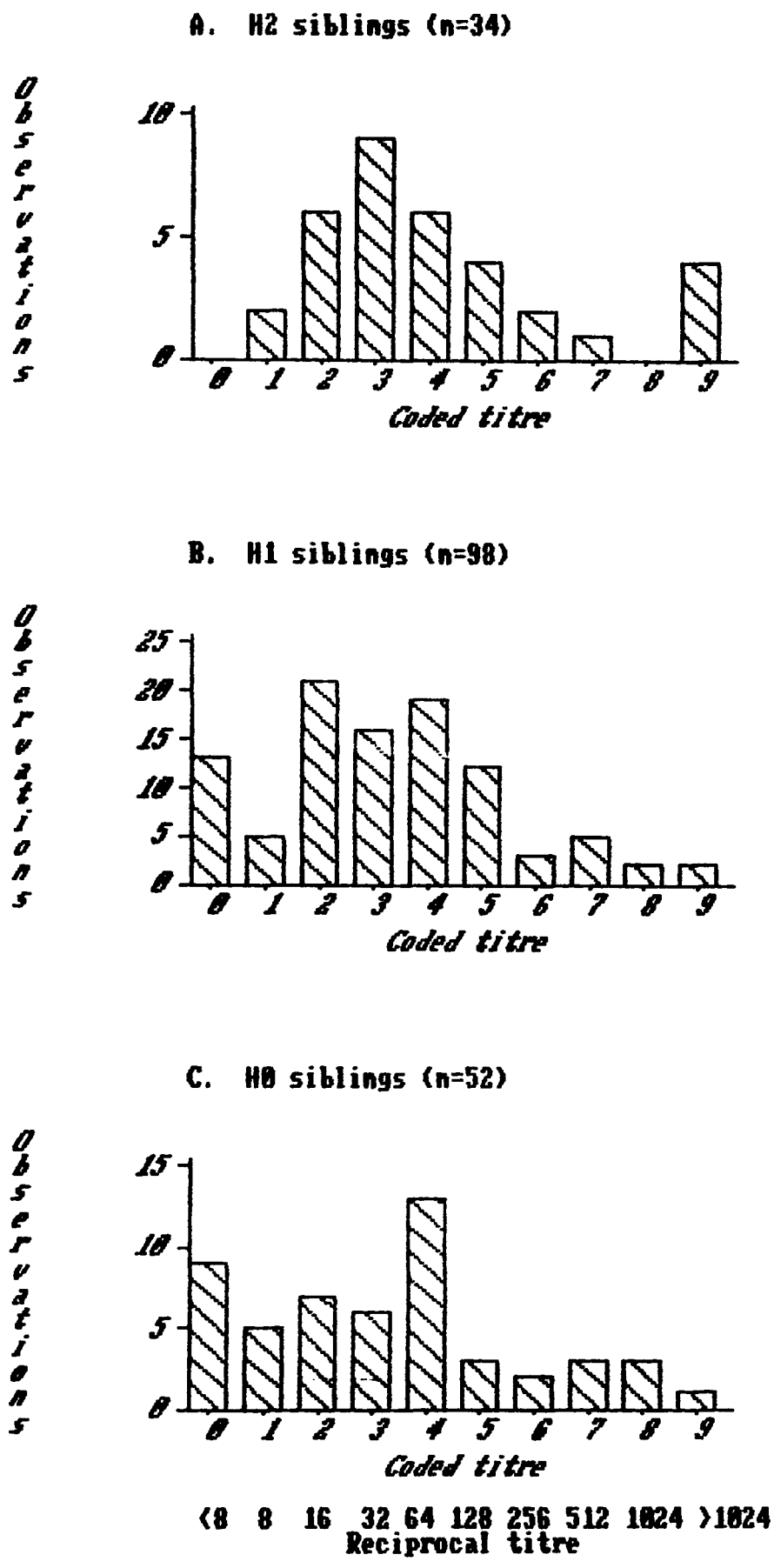


FIGURE VIII-12. POLIOVIRUS TYPE 1 TITRES IN VACCINATED NON-DIABETIC BWFS SIBLINGS, GROUPED BY HLA HAPLOTYPE CONCORDANCE WITH PROBANDS. Each bar represents the number of results at the given titre. The vertical scale varies because the numbers in the three groups differ. Note the numbers with titre weaker than 1/8 (coded 0).

were the H1 and H0 siblings (VIII-1-B-v-c; Table VIII-12). However, 64% of the H2 sibs, whose median (range) age was 13 (3-25) years, fall into the age-range of the Polio 1 non-responders, 12 - 29 (median 19) years. Hence it seems unlikely that age differences are the sole explanation of this interesting observation. For comparison, the age of siblings with titre 1/8 was 18 (9 - 26), and that of siblings with stronger titres was 14 (1 - 32) years.

5-B-i-b. Rubella antibody titres

Rubella titres in non-diabetic siblings are bimodally distributed. They are independent of sex, although a substantial number of older girls had been vaccinated against Rubella. There is a weak positive correlation with age (r_s = 0.205, $p < 0.01$).

The proportion negative for Rubella antibody (22%), although somewhat biased towards the H2 group, is not significantly so (Figure VIII-13; linear trend analysis, 1 d.f.: χ^2 = 2.52). The non-significant trend apparent in this Figure is probably an effect of age.

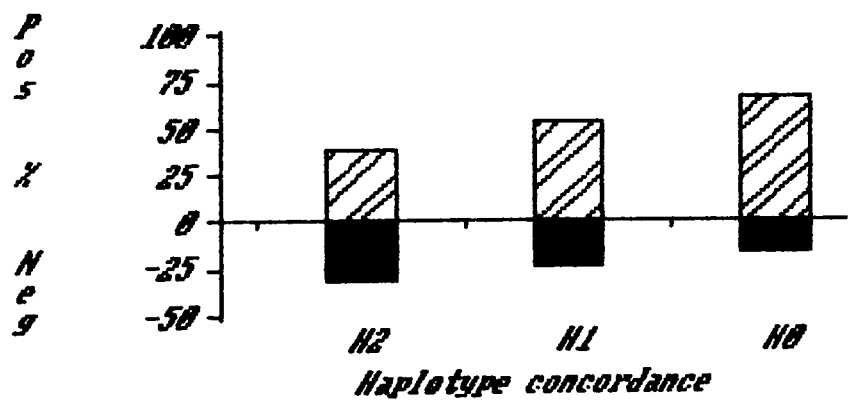


FIGURE VIII-13. RUBELLA TITRES IN BWFS SIBLINGS GROUPED BY HLA-HAPLOTYPE CONCORDANCE WITH PROBANDS. Format as in Figure VIII-11. Titres of 1/19 or stronger are defined as positive.

5-B-1-c. Organ-specific and antinuclear antibodies

Please see the paper published in 1980 (Appendix 7-1). The frequency of thyrogastric antibodies in siblings is shown to be independent of haplotype concordance with diabetic probands, and of whether or not an HLA-A1-B8 haplotype possessed by the sibling is also present in the proband.

Similarly, haplotype concordance of siblings with probands has no effect on frequency of antinuclear antibody (ANA) (Table VIII-51). This holds true even when only the more strongly fluorescent reactions are counted positive (data not shown).

TABLE VIII-51. OCCURRENCE OF ANTINUCLEAR ANTIBODIES (ANA) IN HLA-TYPED SIBLINGS OF PROBANDS IN THE BWFS, GROUPED BY HAPLOTYPE CONCORDANCE WITH PROBAND. A subject is counted positive if ANA is detected at any time during follow-up.

Haplotypes shared	2	1	0	Significance (χ^2 test)
Total number	49	149	72	
Proportion (%) positive for ANA:				
Diffuse pattern	10	5	6	NS
Speckled	17	16	17	NS
Membranous	6	10	10	NS

There is a weak but significant positive association between ANA and thyroid antibody (microsomal and/or thyroglobulin) in parents (Table VIII-52). Similar but individually non-significant trends exist in other groups of subjects of either sex, and when data from all these groups are combined by the method of Woolf and Haldane (1955) the overall association is significant ($p<0.02$). With only 38 siblings and 20 parents positive for ICA, the apparent association between this antibody and ANA is not significant.

TABLE VIII-52. ANTINUCLEAR ANTIBODIES (ANA) IN RELATION TO THYROID AND ISLET-CELL ANTIBODIES IN THE BWFS

Subjects	Thyroid Ab +	Thyroid Ab −	Significance	Islet cell Ab +	Islet cell Ab −	Significance
Proportion (%) positive for ANA:						
Probands	19	16	NS	(not applicable)		
Siblings	34	22	NS	29	22	NS
Parents	41	26	2p = 0.017	50	30	NS

Notes: Probands are excluded from analysis of ANA versus ICA because ICA prevalence depends on duration of diabetes.
Fisher's exact test is used to test for significant associations.

5-B-i-d. Frequency of probable viral infections

(see also Chapter VIII-3-B-iv).

No association was found between frequency of probable viral infections in siblings and haplotype concordance with probands (Kruskal-Wallis non-parametric test; H = 0.1460, 2 d.f.)

5-B-ii. Islet B-cell function and HLA-linked susceptibility to diabetes

Mention has already been made of HbA_1 levels in non-diabetic BWFS family members, and the excess of high levels among H2 siblings (VIII-3-B-vi). This is considered further in the Discussion below.

5-C. DISCUSSION

The observation that non-responders to Polio vaccination are under-represented among children who are genetically susceptible to diabetes is intriguing. Confirmation is of course required in view of the possibility of bias due to age differences. However, it is tempting to speculate that enhanced immune response to picornaviruses might somehow be a factor in genetic susceptibility to Type 1 diabetes.

More directly relevant would be responses to known Coxsackievirus infection, but Coxsackie titres are notoriously liable to non-specific fluctuations (Gamble, personal communication), quite apart from the practical impossibility of knowing whether, and when, an individual might last have been exposed to the wild virus. For what it is worth, Coxsackie titres do not in fact show any quantitative difference between the HLA-haplotype concordance groups of siblings studied here, nor any between diabetic and non-diabetic children.

Further investigation of immune response to picornaviruses is certainly indicated in HLA-genotyped diabetic families. One possible approach would be to give polio vaccine and to assess prospectively the humoral, cell-mediated and lymphokine responses at specific times thereafter.

Rubella virus is of interest because of the high incidence of diabetes after congenital infection (III-8-B-ii-c). The lack of any striking finding here does not exclude a genetic factor in such cases. Again, prospective study of response to vaccination would be interesting. Nevertheless, it is unlikely that any major, HLA-linked diabetes susceptibility gene has a substantial effect on humoral immune response to Rubella, nor indeed on susceptibility to symptomatic viral infections in general (B-i-c).

Thyrogastric antibodies are known to be associated with the HLA-B8 phenotype in the BWFS, at least in probands, and are also more frequent in Type 1 diabetic patients and their relatives than in controls (III-7-B; VIII-1-C-iv). The association with diabetes also applies to antinuclear antibodies (III-7-B). Hence it is reasonable to postulate that an HLA-B8-associated gene enhancing organ-specific, and perhaps antinuclear, autoimmune responsiveness might contribute to genetic susceptibility to Type 1 diabetes. However, this hypothesis is no longer tenable in the light of the findings summarized above (5-B-i-c). Nor, as reported above, do diabetic probands with thyrogastric antibodies show any differences in age or season at onset of symptomatic diabetes, at least within the age-range covered by the BWFS (VIII-2-B-ii-e).

It remains likely that genes controlling autoantibody expression contribute to susceptibility to diabetes, and the association between ANA and thyroid antibodies (5-B-i-c) is compatible with this idea. In that case the genes concerned are not in the HLA region. Indeed, Field et al (1984) present some evidence suggesting interaction between HLA-linked genes and immunoglubulin heavy-chain (Gm) genes in susceptibility to Type 1 diabetes.

The possibility that transient elevations in HbA_1 may follow periods of excess frequency of minor illnesses attributable to viruses (see VIII-3-B-vi-b) invites speculation that individuals with HLA-linked diabetes susceptibility genes may suffer islet B-cell damage in association with certain viral infections, but that this damage is reversible in the absence of some other factor of which ICA is a marker.

In conclusion, these investigations have not yet shown any distinct immunological or metabolic abnormality to explain the mechanism by which genes predispose to islet B-cell destruction. They do, however, support the notion that multiple genes with different effects on immune responses may by their interaction underlie susceptibility to Type 1 diabetes.

PART IX

SUMMARY OF FINDINGS AND CONCLUSIONS

CONTENTS

A Subjects · 317
B Basic genetic and immunological observations · · · · · · · · · · · · 318
i HLA-Genetics · 318
ii Autoimmunity in families with Type 1 diabetes · · · · · · · · · · · 319
C Mode of inheritance of susceptibility · · · · · · · · · · · · · · · · 320
i Heterogeneity · 320
ii Single-gene hypotheses · 321
iii Multifactorial inheritance · 322
D The pre-diabetic period · 322
i Duration · 322
ii ICA as a marker · 322
iii Environmental events · 323
iv Metabolic abnormalities short of diabetes · · · · · · · · · · · · · 323
E Quantitative risk in siblings of affected children · · · · · · · · · 323
i Risk and HLA haplotype concordance with probands · · · · · · · · · 323
ii Predictive value of ICA · 324
F Nature of genetic susceptibility · · · · · · · · · · · · · · · · · · 324
i Modification of immune response to viruses · · · · · · · · · · · · · 324
ii Modification of autoimmune responses · · · · · · · · · · · · · · · 324
G Concluding remarks · 325
i Aetiology of Type 1 diabetes · 325
ii Prediction of Type 1 diabetes · · · · · · · · · · · · · · · · · · · 326

A. SUBJECTS

One hundred and sixty-six Type 1 diabetic probands, together with 313 parents and 289 siblings, were ascertained to the Barts-Windsor Prospective Family Study (BWFS) between April 1978 and June 1979. A further 30 new cases aged < 20 years were notified by August 1980. Additional data from other studies (125 UK multiplex families; 288 Liverpool cases) were made available to the author for certain analyses. Contact was maintained with all BWFS participants at least until 31st October 1980, with serological follow-up of 155 probands, 271 siblings, and 297 parents for a median of 101 (range 13-133) weeks.

B. BASIC GENETIC AND IMMUNOLOGICAL OBSERVATIONS

B-i. HLA-Genetics

In the families studied here, B8+ haplotypes show non-random segregation not explicable by ascertainment bias, in that they occur in offspring of heterozygous fathers significantly more frequently than do the alternative haplotypes. This implies that the B8 allele, or one in linkage disequilibrium with it, enhances fitness of the gamete or zygote. Furthermore, B8 from either parent segregates preferentially to male offspring (VIII-1-B-iv-e).

Siblings grouped by haplotype concordance with probands differ significantly in age, with an increasing trend from H2 through H1 to H0 (VIII-1-B-v-c). This is not due to ascertainment bias, unlike an apparent deficit in numbers of the H2 group when compared with the 'expected' ratio of 25:50:25% for H2, H1 and H0 groups (VIII-1-B-v-a, VIII-1-B-v-d).

Several recognized first-order linkage disequilibria in the HLA region, and a second-order linkage disequilibrium between B18, Cw5 and Bf-F1, are confirmed (VIII-1-B-ii-b).

The intra-HLA recombination rates observed in the BWFS are comparable with those generally found, and contrast with the high rates reported in diabetic families by Suciu-Foca & Rubinstein. The recombination rate between HLA-B and GLO is 8.2% (VIII-1-B-iii).

The established associations of Type 1 diabetes with HLA-A1, B7 (negative), B8, Bw60 and Bw62 are confirmed. A 'new' association with Bw39, of borderline significance here, is clearly genuine in the light of a recent report from Finland (VIII-B-i-a).

Factor B allelotype Bf-F1 occurs in the BWFS families, with a significant predilection for haplotypes present in diabetic probands ('diabetogenic' haplotypes). Bf-S shows no definite association with diabetes when considered alone, but enhances the association between HLA-

B8 and diabetes (VIII-B-i-c). Complement C2 allotypes are equally frequent in 'diabetogenic' and in 'control' haplotypes.

Glyoxalase (GLO) allotypes show neither association with nor disturbance of Hardy-Weinberg equilibrium in diabetes (VIII-B-i-e).

B-ii. Autoimmunity in families with Type 1 diabetes

First and second-degree relatives of BWFS probands show an excess of insulin-dependent diabetes but none of non-insulin-dependent diabetes when compared with the general population. These relatives also show a significant excess of past or present thyrotoxicosis, which is reported by 5.5% of mothers (VIII-1-C-ii).

Thyroid microsomal and/or thyroglobulin antibodies occur in 16% of male and 28% of female diabetic probands (VIII-1-C-iii), 6% of male and 15% of female non-diabetic siblings, and 21% of male and 39% of female non-diabetic parents (VIII-1-C-iv). As well as being greater than would be expected in the general population, these figures show significant effects of sex and age.

The well-known association of ICA with short duration of Type 1 diabetes is present (VIII-1-C-iii; VIII-3-B-ii). ICA appears to be at least as frequent in diabetic boys as in the girls, and possibly more so (VIII-1-C-iii).

At entry to the BWFS 3.8% of non-diabetic subjects were positive for Conv-ICA and 1.8% for CF-ICA. In contrast to thyroid antibodies, ICA is at least as frequent in siblings as in parents, and there is no sex difference (VIII-1-C-iv).

During prospective follow-up, 12% of the siblings and 7% of the parents were ICA positive at some time. However, most of the positive conversions remained 'w+' or reverted to negative, and there is evidence that laboratory definition of a 'w+' result drifted towards greater sensitivity during the period concerned (VIII-3-B-iii).

ICA results by the conventional and the complement-fixation immunofluorescence techniques ('Conv-ICA', 'CF-ICA') are strongly correlated, but to some extent they give different information (VIII-1-C-v).

C. MODE OF INHERITANCE OF SUSCEPTIBILITY TO TYPE 1 DIABETES

C-i. Heterogeneity

The frequency of HLA-B15 differs markedly between Liverpool and BWFS probands. Inclusion of source of data as a variable in these heterogeneity analyses prevents spurious conclusions (VIII-B-ii-d).

Simplex probands aged > 5 years show the well-recognized seasonal variation in time of diagnosis, but multiplex probands do not (VIII-2-B-ii-e; Figure VIII-4). This suggests that, given HLA-linked susceptibility, some other factor (genetic or not) influences both the probability of clinical diabetes and any seasonal effects on its development.

Otherwise, discrete multivariate analysis shows no significant difference between simplex and multiplex probands in frequency of HLA B8, B15 or B8,15, in sex ratio or age at diagnosis (VIII-2-B-ii-d), or in frequency of thyrogastric antibodies (VIII-2-B-ii-e). Furthermore, Hardy-Weinberg equilibrium between HLA-B 8, 15 and other alleles is equally distorted in simplex and in multiplex probands (VIII-2-B-iii-b; Table VIII-40), excluding the hypothesis that susceptibility is recessive in simplex but not in multiplex families (III-10-C-iii-b).

The haplotype concordance analysis in pairs of diabetic siblings (VIII-2-B-i) provides evidence against heterogeneity between duplex and triplex/tetraplex sibships, in that the gene frequency giving best fit to the recessive model among the latter is virtually identical to that among duplex sibships.

In the BWFS and BDA studies, but not in Liverpool, HLA-B15 is significantly associated with male sex of diabetic children ($p < 0.01$) (VIII-2-B-ii-d; VIII-2-B-ii-e). This curious finding requires further confirmation. None of the analyses performed here supports the claim

that HLA-B15 or B8,15 is particularly associated with onset in early childhood rather than later (VIII-2-B-ii-d).

There is a significant interaction between age at diagnosis, season at diagnosis and HLA-B8 (VIII-2-B-ii-e). However, this may not be relevant to the disease process itself, since the interval between initiation of the disease process and clinical diabetes is probably much longer than a season (see below).

A feature of 'Type 1B' diabetes is onset in adult life, whereas the probands in the BWFS and the UK multiplex study were diagnosed diabetic at a median age of 9 years. Hence the concept of Type 1B diabetes is not precluded by the fact that these probands show no supportive evidence: presence of thyrogastric antibody is independent of age at diagnosis and of other variables directly relevant to diabetes (season at diagnosis and number of affected siblings; VIII-2-B-ii-e), while the frequency of persistent ICA is identical in HLA-B8 positive and B8-negative probands (VIII-3-B-ii).

C-ii. Single-gene hypotheses

Analysis of HLA haplotype concordance within 376 pairs of affected siblings from UK and other sources confirms earlier reports: Mendelian dominant inheritance is unequivocally excluded and, while the pattern of HLA haplotype concordance is itself compatible with the recessive model, the minimum gene frequency for this to be so is unacceptably high (VIII-2-B-i-c; VIII-2-C).

Additional and compelling evidence against recessive inheritance is the excess of B8,15 heterozygotes over that expected under Hardy-Weinberg equilibrium, shown to be statistically significant for the first time here in the BWFS (VIII-2-B-iii-a and -b). This also excludes intermediate inheritance. At least two interacting HLA-linked genes must be involved (III-6-E-iii).

C-iii. Multifactorial inheritance

Further support for multifactorial inheritance of susceptibility includes the minor variation in clinical, epidemiological and immunological features of the disease (heterogeneity - see above), the evidence for a contribution from extra-HLA genes (including, perhaps, an autoimmunity gene), and evidence from the literature that the incidence of the disease in identical twins exceeds by a factor of more than four that in other siblings of affected individuals (VIII-2-C).

Given multifactorial inheritance, a direct contribution by HLA genes to susceptibility is feasible. Two novel approaches here (VIII-2-B-iii-c) fail to demonstrate any such direct effect of HLA-A or -B alleles, but similar studies on HLA-DR3 and DR4 are indicated.

D. THE PRE-DIABETIC PERIOD

Five initially non-diabetic family members developed diabetes during the first 2.5 years of the prospective study.

D-i. Duration

The most striking finding is long latency of Type 1 diabetes; minimum periods of up to 30 months (this in a young child) are established by the positive ICA detected at the first test in every case (VIII-3-B-v).

D-ii. ICA as a marker

The use of ICA as a marker of pre-diabetes in first-degree relatives of Type 1 diabetic children is supported both by its predilection for HLA-identical siblings (VIII-3-B-iii) and by its predictive value for diabetes developing during prospective follow-up (VIII-3-B-vii). However, this depends on the criterion used to define a positive result; ICA reported as positive for the first time during follow-up in this study, and never exceeding the 'w+' grade, does not show these associations (VIII-3-B-vii).

D-iii. Environmental events

Probable viral infections were reported more frequently by sisters than by brothers of the BWFS probands, and particularly by the younger siblings (VIII-3-B-iv). Three of the five relatives who developed diabetes while under prospective observation reported upper respiratory infections during the preceding fortnight (Appendix 7-2).

D-iv. Metabolic abnormalities short of diabetes

Glycosuria was reported by 12 non-diabetic siblings during follow-up. Two of these, and one of the remaining 253 informative siblings whose urine tests remained negative, became diabetic by 31st October 1980 (VIII-3-B-vi). Parents did not test reliably.

Variation of HbA_1 in non-diabetic family members suggests that some of them may experience transient hyperglycaemia, possibly after viral infections, without either becoming overtly diabetic at the time or detecting glycosuria at routine monthly testing (VIII-3-B-vi).

E. QUANTITATIVE RISK IN SIBLINGS OF AFFECTED CHILDREN

E-i. Risk and HLA haplotype concordance with probands

New cases of Type 1 diabetes in the Windsor, Maidenhead and Slough districts of East Berkshire, reported to the author during 1978-1980, give an age-corrected incidence rate in the young white Caucasoid population of 17.75 (95% confidence limits 11.51 - 23.98) per 100,000 per year. This represents a potential prevalence of 0.28% at age 14 years (VIII-4-B-i-a).

Compared with this control population, HLA-(B-locus)-identical and haplo-identical siblings of affected probands show age-corrected relative risks for Type 1 diabetes (RR) of 40 (13-121) and 10 (2-40) respectively, implying that 19% (6%-57%) and 5% (1%-19%) of the two respective groups would be diabetic by age 30 years. The frequency in HO siblings is here indistinguishable from that in controls. For all siblings, irrespective of HLA genotype, the figures are: RR = 12 (7-21), potential prevalence at age 30 = 6% (3%-10%) (VIII-4-C-i; VIII-4-D).

E-ii. Predictive value of ICA

Quantitative estimates of predictive value show that the most sensitive screening test to predict diabetes within the next 2.5 years is Conv-ICA with 'w+' as threshold criterion, but there are many false positives (VIII-3-B-vii). The probability of diabetes developing in a first-degree relative of an affected proband within a median of 2 years after a positive Conv-ICA result is 23% for 'w+', 30% for '+' and 60% for '++', while for CF-ICA the estimates are 36%, 30% and 50% respectively. However, these estimates have wide confidence intervals, and relate only to development of diabetes within 1.4 to 2.5 years after the first test. Moreover, without effective cross-standardization of the assay they are not applicable to other laboratories (VIII-3-B-vii).

In this context CF-ICA seems to offer no particular advantage over Conv-ICA, provided the strength of fluorescence is quantified.

Whatever method and threshold criterion is used for ICA, less than 0.7% of ICA-negative first-degree relatives of Type 1 diabetic children is likely to develop clinical diabetes within 2.5 years (VIII-3-B-vii).

F. NATURE OF GENETIC SUSCEPTIBILITY

F-i. Modification of immune response to viruses

Non-responders to Polio vaccination appear to be under-represented among HLA-identical siblings in the BWFS (VIII-5-B-i-a). If this is confirmed it favours a role for an immune response gene in determining susceptibility to Type 1 diabetes.

F-ii. Modification of autoimmune responses

Thyrogastric and antinuclear antibodies in siblings of Type 1 diabetic probands are distributed independently of 'diabetogenic' HLA haplotypes (VIII-5-B-i-c). This weighs against the idea that an HLA-associated gene enhancing autoimmune responsiveness might contribute to genetic susceptibility to Type 1 diabetes. However, such gene(s) are likely to exist outside the HLA region (VIII-5-C).

G. CONCLUDING REMARKS

G-i. Aetiology of Type 1 diabetes

Genetic susceptibility to the disease is multifactorial, involving both HLA-linked and other genes (see above).

The relatively high population frequency of one or more of these apparently deleterious HLA-linked genes may, we suggest, be maintained by intrauterine immune selection. This could explain several curious observations: relatively late birth-order and hence relative youth of susceptible individuals within sibships, distorted zygotic assortment of parental HLA-B8 alleles (see above), and excess incidence of diabetes in boys over that in girls (VIII-1-B-v-d).

The disease process leading to diabetes begins years beforehand, yet both the evidence of individual cases in the literature and, perhaps, seasonality of clinical onset implicate viruses which may be active in the islets even at the time of diagnosis (Literature Survey, III-8-B-ii-b). These apparently conflicting conclusions may be reconciled by proposing that islet B-cell destruction is a cumulative process, each stage being initiated by a new environmental event, the repeated episodes eventually resulting in clinically significant insulin deficiency. Seasonal factors (viruses?) may also have a non-specific "stress" effect, triggering off metabolic decompensation when islet B-cell function is already compromised.

It is not likely that 'purely viral diabetes' is a distinct subgroup of Type 1 diabetes, since if it were at all common the resulting heterogeneity should be detectable, while even the celebrated case of Yoon et al (1979) with betacytotropic Coxsackie B4 virus in the islets also had evidence of long-standing changes typifying Type 1 diabetes (Literature Survey, III-8-B-ii-b).

The disease process could conceivably begin early in life in all cases, possibly initiated by exposure to virus or toxic chemical

substances (III-8) even, perhaps, while in utero. The net rate of islet B-cell loss could then determine age at onset of clinical diabetes (VIII-3-C).

The observation that transiently high HbA_1 may occur particularly in HLA-identical siblings, in the absence of ICA, and with a tendency to follow periods of excess frequency of minor illnesses (VIII-3-B-vi-b), suggests the following hypothesis when considered in the light of the findings in Chapter VIII-5. Individuals with one type of genetic factor (HLA-linked diabetes susceptibility genes) may suffer islet B-cell damage in association with viral infections, but this damage is normally reversible except in the presence of some other, non-HLA-linked factor (autoimmunity genes?) of which low-titre ICA is a marker. When both factors are present, islet autoimmunity is sufficiently intense (manifested as high-titre ICA) to prevent regeneration of islet B-cells. Cumulative damage then results from repeated viral or other insults.

G-ii. Prediction of Type 1 diabetes

There is no doubt that islet-cell antibody is a major risk factor for development of diabetes in a non-diabetic relative of an affected proband. Accurate quantification of the risk is not possible from the observations reported here, since the follow-up period was too short and the numbers developing diabetes too few. Nevertheless, some approximate estimates have been obtained. It is conceivable that the great majority of family members having Conv-ICA ++ is destined to develop diabetes at some time in the future.

By determination of HLA genotypes, siblings of affected children can be divided into high, medium and low-risk categories, and a reasonable estimate of the risk of diabetes given using the empirical figures obtained here. In the future, when further siblings have become diabetic while under prospective surveillance, it should be possible to calculate combined risk estimates using both ICA and genetic information.

PART X

INDICATIONS FOR FUTURE RESEARCH

CONTENTS

A	HLA genetics	327
B	Genetics of Type 1 diabetes	327
C	Pathogenesis	328
D	Prediction	328

A. HLA GENETICS

The non-random segregation shown by B8+ haplotypes in the BWFS is potentially most important. Confirmation is required in further series of HLA-genotyped families, both healthy and diabetic. Other alleles, including those at HLA Class II loci (DR etc), should be investigated in the same way.

B. GENETICS OF TYPE 1 DIABETES

The odd finding that among two out of three series of diabetic children HLA-B15 is particularly associated with male sex merits examination in another population.

There is in this study enough suggestive evidence of heterogeneity within Type 1 diabetes to justify repeat analyses using HLA-DR rather than -B specificities. The search for evidence whether HLA genes may themselves contribute to susceptibility will continue. The new approaches used here (VIII-2-B-iii-c) should be helpful when applied to HLA-DR.

Further investigation of immune response to picornaviruses is certainly indicated in HLA-genotyped diabetic families. One possible

approach would be to give polio vaccine and to assess prospectively the humoral, cell-mediated and lymphokine responses at specific times thereafter.

C. PATHOGENESIS

In view of the evidence for long latency of Type 1 diabetes, extended prospective studies are indicated in the search for aetiologically important environmental factors. Even factors operative in utero and in early infancy need to be considered.

Since the findings on HbA_1 reported here, retrospective glucose assay on stored serum samples has confirmed the occurrence of transient hyperglycaemia without overt diabetes (Tarn et al, 1987). More detailed prospective work may help to elucidate environmental factors with short-term deleterious effects on insulin secretion.

There is also a need for an even more sensitive marker of active islet B-cell damage, such as an extremely sensitive assay for some specific constituent released into the circulation from the damaged cells, or an imaging method (possibly involving magnetic resonance) capable of demonstrating a rise in free radical activity in the islets.

D. PREDICTION

The findings reported here point to a need for further prospective observation of more numerous families over extended periods, when the number of new cases of diabetes should become sufficient for a life-table risk analysis taking into account both HLA genotypes and ICA. Such an extension of the present study is under way.

Furthermore, there is a need for a well-standardized, quantitative ICA assay method which is sufficiently economical in human and substrate resources for large-scale use. Only then is reasonably accurate long-term prediction of Type 1 diabetes in such families as these (and possibly in the general population) likely to be achieved, and only then will the eventual aim of widespread prophylaxis against the disease become ethically feasible.

PART XI — ACKNOWLEDGMENTS

The late Professor Andrew Cudworth believed that a prospective study in HLA-genotyped families of Type 1 diabetic children would be valuable. In 1977 he met Dr. John Lister, of Windsor, and invited the author to set up the project later known as the Barts-Windsor Family Study, with Dr. Lister's diabetic clinic records as a nucleus. Professor Cudworth remained closely involved and a source of considerable support, stimulation and encouragement until his tragically early death in 1982 after a long and distressing illness. He also provided details from two other studies (the BDA UK Multiplex Family Study and his own Liverpool series) for inclusion in heterogeneity analyses. Dr. Lister likewise offered much encouragement, as well as the opportunity to maintain for several years clinical as well as research involvement with many participants in the study. The author owes a great deal to both.

The following clinicians also encouraged the author to study their patients: Dr. Robin Scott; Dr. Bruce Thompson; Dr. Tessa Westcott; Dr. Kim Cheetham; Dr. David Garrow; Dr. Jean Scott.

HLA typing was performed initially in collaboration with Professor Hilliard Festenstein at the London Hospital, and later at St Bartholomew's Hospital with Dr. Eva Wolf. HLA-D typing by Dr Johnny Sachs helped to clarify some families. We are grateful for the meticulous technical assistance of Miss Bridget Watson, Miss Varina Drummond and others. Thanks are due to Drs. John Markwick and Linda Wells for Bf and C2 typing, respectively, to Dr. Simon Welch who determined GLO types and HbA_1 concentrations, and to Dr. Sami Medbak for insulin and C-peptide assays.

Autoantibody estimations - thousands of them - were undertaken in collaboration with Drs. Gian Franco Bottazzo and Betty Dean at the Middlesex Hospital, with technical assistance from Jean McNally and others. Many stimulating discussions with them and with Professor Deborah Doniach were greatly appreciated.

Dr. Robert Gamble, then director of the Public Health Laboratory, Epsom, is thanked not only for viral serology but also for his helpful comments on part of this thesis.

Dr. David Ingram, Reader in Medical Computing in the Department of Medicine, St. Bartholomew's Hospital Medical College, has given invaluable assistance and advice on many aspects of data handling and statistical analysis, including collaboration over development of the GENOBASE system and generous provision of computing facilities over a long period. The recent and impressive further development of GENOBASE by Mr. William Howard, the advice of statisticians at Barts, and the assistance of the Computer Centre of University College, London in providing the mainframe computing facilities needed for some analyses, are also acknowledged.

Dr. Edwin Gale is thanked, not only for developing and expanding the Barts-Windsor Family Study to form the Barts-Oxford Study since coming to Barts in place of Professor Cudworth, but also for helpful criticism of this dissertation. Drs Kate Spencer, Anne Tarn and Polly Bingley have successively kept the 'sharp end' of the study going since 1980. Prof W Gepts is thanked for his encouragement and interest, particularly in Chapter III-3, as well as for an opportunity to study islet histopathology sections and for hospitality in Brussels.

Financial support for various aspects of the project was gratefully received from the following: the Medical Research Council (Project Grant No. G977/954); the Joint Research Board of St. Bartholomew's Hospital, London; the British Diabetic Association; the Juvenile Diabetes Federation; Miles Laboratories (UK).

Above all we are indebted to the many diabetic young people and their families without whose willing participation this work would not have been possible.

Finally, I must acknowledge the patience and understanding of my wife, family and friends, who have for so long suffered neglect in favour of this work.

PART XII — REFERENCES

ADAMS SF. The seasonal variation in the onset of acute diabetes. Arch Int Med 1926; 37: 861-864

ADI FC. Diabetes mellitus associated with epidemic of infectious hepatitis in Nigeria. Br Med J 1974; i: 183-185

AKERBLOM HK, REUNANEN A, KÄÄR M-L. The incidence of insulin-dependent diabetes mellitus in 0-14-year-old children in Finland in 1970-78. Nordic Council Arct Med Res Rep 1980; 26: 60-66

AKERBLOM HK, KÄÄR M-L, REUNANEN A. The epidemiology of insulin-dependent diabetes mellitus in 0- to 14-year-old children in Finland, with particular reference to North Finland. In: Martin JM, Ehrlich RM, Holland FJ (eds), Etiology and pathogenesis of IDDM. Raven Press, New York. 1981; 201-210

ALEJANDRO R, RABINOVITCH A, SEVERYN W, HAJEK S, MILLER J, MINTZ DH. Expression of DR antigens on human islets. Diabetes 1981; 30 (Supp 1):65 A

ALLAN W. Heredity in diabetes. Ann Int Med 1933; 6: 1272-4

ALPER CA, BOENISCH T, WATSON L. Genetic polymorphism in human glycine-rich beta-glycoprotein. J Exp Med 1972; 135: 68-80

ARNOZANT, VAILLARD. Contribution a l'étude du pancreas du lapin - lesions provoquées par la ligature du canal de Wirsung. Arch de Physiol Norm et Path (Paris) 1884; 3: 287-316

BACH FH. The HLA Class II genes and products: the HLA-D region. Immunology Today 1985; 6:89-94

BAETENS D, MALAISSE-LAGAE F, PERRELET A, ORCI L. Endocrine Pancreas: three-dimensional reconstruction shows two types of islets of Langerhans. Science 1979; 206: 1323-5

BAGDADE JD. Phagocytic and microbicidal function in diabetes mellitus. Acta Endocrinologica (København) 1976; 83 (Supp.205): 27-34

BANTING FG, BEST CH. The internal secretion of the pancreas. J Lab Clin Med 1922; a; 7: 251-266

BANTING FG, BEST CH. Pancreatic extracts. J Lab Clin Med 1922 b; 7: 464-472

BARBONI E, MANOCCHIO I. Alterazione pancreatiche in bovini con diabete mellito post-aftoso. Arch Vet Ital 1962; 13: 477-89

BARBONI E, MANOCCHIO I, ASDRUBALI G. Osservazioni sul diabete mellito dei bovini da afta epizootica sperimentale. Veterinaria Italiana 1966; 17: 339-368

BARBOSA J, NOREEN H, DOE RP, YUNIS EJ. The genetic heterogeneity of diabetes: histocompatibility antigens (HLA) in families with maturity-onset type diabetes of the young (MODY), and juvenile, insulin-dependent diabetes (JID). Diabetes 1976; 25 Supp.1: 336 (A)

BARBOSA J, KING R, NOREEN H, YUNIS EJ. The histocompatibility system in juvenile, insulin-dependent diabetic multiplex kindreds. J Clin Investigation 1977; 60: 989-98

BARBOSA J, CHERN MM, NOREEN H, ANDERSON VE, YUNIS EJ. Analysis of linkage between the major histocompatibility system and juvenile, insulin dependent diabetes in multiplex families. J Clin Invest 1978; a; 62: 492-5

BARBOSA J, NOREEN H, KING R, YUNIS EJ. Genetics of juvenile diabetes. N Engl J Med 1978; b; 298: 462.

BARBOSA J, CHERN MM, REINSMOEN N, NOREEN H, RAMSEY R, GREENBERG L. HLA-Dw antigens in unrelated juvenile, insulin-dependent diabetes. Tissue Antigens 1979; a; 14: 426-36

BARBOSA J, WEITKAMP L, GUTTORMSEN S, JOHNSON S, SZALAPSKI E. Bf in early-onset insulin-dependent diabetes. Lancet 1979 b; ii: 1239-40

BARBOSA J, CHERN MM, ANDERSON E, NOREEN H, JOHNSON S, REINSMOEN N, McCARTY R, KING R, GREENBERG L. Linkage analysis between the major histocompatibility system and insulin dependent diabetes in families with patients in two consecutive generations. J Clin Invest 1980; 65:592-601

BARBOSA J, RICH S. Type 1 (insulin-dependent) diabetes and genetic heterogeneity. Diabetologia 1984; 26:170.

BARKER DJP, PHILLIPS DIW. Current incidence of thyrotoxicosis and past prevalence of goitre in twelve British towns. Lancet 1984; ii: 567-570

BARNETT AH, EFF C, LESLIE RDG, PYKE DA. Diabetes in identical twins: a study of 200 pairs. Diabetologia 1981; 20: 87-93

BARTA L, SIMON S. Role of HLA-B8 and Bw15 antigens in diabetic children. N Engl J Med 1977; 296: 397.

BASTENIE PA. Diabète et maladies autoimmunitaires. Diabète et Metabolisme 1975; 1: 3-6

BAUR MP, DANILOVS JA. Population analysis of HLA-A,B,C,DR and other genetic markers. In: Terasaki PI, Histocompatibility Testing 1980, 955-8. UCLA Tissue Typing Laboratory, Los Angeles 1980

BEARDMORE M, REID JJA. Diabetic children. Br Med J 1966; ii: 1383-4

BEAVEN DW, NELSON DH, RENOLD AE, THORN GW. Diabetes mellitus and Addison's disease. A report on eight patients and a review of 55 cases in the literature. N Engl J Med 1959; 261: 443-454

BELLGRAU D, NAJI A, SILVERS WK, MARKMANN JF, BARKER CF. Spontaneous diabetes in BB rats: evidence for a T cell dependent immune response defect. Diabetologia 1982; 23: 359-64

BELL GI, HORITA S, KARAM JH. A polymorphic locus near the human insulin gene is associated with insulin-dependent diabetes mellitus. Diabetes 1984; 33: 176-183

BENGSCH N, KöBBERLING J, ECKERT G, WILLMS B. HLA-typing in juvenile diabetics with and without positive family history and in families with one and two diabetic siblings. Diabetologia 1978; 15: 447-51

BENGTSSON BO, THOMSON G. Measuring the strength of associations between HLA antigens and diseases. Tissue Antigens 1981; 18: 356-363

BERNAL JE, ELLIS J, HAIGH J. Bf in insulin-dependent diabetes mellitus. Lancet 1979; ii: 961.

BERTRAMS J, SODEMANN P, GRÜNEKLEE D, GRIES AF. Bf in early-onset insulin-dependent diabetes. Lancet 1979; ii: 1240.

BERTRAMS J, BAUR MP, GRÜNEKLEE D, GRIES FA. Age-related association of insulin-dependent diabetes mellitus with BfF1 and the HLA-B18,BfF1 haplotype. Diabetologia 1981; 21: 47-49

BERTRAMS J, HINTZEN U, SCHLICHT V, SCHOEPS S. C4: another marker for Type 1 diabetes. Lancet 1982; i: 41.

BETTERLE C, CARETTO A, TIENGO A, TREVISAN A. Complement-fixing islet-cell antibodies in Type 1 diabetes and in susceptible patients with autoimmune diseases. Lancet 1980; i: 1418-9

BISHOP YMM, FEINBERG SE, HOLLAND PW. Discrete multivariate analysis. Massachusetts Institute of Technology Press, Cambridge, Mass. 1975

BLOOM A, HAYES TM, GAMBLE DR. Register of newly diagnosed diabetic children. Br Med J 1975; iii: 580-3

BODANSKY HJ, DRURY PL, CUDWORTH AG, EVANS DAP. Acetylator phenotypes and Type 1 (insulin-dependent) diabetics with microvascular disease. Diabetes 1981; 30: 907-10

BODANSKY HJ, WOLF E, CUDWORTH AG, DEAN BM, NINEHAM LJ, BOTTAZZO GF, MATTHEWS JA, KURTZ AB, KOHNER EM. Genetic and immunological factors in microvascular disease in Type 1 insulin-dependent diabetes. Diabetes 1982; 31: 70-74

BODMER W, BODMER J, ADLER S, PAYNE R, BIALEK J. Genetics of '4' and 'LA' human leukocyte groups. Ann New York Acad Sci 1966; 129: 473-89

BODMER WF. The HLA system: Introduction. Br Med Bull 1978; 34: 213-6

BODMER WF, BODMER JG. Evolution and function of the HLA system. Br Med Bull 1978; 34: 309-316

BODMER WF, BATCHELOR JR, BODMER JG, FESTENSTEIN H, MORRIS PJ (eds). Histocompatibility Testing 1977. Munksgaard, Copenhagen 1978;

BODMER WF. The HLA system and disease. J Roy Coll Physicians 1980; 14: 43-50

BODMER JG, BODMER WF. Histocompatibility 1984. Immunology Today 1984; 5: 251-254

BOITARD C, DEBRAY-SACHS M, POUPLARD A, ASSAN R, HAMBURGER J. Lymphocytes from diabetics suppress insulin release in vitro. Diabetologia 1981; 21: 41-6

BORNSTEIN J, LAWRENCE RD. Two types of diabetes mellitus, with and without available plasma insulin Br Med J 1951; i; 732.

BOTTAZZO GF, FLORIN-CHRISTENSEN A, DONIACH D. Islet cell antibodies in diabetes mellitus with autoimmune polyendocrine deficiencies. Lancet 1974; ii: 1279-82

BOTTAZZO GF, DONIACH D, POUPLARD A. Humoral autoimmunity in diabetes mellitus. Acta Endocrinologica (Kbh) 1976; (Supp.205): 55-61

BOTTAZZO GF, DONIACH D. Pancreatic autoimmunity and HLA antigens. Lancet 1976; ii: 800.

BOTTAZZO GF, CUDWORTH AG, MOUL DJ, DONIACH D, FESTENSTEIN H. Evidence for a primary autoimmune type of diabetes mellitus. Br Med J 1978; a; ii: 1253-5

BOTTAZZO GF, MANN JI, THOROGOOD M, BAUM JD, DONIACH D. Autoimmunity in juvenile diabetics and their families. Br Med J 1978; b; ii: 165-8

BOTTAZZO GF, DEAN BM, GORSUCH AN, CUDWORTH AG, DONIACH D. Complement-fixing islet cell antibodies in Type 1 diabetes: possible monitors of active beta-cell damage. Lancet 1980; i: 668-72

BOTTAZZO GF. Beta-cell damage in diabetic insulitis: are we approaching a solution? Diabetologia 1984; 26: 241-249

BOTTAZZO G, DEAN BM. Evidence of the expression of Class II (HLA-DR) and increased presentation of Class I (HLA-A,B,C) molecules in pancreatic islets in Type 1 (insulin-dependent) diabetes. Diabetologia 1984; 27:259 A.

BOUCHER DW, NOTKINS AL. Virus-induced diabetes mellitus. I. Hyperglycaemia and hypo-insulinaemia in mice infected with encephalomyocarditis virus. J Exp Med 1973; 137: 1226-1239

BRATLID D A preliminary study on the incidence of juvenile diabetes mellitus in Northern Norway. Nordic Council Arct Med Res Rep 1976; 15: 54-7

BRODY JI, MERLIE K. Metabolic and biosynthetic features of lymphocytes from patients with diabetes mellitus: similarities to lymphocytes in chronic lymphocytic leukaemia. Br J Haematol 1970; 19: 193-201

BROSTOFF J. Critique of present in vitro methods for the detection of cell-mediated immunity. Proc Roy Soc Med 1974; 67: 514-6

BRUINING GJ, MOLENAAR JL, de JONGH BM, AARSEN RSR, VISSER HKA. Prediction of Type 1 diabetes mellitus - a report on three cases. Eur J Pediatr 1985; 143: 175-178

BUSCHARD K, ANDERSEN OO, CHRISTAU B, CHRISTY M, KROMANN H, NERUP J, PLATZ A, SVEJGAARD A, THOMSEN M, BOTTAZZO GF. Islet cell antibodies - a marker of subclinical diabetes? A pathogenetic factor? Diabetologia 1977; 13: 386. (A)

BUSCHARD K, MADSBAD S, RYGAARD J. Passive transfer of diabetes mellitus from man to mouse. Lancet 1978; i: 908-10

BUSCHARD K, RYGAARD J, MADSBAD S. Suppressor cell activity in insulin-dependent Type 1 diabetes mellitus. A prospective study. Diabetologia 1980; 19: 262 (A)

BUSI RB, WELLS LJ, VOLKERS WS, EBELI-STRUJIK AC, KHAN PM. Distribution of glyoxalase-I (GLO) variants in Western Europe and the Indian subcontinent. Hum Genet 1979; 49: 105-113

CALDER EA, IRVINE WJ, DAVIDSON N McD, WU F. T, B and K cells in autoimmune thyroid disease. Clin Exp Immunol 1976; 25: 17-22

CALNAN M, PECKHAM CS. Incidence of insulin-dependent diabetes in the first sixteen years of life. Lancet 1977; i: 589-90

CAMMIDGE PJ. Diabetes mellitus and heredity. Br Med J 1928; ii: 738-41

CAMMIDGE PJ. Heredity as a factor in aetiology of diabetes mellitus. Lancet 1934; i: 393-5

CARPENTER CCJ, SOLOMON N, SILVERBERG SG, BLEDSOE T, NORTHCUTT RC, KLINENBERG JR, BENNETT IL, HARVEY A McG. Schmidt's syndrome (thyroid and adrenal insufficiency): a review of the literature and a report of fifteen new cases including ten instances of coexistent diabetes mellitus. Medicine (Baltimore) 1964; 43: 153-80

CATHELINEAU G, CATHELINEAU L, HORS J, SCHMID M, DAUSSET J. Les groupes HLA dans le diabète à début précoce. Nouv Presse Med 1976; 5: 586.

CATTANEO R, SAIBENE V, POZZA G. Peripheral T-lymphocytes in juvenile-onset diabetes (JOD) and in maturity-onset diabetes (MOD). Diabetes 1976; 25: 223-6

CECIL RL. A study of the pathological anatomy of the pancreas in 90 cases of diabetes mellitus. J Exp Med 1909; 11: 266-90

CEPPELLINI R, CURTONI ES, MATTIUZ PL, MIGGIANO V, SERRA A. Genetics of leukocyte antigens. A family study of segregation and linkage. In: Curtoni ES, Mattiuz PL, Tosi RM, Histocompatibility Testing 1967, pp 149-85. Munksgaard, Copenhagen, 1967;

CHAMPSAUR H, DUSSAIX E, SAMOLYK D, FABRE M, BACH CG, ASSAN R. Diabetes and Coxsackievirus B5 infection. Lancet 1980; i: 251.

CHRISTAU B, KROMANN H, ANDERSEN OO, CHRISTY M, BUSCHARD K, ARNUNG K, KRISTENSEN IH, PEITERSEN B, STEINRUD J, NERUP J. Incidence, seasonal and geographical patterns of juvenile-onset insulin-dependent diabetes mellitus in Denmark. Diabetologia 1977; 13: 281-4

CHRISTY M, GREEN A, CHRISTAU B, KROMANN H, NERUP J. Epidemiologic studies of insulin-dependent diabetes mellitus. Diabetes Care 1979; 2: 127-30

CHRISTY M, GREEN A, CHRISTAU B, KROMANN H, NERUP J, PLATZ P, THOMSEN M, RYDER L, SVEJGAARD A. Studies of the HLA system and insulin-dependent diabetes mellitus. Diabetes Care 1979; 2: 209-14

CLARKE CA, PRICE EVANS DA, HARRIS R, McCONNELL RB, WOODROW JC. Genetics in medicine: a review. Part I. Q J Med 1968; NS 37: 1-61

CLERGET-DARPOUX F, BONAITI-PELLIÉ C. Epistasis effect: an alternative to the hypothesis of linkage disequilibrium in HLA associated diseases. Ann Human Genet 1980; 44: 195-204

COHEN T, NELKEN L, WOLFSOHN H. Juvenile diabetes mellitus in immigrant populations in Israel. Diabetes 1970; 19: 585-90

COLEMAN TJ, TAYLOR KW, GAMBLE DR. The development of diabetes following Coxsackie B virus infection in mice. Diabetologia 1974; 10: 755-759

CONN JW. The prediabetic state in man. Diabetes 1958; 7: 347-357

CRAIGHEAD JE, McLANE MF. Diabetes mellitus: induction in mice by encephalomyocarditis virus. Science 1968; 162: 913-914

CRAIGHEAD JE. Viral diabetes mellitus in man and experimental animals. Am J Med 1981; 70: 127-34

CROSSLEY JR, UPSDELL M. The incidence of juvenile diabetes mellitus in New Zealand. Diabetologia 1980; 18: 29-34

CUDWORTH AG, WOODROW JC. HL-A antigens and diabetes mellitus. Lancet 1974; ii: 1153.

CUDWORTH AG, WOODROW JC. Evidence for HL-A linked genes in 'juvenile' diabetes mellitus. Br Med J 1975; iii: 133-5

CUDWORTH AG, WOODROW JC. Genetic susceptibility in diabetes mellitus: analysis of the HLA association. Br Med J 1976; ii: 846-8

CUDWORTH AG. Genetic susceptibility in Type 1 diabetes mellitus. PhD thesis, University of Liverpool 1977

CUDWORTH AG, GAMBLE DR, WHITE GBB, LENDRUM R, WOODROW JC, BLOOM A. Aetiology of juvenile onset diabetes: a prospective study. Lancet 1977; i: 385-388

CUDWORTH AG, USHER N, WOODROW JC. Factor B phenotypes in 'juvenile-onset' diabetes. Diabetologia 1977; 13: 388 (A)

CUDWORTH AG. Type 1 diabetes. Diabetologia 1978; 14: 281-91

CUDWORTH AG, FESTENSTEIN H. HLA genetic heterogeneity in diabetes mellitus. Br Med Bull 1978; 34: 285-9

CUDWORTH AG, GORSUCH AN, LISTER J, PRITCHARD J. Evidence that Type 1 diabetes and organ-specific autoimmunity have different genetic determinants. In: Waldhaüsl W, Alberti KGM M (eds), 10th Congress of the International Diabetes Federation. Internat Congress Series No. 481, Excerpta Medica, Amsterdam 1979 a; 48-49

CUDWORTH AG, WOLF E, GORSUCH AN, FESTENSTEIN H. A new look at HLA genetics with particular reference to Type 1 diabetes mellitus. Lancet 1979 b; ii: 389-91

CURIE-COHEN M. HLA-antigens and susceptibility to juvenile diabetes: do additive relative risks imply genetic heterogeneity? Tissue Antigens 1981; 17: 136-48

DANIELS CA, BORSOS T, RAPP HJ, SNYDERMAN R, NOTKINS A. Neutralization of sensitized virus by the fourth component of complement. Science 1969; 165:508-509

DARLOW JM, SMITH C, DUNCAN LJP. A statistical and genetical study of diabetes. III. Empiric risks to relatives. Ann Human Genet 1973; 37: 157-174

DEAN BM, McNALLY JM, DONIACH D. IgG subclass distribution in islet-cell and other organ-specific autoantibodies. Diabetologia 1980; 19:268 (A)

DEAN BM, WALKER R, BONE AJ, BAIRD JD, COOKE A. Pre-diabetes in the spontaneously diabetic BB/E rat: lymphocyte subpopulations in the pancreatic infiltrate and expression of rat MHC Class II molecules in endocrine cells. Diabetologia 1985; 28:464-466

DEGNBOL B, GREEN A. Diabetes mellitus among first- and second-degree relatives of early onset diabetics. Ann Hum Genet 1978; 42: 25-47

DELESPESSE G, DUCHATEAU J, BASTENIE PA, LAVAUX JP, COLLETT H, GOVAERTS A. Cell-mediated immunity in diabetes mellitus. Clin Exp Immunol 1974; 18: 461-467

Del PRETE GF, BETTERLE C, PADOVAN D, ERLE G, TOFFOLO A, BERSANI G. Incidence and significance of islet cell autoantibody in different types of diabetes mellitus. Diabetes 1977; 26: 909-15

De MEYER J. Contribution a l'étude de la pathogenie du diabète pancréatique. Arch Internat Physiol 1909; 8; 121-180

De MEYER J. Observations sur les pancréas d'animaux injectés de serum anti-pancréatique et sur les formes de transition acino-insulaires du pancréas de chien. Arch Internat Physiol 1911; 11: 131-53

DE MOERLOOSE P, JEANNET M, BALLY C, RAFFOUX C, POINTEL JP, SIZONENKO P. HLA and DRW antigens in insulin-dependent diabetes. Br Med J 1978; i: 823-4

De MOUZON A, OHAYON E, DUCOS J, HAUPTMANN G. Bf and C4 markers for insulin-dependent diabetes in Basques. Lancet 1979; ii: 1364.

DESCHAMPS I, LESTRADET H, MARCELLI-BARGE A, BENAJAM A, BUSSON M, HORS J, DAUSSET J. Properdin factor B alleles as markers for insulin-dependent diabetes. Lancet 1979; ii: 793.

DESCHAMPS I, LESTRADET H, BONAITI C, SCHMID M, BUSSON M, BENAJAM J, MARCELLI-BARGE A, HORS J. HLA genotype studies in juvenile insulin-dependent diabetes. Diabetologia 1980; 19: 189-93

De VRIES RRP, FAT RFMLA, NIJENHUIS LE, van ROOD JJ. HLA-linked genetic control of host response to Mycobacterium leprae. Lancet 1976; 2: 1328-30

DEWALD G, RITTNER C. Polymorphism of the second component of human complement (C2). Observation of the rare phenotype (C2 2 = C2 B) and data on the location of the C2 locus in the HLA region. Vox Sanguinis 1979; 37: 47-54

Di MARIO U, POZZILLI P, SCARSELLA M, BOLOGNESI A, JAVICOLI M. Depressed lymphocyte transformation to phytohaemagglutinin in normal subjects by intravenous glucose administration. Diabetologia 1977; 13: 390 (A)

DOBERSEN MJ, SCHARFF JE, GINSBERG-FELLNER F, NOTKINS AL. Cytotoxic autoantibodies to beta cells in the serum of patients with insulin-dependent diabetes mellitus. N Engl J Med 1980; 303: 1493-8

DOBERSEN MJ, SCHARFF JE, GINSBERG-FELLNER F, NOTKINS AL. Preferential lysis of beta cells by islet cell surface antibodies. Diabetologia 1981; 21: 264 (A)

DONIACH I, MORGAN AG. Islets of Langerhans in juvenile diabetes mellitus. Clin Endocrinol 1973; 2: 233-48

DONIACH D, BOTTAZZO GF, RUSSELL RCG. Goitrous autoimmune thyroiditis. Clinics Endoc Metab 1979; 8: 63-80

DORCHY H, LEMIERE B, TOUSSAINT D, GAUSSET P. Anticorps anti-cellules des îlots de Langerhans et spécifiques d'organes chez les jeunes diabétiques. Nouv Presse Med 1981; 10: 2795-8

DORNAN J, ALLAN P, NOEL EP, LARSEN B, FARID NR. Properdin Factor B (Bf) allele BfF1 specifies an HLA-B18 diabetogenic haplotype. Diabetes 1980; 29: 423-427

DORNAN J, ALLAN P, NOEL EP, SAMPSON L, FARID NR. Alleles of the second component of complement (C2) in insulin-dependent diabetes mellitus. Hum Hered 1981; 31: 211-213

DULIN WE, SORET MG. Chemically and hormonally induced diabetes. In: Volk BY, Wellmann KF (eds), The Diabetic Pancreas, 425-465. Ballière Tyndall, London. 1977

DULIN WE, GERRITSEN GC, CHANG AY. Experimental and spontaneous diabetes in animals. In: Ellenberg M, Rifkin H (eds.), Diabetes Mellitus, Theory and Practice (3rd edition), 361-408. Medical Exam. Publishing Company, New York. 1983

DUNSWORTH TS, RICH SS, SWANSON J, BARBOSA J. No evidence for linkage between diabetes and the Kidd marker. Diabetes 1982; 31:991-993

DURRUTY P, RUIZ F, GARCIA de los RIOS M. Age at diagnosis and seasonal variations in the onset of insulin dependent diabetes in Chile (Southern Hemisphere) Diabetologia 1979; 17: 357-60

EDWARDS JH. The simulation of Mendelism. Acta Genetica 1960; 10: 63-70

EVERITT BS. The analysis of contingency tables. Chapman & Hall, London 1977;

FABER OK, THOMSEN M, BINDER C, PLATZ P, SVEJGAARD A. HLA antigens in a family with maturity-onset type diabetes mellitus. Acta Endocrinologica 1978; 88: 329-338

FALCONER DS. The inheritance of liability to certain diseases, estimated from the incidence among relatives. Ann Hum Genet 1965; 29: 51-71

FALCONER DS. The inheritance of liability to diseases with variable age of onset, with particular reference to diabetes mellitus. Ann Hum Genet 1967; 31: 1-20

FALCONER DS, DUNCAN LJP, SMITH C. A statistical and genetical study of diabetes: I. Prevalence and morbidity. Ann Hum Genet 1971; 34: 347-69

FARID NR, BARNARD JM, PEPPER B, NOEL EP, KELLY F, DAVIS AJ, HOBEIKA C, MARSHALL WH. The association of HLA with juvenile diabetes mellitus in Newfoundland. Tissue Antigens 1978; 12: 215-22

FELDMAN JL, BECKER MJ, MOUTSOPOULOS H, FYE K, BLACKMAN M, EPSTEIN WV, TALAL N. Antibody dependent cell-mediated cytotoxicity in selected autoimmune diseases. J Clin Invest 1976; 58: 173-9

FESTENSTEIN H, DEMANT P. HLA and H-2: basic immunogenetics, biology and clinical relevance. Arnold, London, 1978;

FIALKOW DJ, ZAVALA CZ, NIELSEN R. Thyroid autoimmunity: increased frequency in relatives of insulin dependent diabetes patients. Ann Int Med 1975; 83: 170-6

FIELD LL, ANDERSON CE, NEISWANGER K, HODGE SE, SPENCE MA, ROTTER JI. Interaction of HLA and immunoglobulin antigens in Type 1 (insulin-dependent) diabetes. Diabetologia 1984; 27: 504-508

FINKELSTEIN S, ZELLER E, WALFORD RL. No relation between HLA and juvenile diabetes. Tissue Antigens 1972; 2: 74-7

FLEEGLER FM, ROGERS KD, DRASH A, ROSENBLOOM AL, TRAVIS LB, COURT JM. Age, sex and season of onset of juvenile diabetes in different geographic areas. Pediatrics 1979; 63: 374-9

FLOOD TM, BRINK SJ, GLEASON RE. Increased incidence of Type 1 diabetes in children of older mothers. Diabetes Care 1982; 5:571-573

FORD, EB. Genetic polymorphism. All Souls Studies, V. Faber, London, 1965

GABRA GS, MITCHELL R. Haemoglobin A1 and diabetes. Br Med J 1981; 282: 149-150

GALLUZZO A, GIORDIANO C, SPARACINO C, BOMPIANI GD. Studies on suppressor cell function in recent-onset juvenile diabetes. Diabetologia 1980; 19: 275 (A)

GAMBLE DR, KINSLEY ML, FITZGERALD MG, BOLTON R, TAYLOR KW. Viral antibodies in diabetes mellitus. Brit Med J 1969 a; iii: 627-630

GAMBLE DR, TAYLOR KW. Seasonal incidence of diabetes mellitus. Br Med J 1969; iii: 631-3

GAMBLE DR, TAYLOR KW, CUMMING H. Coxsackie viruses and diabetes mellitus. Brit Med J 1973; iv: 260-262

GAMBLE DR. An epidemiological study of childhood diabetes affecting two or more siblings. Diabetologia 1980; a; 19: 341-344

GAMBLE DR. Relation of antecedent illness to development of diabetes in children. Brit Med J 1980; b; 281: 99-101

GEPTS W. Pathological anatomy of the pancreas in juvenile diabetes mellitus. Diabetes 1965; 14: 619-33

GEPTS W (chairman) et al. Immune-pathogenetic aspects of diabetes mellitus: general discussion. In: Ortved Andersen O, Deckert T, Nerup J (eds), Immunological Aspects of Diabetes Mellitus. Acta Endocrinologica (Kbh) 1976; (Supp 205):139-144

GEPTS W, De MEY J. Islet cell survival determined by morphology. An immuno-cytochemical study of the islets of Langerhans in juvenile diabetes mellitus. Diabetes 1978; 27 (Supp 1): 251-61

GEPTS W, LeCOMPTE PM. The pancreatic islets in diabetes. Am J Med 1981; 70: 105-15

GIBBS P. Three cases of acute ketotic diabetes mellitus with myocarditis: a common viral origin? Br Med J 1974; iii: 781-783

GINSBERG-FELLNER F, DOBERSEN M, WITT M, NOTKINS A, RUBINSTEIN P, RAYFIELD EJ. HLA antigens, islet cell antibodies and carbohydrate metabolism in siblings of children with insulin-dependent diabetes mellitus. Diabetes 1979; 28:396. (A)

GINSBERG-FELLNER F, DOBERSEN MJ, WITT ME, RAYFIELD EJ, RUBINSTEIN P, NOTKINS AL. HLA antigens, cytoplasmic islet cell antibodies, and carbohydrate tolerance in families of children with insulin-dependent diabetes mellitus. Diabetes 1982; 31: 292-298

GLADISCH R, HOFMANN W, WALDHERR R. Myokarditis und Insulitis nach Coxsackie-virus Infekt. Zeitschrift Kardiol 1976; 65: 837-849

GOLDSTEIN DE, DRASH A, GIBBS J, BLIZZARD RM. Diabetes mellitus: the incidence of circulating antibodies against thyroid, gastric and adrenal tissue. J Pediatr 1970; 77: 304-6

GORSUCH AN, GAMBLE DR, LISTER J, FESTENSTEIN H. The aetiology of Type 1 diabetes - a prospective family study. In: Waldhaüsl W, Alberti KGMM (eds.), 10th congress of the International Diabetes Federation. Internat Congress Ser No 481, Excerpta Medica, Amsterdam 1979; 76-77

GORSUCH AN, DEAN BM, BOTTAZZO GF, LISTER J, CUDWORTH AG. Evidence that Type 1 diabetes and thyrogastric autoimmunity have different genetic determinants. Br Med J 1980; 280: 145-7

GORSUCH AN, SPENCER KM, LISTER J, McNALLY JM, DEAN BM, BOTTAZZO GF, CUDWORTH AG. Evidence for a long pre-diabetic period in Type 1 (insulin-dependent) diabetes mellitus. Lancet 1981; ii: 1363-5

GORSUCH AN, DEAN BM, LISTER J, BOTTAZZO GF, CUDWORTH AG. Antinuclear antibodies in HLA-genotyped families of Type 1 (insulin-dependent) diabetic children. Diabetologia 1982a; 23: 170.

GORSUCH AN, SPENCER KM, LISTER J, WOLF E, BOTTAZZO GF, CUDWORTH AG. Can future Type 1 diabetes be predicted? A study in families of affected children. Diabetes 1982b; 31: 862-866

GORSUCH AN, SPENCER KMM, WOLF E, CUDWORTH AG. HLA and family studies. In: Köbberling J, Tattersall R (eds), The genetics of diabetes mellitus. Proc Serono Symposia, Vol 47 Academic Press, London. 1982c 43-53

GORSUCH AN, GAMBLE DR, INGRAM D. HLA-linked susceptibility to Type 1 diabetes (IDDM) and polio-virus antibody titres after polio vaccination. Diab Res Clin Practice 1985; Supp 1; 199-200

GORWITZ K, HOWEN GG, THOMPSON T. Prevalence of diabetes in Michigan school-age children. Diabetes 1976; 25: 122-7

GRAY RS, CLARKE BF. Primary autoimmune diabetes mellitus. Br Med J 1978; ii: 1715.

GRAY RS, DUNCAN LJP, CLARKE BF. Seasonal onset of insulin dependent diabetes in relation to sex and age at onset. Diabetologia 1979; 17: 29-32

GREEN JR, WOODROW JC. Sibling method for detecting HLA-linked genes in disease. Tissue Antigens 1977; 9: 31-5

GREEN A, HAUGE M, HOLM NV, RASCH LL. Epidemiological studies of diabetes mellitus in Denmark: II - A prevalence study based on insulin prescriptions. Diabetologia 1981; 20:468-470

GREENBERG DA, ANDERSON CE. The search for heterogeneity in insulin-dependent diabetes mellitus: evidence for familial and non-familial forms. Am J Med Genet 1983; 14: 487-499

GUNDERSEN E. Is diabetes of infectious origin? J Infect Dis 1927; 41: 197-202

HALDANE JBS. The estimation and significance of the logarithm of a ratio of frequencies. Ann Human Genet 1955; 20: 309-311

HAMILTON DV, MUNDIA S, LISTER J. Mode of presentation of juvenile diabetes. Br Med J 1976; ii: 211-212

HANN S, KAYE R, FALKNER B. Subpopulations of peripheral lymphocytes in juvenile diabetes. Diabetes 1976; 25: 101-3

HANSEN S. Über die Vererbung des Diabetes mellitus. Acta Med Scand 1925; 62: 85-100

HANSEN AE, GÜRTLER H. Distorted HLA segregation or biased ascertainment? Lancet 1980; i: 41.

HARKNESS J. Prevalence of glycosuria and diabetes: a comprehensive survey in an urban community. Br Med J 1962; ii: 1503-7

HARRIS H. The familial distribution of diabetes mellitus: a study of the relatives of 1241 diabetic propositi. Ann Eugen (London) 1950; 15: 95-119

HAUPTMANN G. C4-deficiency in early onset insulin dependent diabetes: a hypothesis. Lancet 1980; 1: 1034.

HAYLES AB, KENNEDY RLJ, BEAHRS OH, WOOLNER LB. Exophthalmic goiter in children. J Clin Endocrinol Metab 1959; 19: 138-151

HEDING LG. Radioimmunological determination of human C-peptide in serum. Diabetologia 1975; 11: 541-548

HEDLEY AJ, JONES RB, GALE EAM, TATTERSALL RB. Prevalence of insulin-treated diabetes mellitus. Br Med J 1982; 285:509.

HELGASON T, JONASSON MR. Evidence for a food additive as a cause of ketosis-prone diabetes. Lancet 1981; ii. 717-20

HELGASON T, EWEN SWB, ROSS IS, STOWERS JM. Diabetes produced in mice by smoked/cured mutton. Lancet 1982; ii: 1017-1022

HELLESTROM C, ANDERSSON A, GUNNARSSON R. Regeneration of islet cells. In: Ortved Andersen O, Deckert T, Nerup J (eds), Immunological Aspects of Diabetes Mellitus. Acta Endocrinologica (Kbh) 1976; (Supp 205):145-158

HELMKE K, BROKHAUS R, OTTEN A, WILLEMS A, MASER E, FEDERLIN K. Islet cell antibodies in juvenile diabetes in children with mumps infection. Diabetologia 1980; 19: 282 (A)

HIERHOLZER JC, FARRIS WA. Followup of children infected in a Coxsackie B-3 and B-4 outbreak. J Infect Dis 1974; 129: 741-746

HODGE SE, ANDERSON CE, NEISWANGER K, FIELD LL, SPENCE MA, SPARKES RS, SPARKES M, CHRIST M, TERASAKI PI, RIMOIN DL, ROTTER JI. Close genetic linkage between diabetes mellitus and Kidd blood group. Lancet 1981; ii: 893-895

HODGE SE, ANDERSON CE, NEISWANGER K, RUBIN R, SPARKES RS, SPARKES MC, CRIST M, SPENCE MA, TERASAKI PI, RIMOIN DL, ROTTER JI. Association studies between Type 1 (insulin-dependent) diabetes and 27 genetic markers: lack of association between Type 1 diabetes and Kidd blood group. Diabetologia 1983; 25: 343-347

HOLLANDER PH, ASPLIN CM, KNIAZ D, HANSEN JA, PALMER JP. Beta-cell dysfunction in non diabetic HLA identical siblings of insulin-dependent diabetics. Diabetes 1982; 31:149-153

HOWARD JC. The major histocompatibility complex of the rat: a partial review. Metabolism 1983; 32 (7 Supp 1): 41-50

HUANG S-W, MACLAREN NK. Insulin-dependent diabetes: a disease of autoaggression. Science 1976; 192: 64-66

HUANG S-W, MACLAREN NK. Antibodies to nucleic acids in juvenile-onset diabetes. Diabetes 1978; 27: 1105-1111

HUANG S-W, HAEDT LH, RICH S, BARBOSA J. Prevalence of antibodies to nucleic acids in insulin-dependent diabetics and their relatives. Diabetes 1981; 30: 873-874

HUGH-JONES P. Diabetes in Jamaica. Lancet 1955; ii: 891-897

ILONEN J, MUSTONEN A, AKERBLOM H, HUTTUNEN N-P. Complement-fixing islet-cell antibodies before and after onset of insulin-dependent diabetes. Lancet 1980; ii: 805.

IRVINE WJ, CLARKE BF, SCARTH L, CULLEN DR, DUNCAN LJ. Thyroid and gastric autoimmunity in patients with diabetes mellitus. Lancet 1970; ii: 163-8

IRVINE WJ, BARNES EW. Addison's disease, ovarian failure and hypoparathyroidism. Clinics Endoc Metab 1975; 4: 379-434

IRVINE WJ, GRAY RS, McCALLUM CJ. Pancreatic-islet cell antibody as a marker for asymptomatic and latent diabetes and prediabetes. Lancet 1976; a; ii: 1097-1102

IRVINE WJ, MacCUISH AC, CAMPBELL CJ, DUNCAN LJP. Organ-specific cell-mediated autoimmunity in diabetes mellitus. Acta Endocrinologica (Kbh) 1976; b;(Supp.205):65-76

IRVINE WJ, McCALLUM CJ, GRAY RS, CAMPBELL CJ, DUNCAN LJP, FARQUHAR JW, VAUGHAN H, MORRIS PJ. Pancreatic islet cell antibodies in diabetes mellitus correlated with the duration and type of diabetes, coexistent autoimmune disease and HLA type. Diabetes 1977; 26: 138-47

IRVINE WJ, SAWERS JSA, FEEK CM, PRESCOTT RJ, DUNCAN LJP. The value of islet cell antibody in predicting secondary failure of oral hypoglycaemic agent therapy in diabetes mellitus. J Clin Lab Immunol 1979; 2: 23-6

JACKSON WPU. Present state of prediabetes. Diabetes 1960; 9: 373-378

JACKSON WPU. Nutrition and the proper use of food supplies. IDF Bulletin 1981; (3): 20.

JENSON AB, ROSENBERG HS, NOTKINS AL. Pancreatic islet cell damage in children with fatal viral infections. Lancet 1980; ii: 354-358

JOYSEY VC, WOLF E. HLA-A, -B and -C antigens, their serology and cross-reaction. Br Med Bull 1978; 34: 217-222

JUNKER K, EGEBURG Y, KROMANN H, NERUP J. An autopsy study of the islets of Langerhans in acute onset juvenile diabetes mellitus. Acta Pathol Microbiol Scand (Sect A) 1977; 85: 699-706

KADRNKA-LOVRENCIC M, RABATIC S, MAZURAN M, OBERITER V, DEKARIS D. Cell-mediated immunity in children with juvenile diabetes mellitus. Diabetologia 1978; 14:244. (A)

KAHN CR, MANN D, ROSENTHAL AS, GALLOWAY JA, JOHNSON AH, MENDELL N. The immune response to insulin in man: interaction of HLA alloantigens and the development of the immune response. Diabetes 1982; 31: 716-23

KAPOOR-PILLARISETTI A, MOWBRAY JF, BROSTOFF J, CRONIN EA. HLA dependence of sensitivity to nickel and chromium. Tissue Antigens 1981; 17:261-264

KARAM JH, PROSSER PR, LEWITT PA. Islet cell surface antibodies in a patient with diabetes mellitus after rodenticide ingestion. N Engl J Med 1978; 299:1191.

KARAM JH, LEWITT PA, YOUNG CW, NOWLAIN RE, FRANKEL BJ, FUJIYA H, FREEDMAN Z, GRODSKY G. Insulinopenic diabetes after rodenticide (Vacor) ingestion: a unique model of acquired diabetes in man. Diabetes 1980; 29: 971-978

KARMAZSIN L, AMBRO I, STENSZKY V, KOZMA L, BALAZS C, SVETLANA K. HLA-antigens and some autoimmune features of juvenile diabetes mellitus. Acta Paediatr Acad Sci Hung 1979; 20: 11-19

KEEN H, TRACK NS. Age of onset and inheritance of diabetes: importance of examining relatives. Diabetologia 1968; 4: 317-321

KIESEL U, KOLB H. Low-dose streptozotocin-induced autoimmune diabetes is under the genetic control of the major histocompatibility complex in mice. Diabetologia 1982; 23: 69-71

KIRJARINTA L, ERIKSSON AW. Preliminary studies on diabetes mellitus in Finnish Lapps. Nordic Council Arct Med Res Rep 1976; 15: 41-47

KIRK RL, THEOPHILUS J, WHITEHOUSE S, COURT J, ZIMMET P. Genetic susceptibility to diabetes mellitus: the distribution of properdin factor B (Bf) and glyoxalase (GLO) phenotypes. Diabetes 1979; 28: 949-51

KLöPPEL G. "Insulin" induced insulitis. In: Ortved Andersen O, Deckert T, Nerup J, Immunological Aspects of Diabetes Mellitus. Acta Endocrinologica (Kbh) 1976 (Supp 205): 1976; 107-120

KNOWLER WC, BENNETT PH, BOTTAZZO GF, DONIACH D. Islet cell antibodies and diabetes mellitus in Pima Indians. Diabetologia 1979; 17: 161-4

KöBBERLING J. Untersuchungen zur Genetik des Diabetes mellitus. Eine geeignete Methode zur Durchführung von Alterskorrekturen. Diabetologia 1969; 5: 392-396

KöBBERLING J. Genetic heterogeneities within idiopathic diabetes. In: Creutzfeldt W, Köbberling J, Neel JV (eds), The genetics of diabetes mellitus. Springer, Berlin 1976; 79-87

KOIVISTO VA, AKERBLOM HK, WASZ-HOCKERT O. The epidemiology of childhood diabetes mellitus in Northern Finland. Nordic Council Arct Med Res Rep 1976; 15: 58-65

KOLB H, FREYTAG G, KIESEL U, KOLB-BACHOFEN V. Spontaneous autoimmune reactions against pancreatic islets in mouse strains with generalized autoimmune disease. Diabetologia 1980; 19: 216-221

KOLB H, SCHERNTHANER G, GRIES FA. Immunological aspects of Type 1 (insulin-dependent) diabetes: pathogenesis and therapy. Diabetologia 1983; 24 :394-395

KöMPFF J, BISSBORT S, GUSSMANN S, RITTER H. Polymorphism of red cell glyoxalase-I (E.C.:4.4.1.5): A new genetic marker in man. Humangenetik 1975; 27: 141-143

KöMPFF J, BISSBORT S, GöHLER F, SCHUNTER F, WERNET P. Map order of the linkage group GLO-Bf-HLA-A-PGM$_3$ on human chromosome 6. Scand J Immunol 1977; 6:453-457

KRAUS EJ. Die pathologisch-anatomischen Veränderungen des Pankreas beim Diabetes mellitus. In: Henke F, Lubarsch O (eds), Handbuch der spez. path. Anat. und Histologie, Vol V, Part 2. Springer, Berlin 1929; 622-747 (esp.707-19)

KYLLO CJ, NUTTALL FQ. Prevalence of diabetes mellitus in school-age children in Minnesota. Diabetes 1978; 27: 57-60

LANDING BH, PETTIT MD, WIENS RL, KNOWLES H, GUEST GM. Antithyroid antibody and chronic thyroiditis in diabetes mellitus. J Clin Endocrinol Metab 1963; 23: 119-20

LANE MA. The cytological characteristics of the areas of Langerhans. Am J Anat 1907; 7: 409-422

LANGERHANS P. Beitrage zur mikroskopischen Anatomie der Bauchspeicheldruse. Inaugural dissertation, Berlin. 1869

LANGENBECK U, JORGENSEN G. The genetics of diabetes mellitus - a review of twin studies. In: Creutzfeldt W, Köbberling J, Neel JV (eds), The genetics of diabetes mellitus. Springer, Berlin 1976; 21-25

LAPORTE RE, FISHBEIN HA, DRASH AL, KULLER LH, SCHNEIDER BB, ORCHARD TJ, WAGNER DK. The Pittsburgh insulin-dependent diabetes mellitus (IDDM) registry. The incidence of juvenile-onset diabetes mellitus in Allegheny County, Pennsylvania (1965-76). Diabetes 1981; 30: 279-84

LECOMPTE PM. "Insulitis" in early juvenile diabetes. Arch Path 1958; 66: 450-7

LECOMPTE PM, LEGG MA. Insulitis (lymphocytic infiltration of pancreatic islets) in late onset diabetes. Diabetes 1972; 21: 762-9

LEDDY JP, SIMONS RL, DOUGLAS RG. Effect of selective complement deficiency on the rate of neutralization of envelopped viruses by human sera. J Immunol 1977; 118:28-34

LENDRUM R, WALKER G, GAMBLE DR. Islet cell antibodies in juvenile diabetes mellitus of recent onset. Lancet 1975; i: 880-3

LENDRUM R, WALKER JG. Serum antibodies in human pancreatic disease. Gut 1975; 16: 365-71

LENDRUM R, WALKER G, CUDWORTH AG, THEOPHANIDES C, PYKE DA, BLOOM A, GAMBLE DR. Islet cell antibody in diabetes mellitus. Lancet 1976; ii: 1273-6

LERNMARK A, FREEDMAN ZR, HOFMANN C, RUBENSTEIN AH, STEINER DF, J ACKSON RL, WINTER RJ, TRAISMAN HS. Islet-cell-surface antibodies in juvenile diabetes mellitus. N Engl J Med 1978; 299: 375-80

LERNMARK A, HÄGGLÖF B, FREEDMAN ZR, IRVINE WJ, LUDVIGSSON J, HOLMGREN G. A prospective analysis of antibodies reacting with pancreatic islet cells in insulin-dependent diabetic children. Diabetologia 1981; 20: 471-4

LESLIE RDG, PYKE DA. Chlorpropamide-alcohol flushing: a dominantly inherited trait associated with diabetes. Br Med J 1978; ii: 1519-1521

LESLIE RDG, PYKE DA. Letter to the editor. Diabetologia 1979; 16: 139.

LESTRADET H, BATTISTELLI J, LEDOUX M. L'heredité dans le diabète infantile. Diabète 1972; 20: 17-21

LESTRADET H, BESSE J. Prévalence et incidence du diabète juvenile insulino-dépendante en France. Diabète et Metabolisme 1977; 3: 229-234

LEVIT SG, PESSIKOVA LN. The genetics of diabetes mellitus. Proc Maxim Gorky Med Genet Res Inst Moskva 1934; 3:132-47 (not available in UK; cited by Harris 1950)

LEVINE BB, STEMBER RH, FOTINO M. Ragweed hay fever: genetic control and linkage with HL-A haplotypes. Science 1972; 178:1201-1203

LIKE AA, ROSSINI AA. Streptozotocin-induced pancreatic insulitis: new model of diabetes mellitus. Science 1976; 193:415-417

LIKE AA, ROSSINI AA. Streptozotocin-induced pancreatic insulitis in mice: morphological and physiological studies. Lab Invest 1978; 38:470-486

LIPSICK J, BEATTIE G, OSLER AG, KAPLAN NO. Passive transfer of lymphocytes from diabetic man to athymic mouse. Lancet 1979; i: 1290-1

LISTER J, TANNER JM. The physique of diabetics. Lancet 1955; ii: 1002-4

LUCARELLI P, SCACCHI R, CORBO RM, PALMARINO R, ORSINI M, CAMPEA L, CARAPELLA E, PASCONE R. Genetic polymorphisms in juvenile-onset diabetes. Human Hered 1978; 28:89-95

LUDVIGSSON J, SAFWENBERG J, HEDING LG. HLA-types, C-peptide and insulin antibodies in juvenile diabetes. Diabetologia 1977; 13: 13-17

LUDVIGSSON J, HEDING LG. Beta-cell function in children with diabetes. Diabetes 1978; (Supp 1); 27: 230-4

LUDWIG H, SCHERNTHANER G, MAYR WR. Is HLA-B7 a marker associated with a protective gene in juvenile-onset diabetes mellitus? N Engl J Med 1976; 294: 1066.

LUDWIG H, SCHERNTHANER G, MAYR WR. The importance of HLA genes to susceptibility in the development of juvenile diabetes mellitus. Diabète Metab 1977; a; 3: 43-8

LUDWIG H, SCHERNTHANER G, MAYR WR. Inselzellantikörper beim juvenilen Diabetes mellitus. Dtsch Med Wochenschr 1977; b; 102:1221-1224,1226

McCANN VJ, DAVIS RE, WELBORN TA. Distribution of glyoxalase phenotypes in diabetic patients in Western Australia. Diabetologia 1981; 21: 76 (A)

MacCUISH AC, BARNES EW, IRVINE WJ, DUNCAN LJP. Antibodies to pancreatic islet cells in insulin dependent diabetics with autoimmune disease. Lancet 1974; a; ii: 1529-31

MacCUISH AC, JORDAN J, CAMPBELL CJ, DUNCAN LJP, IRVINE WJ. Cell mediated immunity to human pancreas in diabetes mellitus. Diabetes 1974; b; 23:693-697

MacCUISH AC, URBANIAK SJ, CAMPBELL CJ, DUNCAN LJP, IRVINE WJ. Phytohaemagglutinin transformation and circulating lymphocyte subpopulations in insulin dependent diabetes patients. Diabetes 1974; c; 23:708-712

MacCUISH AC, JORDAN J, CAMPBELL CJ, DUNCAN LJP, IRVINE WJ. Cell-mediated immunity in diabetes mellitus: lymphocyte transformation by insulin and insulin fragments in insulin-treated and newly-diagnosed diabetics. Diabetes 1975; 24: 36-43

MacDONALD MJ. Equal incidence of adult onset diabetes mellitus among ancestors of juvenile onset diabetics and normal subjects. Diabetologia 1974; 10: 767-73

McLAREN NK, HUANG SW, FOGH J. Antibody to cultured human insulinoma cells in insulin-dependent diabetics. Lancet 1975; i: 997-1000

MACLEAN N, OGILVIE RF. Quantitative estimation of the pancreatic islet tissue in diabetic subjects. Diabetes 1955; 4: 367-76

MACLEAN N, OGILVIE RF. Observations on the pancreatic islet tissue of young diabetic subjects. Diabetes 1959; 8: 83-91

MACMILLAN DR, KOTOYAN M, ZEIDNER D, HAFEZI B. Seasonal variation in onset of diabetes in children. Pediatrics 1977; 59: 113-115

McMILLAN DE, GEEVERGHESE PJ. Dietary cyanide and tropical malnutrition diabetes. Diabetes Care 1979; 2: 202-8

McNEIL BJ, KEELER E, ADELSTEIN SJ. Primer on certain elements of medical decision making. New Engl J Med 1975; 293: 211-215

MADSBAD S, BOTTAZZO GF, CUDWORTH AG, DEAN BM, FABER OK, BINDER C. Islet cell antibodies and beta-cell function in insulin-dependent diabetics. Diabetologia 1980; 18:45-47

MALAISSE-LAGAE F, STEFAN Y, COX J, PERRELET A, ORCI L. Identification of a lobe in the adult human pancreas rich in pancreatic polypeptide. Diabetologia 1979; 17: 361-5

MANCINI AM, COSTANZI G, ZAMPA GA. Human insulin antibody detected by immunofluorescence techniques. Lancet 1964; i: 726.

MARLISS EB, NAKHOODA AF, POUSSIER P, SIMA AAF. The diabetic syndrome of the 'BB' Wistar rat: possible relevance to Type 1 (insulin-dependent) diabetes in man. Diabetologia 1982; 22: 225-32

MARSH DG, BIAS WB, HSU SH, GOODFRIEND L. Association of the HLA-A7 cross-reactive group with a specific reaginic antibody response in allergic man. Science 1973; 179:691-693

MATTIUZ PL, IHDE D, PIAZZA A, CEPPELLINI R, BODMER WF. New approaches to the population genetic and segregation analysis of the HLA system. In: Terasaki P (ed.), Histo-compatibility Testing 1970, 193-205. Munksgaard, Copenhagen, 1971

MENOZZI P, PIAZZA A, CAVALLI-SFORZA L. Synthetic maps of human gene frequencies in Europeans. Science 1978; 201: 786-92

MENSER MA, FORREST JM, HONEYMAN MC, BURGESS JA. Diabetes, HL-A antigens, and congenital rubella. Lancet 1974; ii: 1508-1509

MENSER MA, FORREST JM, BRANSBY RD. Rubella infection and diabetes mellitus. Lancet 1978; i: 57-60

MEO T, ATKINSON J, BERNOCO M, BERNOCO D, CEPPELLINI R. Mapping of the HLA locus controlling C2 structural variants and linkage disequilibrium between alleles C2 and Bw15. Eur J Immunol 1976; 6: 916-919

MILLER LV, STOKES JD, SILPIPAT C. Diabetes mellitus and autonomic dysfunction after Vacor rodenticide ingestion. Diabetes Care 1978; 1: 73-76

MILLER AP, RICH S, BARBOSA J. Insulin dependent diabetic families: sex ratio and HLA haplotype segregation. Lancet 1981; i: 388.

MIROUZE J, SEIGNALET J, SELAM J-L, LAPINSKI H, JAFFIOL C. Antigènes HLA chez les diabétiques insulino-dépendants. Nouv Presse Med 1976; 5: 1628-30

MØLLER-PETERSEN J. Nomogram for predictive values and efficiencies of tests. Lancet 1985; i: 348.

MORGAN CR, LAZAROW A. Immunoassay of plasma insulin levels of normal, sub-diabetic and diabetic rats. Diabetes 1963; 12: 115-126

MORRIS PJ, VAUGHAN H, IRVINE WJ, McCALLUM FJ, GRAY RS, CAMPBELL CJ, DUNCAN LJP, FARQUHAR JW. HLA and islet cell antibody in diabetics. Lancet 1976; ii: 652-3

MOURANT AE, KOPEC AC, DOMANIEWSKA-SOBCZAK K. Blood groups and diseases. Oxford University Press, Oxford, 1978

MÜLLER GL De diabete praesertim mellitio. Dissertatio inauguralis, Gottingae Typogr. Ac. 1822

MÜLLER R, KOLB H, KUSCHAK D, JÖRGENS V, GRIES FA. Analysis of T-lymphocyte subpopulations in juvenile onset diabetes. Clin Exp Immunol 1980; 39: 130-135

MUNRO A, WALDMANN H. The major histocompatibility system and the immune response. Br Med Bull 1978; 34: 253-258

MUSTONEN A, ILONEN J, TIILIKAINEN A, KATAJA M, AKERBLOM HK. An analysis of epidemiological data in HLA-typed diabetic children. Diabetologia 1985; 28: 397-400

MUTCH WJ, BOTTAZZO GF, STOWERS JM. Islet cell antibodies in subjects with impaired glucose tolerance. Diabetologia 1981; 20:673. (A)

NAGAOKA K, SAKURAMI I, NABEYA N, IMURA H, KUNO S. Antimicrosomal antibodies, gastric parietal cell antibodies and antinuclear factors in insulin-dependent diabetes mellitus. Endocr Japon 1979; 26: 599-603

NAKHOODA AF, LIKE AA, CHAPPEL CI, MURRAY FT, MARLISS EB. The spontaneously diabetic Wistar rat. Metabolic and morphologic studies. Diabetes 1977; 26:100-112

NEEL JV. Diabetes mellitus - a geneticist's nightmare. In: Creutzfeldt W, Köbberling J, Neel JV (eds). The Genetics of Diabetes Mellitus. Springer, Berlin, 1976; 1-11

NELSON PG, PYKE DA, CUDWORTH AG, WOODROW JC, BATCHELOR JR. Histocompatibility antigens in diabetic identical twins. Lancet 1975; ii: 193-4

NELSON PG, PYKE DA. Genetic diabetes not linked to the HLA locus. Br Med J 1976; i: 196-7

NERUP J, ORTVED ANDERSEN O, BENDIXEN G, EGEBERG J, POULSEN JE. Antipancreatic hypersensitivity in diabetes mellitus. Diabetes 1971; 20: 424-427

NERUP J, BINDER C. Thyroid, gastric and adrenal auto-immunity in diabetes mellitus. Acta Endocrinologica 1973; 72: 279-286

NERUP J, ORTVED ANDERSEN O, BENDIXEN G, EGEBERG J, GUNNARSSON R, KROMANN G, POULSEN JE. Cell-mediated immunity in diabetes mellitus. Proc Roy Soc Med 1974 a; 67: 506-513

NERUP J, PLATZ P, ORTVED ANDERSON O, CHRISTY M, LYNGSØE J, POULSEN JE, RYDER LP, STAUB NIELSEN L, THOMSEN M, SVEJGAARD A. HL-A antigens and diabetes mellitus. Lancet 1974 b; ii: 864-6

NERUP J, CATHELINEAU G, SEIGNALET J, THOMSEN M. HLA and endocrine disease. In: Dausset J, Svejgaard A (eds.), HLA and disease, 149-167. Munksgaard, Copenhagen, 1977 a

NERUP J, ORTVED ANDERSEN O, BUSCHARD K, CHRISTAU B, CHRISTY M, KROMANN H, PLATZ P, RYDER A, SVEJGAARD A, THOMSEN M. HLA-D typing in insulin-dependent diabetes mellitus (IDDM). Diabetologia 1977 b; 13: 421 (A)

NEUFELD M, MACLAREN NK, RILEY WJ, LEZOTTE D, McLAUGHLIN JV, SILVERSTEIN J, ROSENBLOOM AL. Islet cell and other organ-specific antibodies in U.S. Caucasians and Blacks with insulin-dependent diabetes mellitus. Diabetes 1980; 29: 589-592

NISSLEY SP, DRASH AL, BLIZZARD RM, SPERLING M, CHILDS B. Comparison of juvenile diabetics with positive and negative organ-specific antibody titres - evidence for genetic heterogeneity. Diabetes 1973; 22: 63-5

NORTH AF, GORWITZ K, SULTZ HA. A secular increase in the incidence of juvenile diabetes mellitus. J Pediat 1977; 91: 706-710

OLI JM, NWOKOLO C. Diabetes after infectious hepatitis: a follow-up study. Br Med J 1979; i: 926-927

ONODERA T, TONIOLO A, RAY UR, JENSON AB, KNAZEK RA, NOTKINS AL. Virus-induced diabetes mellitus. XX. Polyendocrinopathy and autoimmunity. J Exp Med 1981; 153:1457-1473

ONODERA T, SUZUKI H, TONIOLO A, NOTKINS AL. Virus-induced diabetes: cytomegalovirus and multiple environmental insults. Diabetologia 1983; 24: 218-219

OFFICE OF POPULATION CENSUSES AND SURVEYS. Census 1981: County Report - Berkshire, Part 1. HMSO 1982

ORCHARD TJ, WAGENER DK, RABIN BS, LAPORTE RE, CAVENDER D, KULLER LH, DRASH AL, BECKER DJ. Glucose tolerance in siblings of Type 1 diabetic patients : relationship to HLA status. Diabetologia 1986; 29:39-45

ORCI L, BAETENS D, RAVAZZOLA M, STEFAN Y, MALAISSE-LAGAE F. Pancreatic polypeptide and glucagon: non-random distribution in pancreatic islets. Life Sci 1976; 19: 1811-6

ORCI L, STEFAN Y, MALAISSE-LAGAE F, PERRELET A. Instability of pancreatic endocrine cell populations throughout life. Lancet 1979; i: 615-6

OWERBACH D, LERNMARK A, PLATZ P, RYDER LP, RASK L, PETERSON PA, LUDVIGSSON J. HLA-D region β-chain DNA endonuclease fragments differ between HLA-DR identical healthy and insulin-dependent diabetic individuals. Nature 1983; 303: 815-817

PALMER JP, COONEY MK, WARD RH, HANSEN JA, BRODSKY JB, RAY CG, CROSSLEY JR, ASPLIN CM, WILLIAMS RH. Reduced Coxsackie antibody titres in Type 1 (insulin-dependent) diabetic patients presenting during an outbreak of Coxsackie B3 and B4 infection. Diabetologia 1982; 22: 426-429

PALUMBO PJ, ELVEBACK LR, CHU CP, CONNOLLY DC, KURLAND LT. Diabetes mellitus: Incidence, prevalence, survivorship and causes of death in Rochester, Minnesota, 1945-1970. Diabetes 1976; 25: 566-73

PAPASPYROS NS. The history of diabetes mellitus (2nd edition). Thieme, Stuttgart 1964

PARR CW, BAGSTER IA, WELCH SG. Human red cell glyoxalase-I polymorphism. Biochem Genet 1977; 15: 109-113

PATTERSON K, CHANDRA RS, JENSON AB. Congenital rubella, insulitis and diabetes mellitus in an infant. Lancet 1981 a; i: 1048-9.

PATTERSON CC, SMITH PG, MANN JI,THOROGOOD M, HEASMAN MA, CLARKE JA. Distribution of juvenile-onset diabetes cases in Scotland (1968- 76). Diabetologia 1981 b; 20: 673 (A)

PENROSE LS. The genetical background of common diseases. Acta Genetica 1953; 4: 257-65

PERLMAN LV. Familial incidence of diabetes in hyperthyroidism. Ann Intern Med 1961; 55: 796-799

PETTIT MD, LANDING BH, GUEST G. Antithyroid antibody in juvenile diabetics. J Clin Endoc Metab 1961; 21: 209-10

PHILLIPS WR, PAULI R. Simultaneous onset of insulin-dependent diabetes mellitus in siblings. Lancet 1981; ii: 807.

PINCUS G, WHITE P. On the inheritance of diabetes mellitus. I. An analysis of 675 family histories. Am J Med Sci 1933; 186: 1-14

PINCUS G, WHITE P. On the inheritance of diabetes mellitus. II. Further analysis of family histories. Am J Med Sci 1934 a; 188: 159-168

PINCUS G, WHITE P. On the inheritance of diabetes mellitus. III. The blood sugar va lues of the relatives of diabetics. Am J Med Sci 1934 b; 188: 782-90

PLATZ P, THOMSEN M, SVEJGAARD A, CUDWORTH AG, WOODROW JC, NERUP J. More on genetics of diabetes mellitus. New Engl J Med 1978; 298: 1200-1201

PLATZ P, JAKOBSEN BK, MORLING N, RYDER LP, SVEJGAARD A, THOMSEN M, CHRISTY M, KROMANN H, BENN J, NERUP J, GREEN A, HAUGE M. HLA-D and -DR antigens in genetic analysis of insulin dependent diabetes mellitus. Diabetologia 1981; 21: 108-15

PLATZ P, JAKOBSEN BK, SVEJGAARD A, THOMSEN BS, JENSEN KB, HENNINGSEN K, LAMM LU. No evidence for linkage between HLA and maturity onset type of diabetes in young people . Diabetologia 1982; 23: 16 - 18

POZZILLI P, SENSI M, GORSUCH AN, BOTTAZZO GF, CUDWORTH AG. Evidence for raised K cell levels in Type 1 diabetes. Lancet 1979; ii: 173-5

POZZILLI P, SENSI M, DEAN BM, GORSUCH AN, CUDWORTH AG. The importance of measuring Fc+ (killer) cells in insulin-dependent (Type 1) diabetes. Acta Diabet Lat 1980; 17: 119-122

PRINCE GA, JENSON AB, BILLUPS LC, NOTKINS AL. Infection of human pancreatic beta cell cultures with mumps virus. Nature 1978; 271: 158-161

PROUST AJ, SMITHURST BA. Epidemiology of diabetes mellitus in Australia. Med J Australia 1968; ii: 769-72

PUJOL-BORRELL R, KHOURY EL, BOTTAZZO GF. Islet cell surface antibodies in Type 1 (insulin dependent) diabetes mellitus: use of human fetal pancreas cultures as substrate. Diabetologia 1982; 22: 89-95

PYKE DA, LESLIE RDG. Chlorpropamide-alcohol flushing: a definition of its relation to non insulin dependent diabetes. Br Med J 1978; ii: 1521-2

PYKE DA. Current concepts in aetiology: non-insulin-dependent diabetes. In: Bellingham AJ (ed), Adv. Med. 16. Pitman Medical, London 1980; 136-40

RAHIER J, WALLON J, HENQUIN JC. Abundance of somatostatin cells in the human neonatal pancreas. Diabetologia 1980; 18: 251-4

RAHIER J, GOEBBELS RM, HENQUIN JC. Cellular composition of the human diabetic pancreas. Diabetologia 1983; 24:366-371

RAKIETEN N, RAKIETEN ML, NADKARNI MV. Studies on the diabetic action of streptozotocin (NSC-37917). Cancer Chemother Rep 1963; 29: 91-98

RAUM D, ALPER CA, STEIN R, GABBAY KH. Genetic marker for insulin dependent diabetes mellitus. Lancet 1979; i: 1208-10

RAYFIELD EJ, SETO Y, WALSH S, McEVOY RC. Virus induced alterations in insulin release in hamster islets of Langerhans. J Clin Invest 1981; 68: 1172-1181

RENOLD AE, STEINKE JS, SOELDNER JS, ANTONIADES HN, SMITH RE. Immunological response to the prolonged administration of heterologous and homologous insulin in cattle. J Clin Invest 1966; 45: 702-13

RENOLD AE, GONET AE, VECCHIO D Immunopathology of the endocrine pancreas. In: Miescher PA, Müller-Eberhard HJ (eds), Textbook of Immuno-pathology, II. Grune & Stratton, New York & London, 1969; Volume 2: 595-601.

RILEY WJ, MACLAREN NK, NEUFELD M. Adrenal autoantibodies and Addison disease in insulin dependent diabetes mellitus. J Pediatr 1980; 97: 191-5

RIMOIN DL. Genetic syndromes associated with glucose intolerance. Creutzfeldt W, Köbberling J, Neel JV (eds). The Genetics of Diabetes Mellitus. Springer, Berlin, 1976; 43-63

RITTENHOUSE HG, OXENDER DL, PEK S, AR D. Complement-mediated cytotoxic effects on pancreatic islets with sera from diabetic patients. Diabetes 1980; 29: 317-22

RITTNER C, BERTRAMS J. On the significance of C2, C4 and Factor B polymorphisms in disease. Hum Genet 1981; 56: 235-47

ROBERTSON JS. The pancreatic lesion in adult mice infected with a strain of pleurodynia virus. Austral J Exp Biol 1954; 32: 393-409

RODGER B, WHITTINGHAM S, MARTIN FIR, HAWKINS BR, DAWKINS RL, WELBORN TA. A population survey of pancreatic islet cell antibodies. Clin Exp Immunol 1980; 39: 125-9

ROSS ME, ONODERA T, BROWN KS, NOTKINS AL. Virus-induced diabetes mellitus IV. Genetic and environmental

factors influencing the development of diabetes after infection with the M variant of encephalomyocarditis virus. Diabetes 1976; 25: 190-7

ROSSINI AA, WILLIAMS RM, MORDES JP, APPEL MC, LIKE AA. Spontaneous diabetes in the gnotobiotic BB/W rat. Diabetes 1980; 28: 1031-2

ROTTER JI, RIMOIN DL. Heterogeneity in diabetes mellitus - update, 1978. Diabetes 1978; 27: 599-605

RUBINSTEIN P, SUCIU-FOCA N, NICHOLSON JF, FOTINO M, MOLINARO A, HARISIADIS L, HARDY MA, REEMTSMA K, ALLEN FH. The HLA system in the families of patients with juvenile diabetes mellitus. J Exp Med 1976; 143: 1277-82

RUBINSTEIN P, SUCIU-FOCA N, NICHOLSON JF. Genetics of juvenile diabetes mellitus: a recessive gene closely linked to HLA D and with 50% penetrance. N Engl J Med 1977; 297: 1036-40

RUBINSTEIN P, WALKER ME, FEDUN B, WITT ME, COOPER LZ, GINSBERG-FELLNER F. The HLA system in congenital rubella patients with and without diabetes. Diabetes 1982; 31:1088-1091

RYDER LP, SVEJGAARD A. Associations between HLA and disease - Report from the HLA and disease registry of Copenhagen, 1976. Munksgaard, Copenhagen, 1976

RYDER LP, ANDERSEN E, SVEJGAARD A. HLA and Disease Registry: 3rd report. Munksgaard, Copenhagen, 1979

SACHS JA, CUDWORTH AG, JARAQUEMADA D, GORSUCH AN, FESTENSTEIN H. Type 1 diabetes and the HLA-D locus. Diabetologia 1980; 18: 41-3

SAI P, BOITARD C, DEBRAY-SACHS M, POUPLARD A, ASSAN R, HAMBURGER J. Complement-fixing islet cell antibodies from some diabetic patients alter insulin release in vitro. Diabetes 1981; 30:1051-1057

SAKURAMI T, NABEYA N, NAGAOKA K, MATSUMORI A, KUNO S, HONDA A. Antibodies to Coxsackie B viruses and HLA in Japanese with juvenile-onset Type 1 (insulin-dependent) diabetes mellitus. Diabetologia 1982; 22:375-377

SCHERNTHANER G, LUDWIG H, MAYR W. B-lymphocyte allo-antigens and insulin-dependent diabetes mellitus. Lancet 1977; ii: 1128.

SCHERNTHANER G, LUDWIG H, MAYR WR. Immunoglubulin G-insulin antibodies and immune region-associated alloantigens in insulin-dependent diabetes mellitus. J Clin Endocrinol Metab 1979; 48: 403-7

SCHERNTHANER G, BORKENSTEIN M, FINK M, MAYR WR, MENZEL J, SCHOBER E. Immunogenicity of human insulin (Novo) or pork monocomponent insulin in HLA-DR-typed insulin-dependent diabetic individuals. Diabetes Care 1983; 6 (Supp 1): 43-48

SCHMIDT MB. über der Beziehung der Langerhans'schen Inseln des Pankreas zum Diabetes mellitus. Münch Med Wochnschr 1902; 49: 51-4

SELAM JL, CLOT J, ANDARY M, MIROUZE J. Circulating lymphocyte subpopulations in juvenile insulin dependent diabetes. Diabetologia 1979; 16: 35-40

SELYE H. Textbook of Endocrinology. Acta Endocrinologica, Montreal 1947;

SENSI M, POZZILLI P. Cell mediated immunity in the families of Type 1 diabetics and in newly diagnosed patients. In: Waldhausl W, Alberti KGMM (eds), 10th Congr Internat Diab Fed, Excerpta Medica, Amsterdam, 1979; 211 (A)

SENSI M, POZZILLI P, CUDWORTH AG. Antibody-dependent and natural killer cytotoxicity in Type 1 diabetes. Acta Diabet Lat 1981 a; 18: 339-344

SENSI M, POZZILLI P, GORSUCH AN, BOTTAZZO GF, CUDWORTH AG. Increased killer cell activity in insulin dependent (Type 1) diabetes mellitus. Diabetologia 1981 b; 20: 106-9

SENSI M, POZZILLI P, VENTRIGLIA L, DONIACH I, CUDWORTH AG. Histology of the islets of Langerhans following administration of human lymphocytes into athymic mice. Clin Exp Immunol 1982; 49:81-86

SERJEANTSON S. Distorted HLA segregation or biased ascertainment? Lancet 1980; i: 40-1

SERRA AS, FARNDON JR, SHENTON BK, JOHNSTON IDA. Passive transfer of lymphocytes from diabetic man to athymic mouse. Lancet 1979; i: 1292.

SHARPEY-SCHAFER E. The endocrine organs. Longmans, Green & Todd, London, 1926

SHAW DUNN J, SHEEHAN HL, McLETCHIE NGB. Necrosis of islets of Langerhans produced experimentally. Lancet 1943 a; i: 484-487

SHAW DUNN J, McLETCHIE NGB. Experimental alloxan diabetes in the rat. Lancet 1943 b; ii: 384-387

SHAW DUNN J, KIRKPATRICK J, McLETCHIE NGB, TELFER SV. Necrosis of the islets of Langerhans produced experimentally. J Path Bact 1943; c; 55: 245-257

SILVER J, WALKER LE, REISFELD RA, PELLEGRINO MA, FERRONE S. Structural studies of murine I-E and human DR antigens. Molecular Immunology 1979; 16: 37-42

SILWER H. Incidence of diabetes mellitus in a Swedish county. In: Silwer H, Oscarsson PN (eds.), Incidence and coincidence of diabetes mellitus and pulmonary tuberculosis in a Swedish county. Acta Med Scand 1958; Supp 335: 5-22

SIMKINS S. Antithyroglobulin antibodies in diabetes mellitus. Diabetes 1968; 17: 136-140

SIMPSON NE. The genetics of diabetes: a study of 233 families of juvenile diabetics. Ann Hum Genet (London) 1962; 26: 1-22

SIMPSON NE. Diabetes in the families of diabetics. Can Med Assoc J 1968; 98: 427-432

SIMPSON NE. Heritabilities of liability to diabetes when sex and age at onset are considered. Ann Hum Genet 1969; 32: 283-303

SIMPSON NE The genetics of diabetes mellitus - a review of family data. In: Creutzfeldt W, Köbberling J, Neel JV (eds), The Genetics of Diabetes Mellitus. Springer, Berlin, 1976; 12-20

SINGAL DP, BLAJCHMAN MA. Histocompatibility (HL-A) antigens, lymphocytotoxic antibodies and tissue antibodies in patients with diabetes mellitus. Diabetes 1973; 22: 429-432

SMITHELLS RW, SHEPPARD S, MARSHALL WC, PECKHAM C. Congenital rubella and diabetes mellitus. Lancet 1978; i: 439.

SOLOMON IL, BLIZZARD RM. Autoimmune disorders of endocrine glands. J Pediatr 1963; 63: 1021-33

SPENCER KM, GORSUCH AN, CUDWORTH AG, BOTTAZZO GF. Diabetes and month of birth. Lancet 1982; i: 449.

SPENCER KM, TARN A, DEAN BM, LISTER J, BOTTAZZO GF. Fluctuating islet-cell autoimmunity in unaffected relatives of patients with insulin-dependent diabetes. Lancet 1984; i: 764-766

SPIELMAN RS, BAKER L, ZMIJEWSKI CM. Gene dosage and susceptibility to insulin-dependent diabetes. Ann Human Genet 1980; 44: 135-150

SRIKANTA S, GANDA OP, EISENBARTH GS, SOELDNER JS. Islet-cell antibodies and beta-cell function in monozygotic triplets and twins initially discordant for Type 1 diabetes mellitus. N Engl J Med 1983; 308:322-325

SSOBOLEW LW. Zur normalen und pathologischen Morphologie der inneren Sekretionen der Bauchspeicheldrüse. Arch Path Anat (Virchow) 1902; 168: 91-128

STANSFIELD OH, WARREN S. Inflammation involving the islets of Langerhans in diabetes. N Engl J Med 1928; 198: 686-9

STEEL JM, GRAY RS, CLARKE BF. Obstetric history of diabetics: its relevance to aetiology of diabetes mellitus. Brit Med J 1979; i: 1303-1305

STEINBERG AG, WILDER RM. A study of the genetics of diabetes mellitus. Am J Hum Genet 1952; 4: 113-35

STEINBERG AG. Heredity in diabetes mellitus. Diabetes 1961; 10: 269-79

STERKY G, HOLMGREN G, GUSTAVSON KH, LARSON Y, LUNDMARK KM, NILSSON KD, SAMUELSON G, THALME B, WALL S. The incidence of diabetes mellitus in Swedish children 1970-75. Acta Pediatrica Scandinavica 1978; 67: 139-143

SUAREZ BK, HODGE SE. A simple way to detect linkage for rare recessive diseases: an application to juvenile diabetes. Clin Genet 1979; 15: 126-36

SUAREZ BK, Van EERDEWEGH P. Type 1 (Insulin-dependent) diabetes mellitus: is there strong evidence for a non-HLA linked gene? Diabetologia 1981; 20:524-529

SUCIU-FOCA N, RUBINSTEIN P. Intra-HLA recombinations in juvenile diabetes mellitus. Lancet 1976; i: 371-2

SUCIU-FOCA N, RUBINSTEIN P, NICHOLSON J, SUSINNO E, WEINER J, GODFREY M, HARDY M, RAYFIELD E, REEMTSMA K. Juvenile diabetes mellitus and the HLA system. Transplantation Proc 1979; 11: 1309-13

SULTZ HA, HART BA, ZIELEZNY M, SCHLESINGER ER. Is mumps virus an etiologic factor in juvenile diabetes mellitus? J Pediat 1975; 86: 654-6

SVEJGAARD A, RYDER LP. Associations between HLA and disease; notes on methodology and a report from the HLA and Disease Registry. In: Dausset J, Svejgaard A (eds.), HLA and disease, 46-71. Munksgaard, Copenhagen, 1977.

SVEJGAARD A, PLATZ P, RYDER LP Insulin-dependent diabetes mellitus. In: Terasaki PI (ed.), Histocompatibility Testing 1980, 638-56. UCLA Tissue Typing Laboratory, Los Angeles, 1980

SVENDSEN PA, CHRISTIANSEN JS, WELINDER B, NERUP J. Fast glycosylation of haemoglobin. Lancet 1979; i: 603.

TARN AC, SMITH CP, SPENCER KM, BOTTAZZO GF, GALE EAM. Type 1 (insulin-dependent) diabetes: a disease of slow clinical onset? Br Med J 1987; 294: 342-345

TATTERSALL RB, PYKE DA. Diabetes in identical twins. Lancet 1972; ii: 1120-5

TATTERSALL RB, FAJANS SS. A difference between the inheritance of classical juvenile-onset and maturity-onset type diabetes of young people. Diabetes 1975; 24: 44-53

TATTERSALL R. The inheritance of maturity-onset type diabetes in young people. In: Creutzfeldt W, Köbberling J, Neel JV (eds), The genetics of diabetes mellitus. Springer-Verlag, Berlin. 1976; 88-94.

TERASAKI PI, BERNOCO D, PARK MS, OZTURK G, IWAKI Y. Microdroplet testing for HLA-A, -B, -C and -D antigen. Am J CLin Path 1978; 69: 103-120

TERASAKI PI. Histocompatibility testing 1980. UCLA, Los Angeles 1980

THAIN ME, HAMILTON JR, EHRLICH RM. Coexistence of diabetes mellitus and celiac disease. J Pediatr 1974; 85: 527-9

THEN BERGH H. Die Erbbiologie des Diabetes mellitus. Vorlaüfiges Ergebnis der Zwillingsuntersuchungen. Arch Rassen- und Gesellschaftsbiologie 1938; 32: 289-340

THOMSEN M, PLATZ P, ORTVED ANDERSEN O, CHRISTY M, LYNGSØE J, NERUP J, RASMUSSEN K, RYDER LP, STAUB NIELSEN L, SVEJGAARD A. MLC typing in juvenile diabetes mellitus and idiopathic Addison's disease. Transpl Rev 1975; 22: 125-147

THOMSON G, BODMER WF. The genetic analysis of HLA and disease associations. In: Dausset J, Svejgaard A (eds.), HLA and disease, 86-93. Munksgaard, Copenhagen, 1977;

THOMSON G. A two locus model for juvenile diabetes. Ann Hum Genet 1980; 43: 383-398

THURNEYSSEN O, JANSEN FK, VIALETTES B, VAGUE Ph, SELAM JL, MIROUZE J. Passive transfer of lymphocytes from diabetic man to athymic mouse. Lancet 1979; i: 1291-2

TIENGO A, Del PRETE GF, NOSADINI R, BETTERLE C, GAROTTI C, BERSANI G. Insulin and glucagon secretion in diabetic and non-diabetic patients with circulating islet cell antibodies. Diabetologia 1977; 13:451-458

TONIOLO A, ONODERA T, YOON J-W, NOTKINS AL. Induction of diabetes by cumulative environmental insults from viruses and chemicals. Nature 1980; 288: 383-385

TONIOLO A, ONODERA T, JORDAN G, YOON J-W, NOTKINS AL. Virus-induced diabetes mellitus. Glucose abnormalities produced in mice by the six members of the Coxsackie B virus group. Diabetes 1982; 31: 496-499

TöRESON WE, FELDMAN R, LEE JC, GRODSKY GM. Pathology of diabetes mellitus produced in rabbits by means of immunization with beef insulin. Am J Clin Path 1964; 42: 531 (A)

TöRESON WE, LEE JC, GRODSKY GM. The histopathology of immune diabetes in the rabbit. Am J Path 1968; 52: 1099-1115

UNGAR B, STOCKS AE, MARTIN FIR, WHITTINGHAM S, MACKAY I. Intrinsic-factor antibody, parietal-cell antibody and latent pernicious anaemia in diabetes mellitus. Lancet 1968; ii: 415-8

VAANDRAGER GJ, BRUINING GJ, VEENHOF FJ, DRAGER NM. Incidence of childhood diabetes in The Netherlands. A decrease from north to south over North-Western Europe? Diabetologia 1984; 27:203-206

VAGUE P, MELIS C, MERCIER P, VIALETTES B, LASSMANN V. The increased frequency of the Lewis negative blood group in a diabetic population. Diabetologia 1978; 15: 33-6

Van ROOD JJ, van LEEUWEN A, PLOEM JS. Simultaneous detection of two cell populations by two-colour fluorescence and application to the recognition of B cell determinants. Nature 1976; 262: 795-797

Van de WINKEL M, SMETS G, GEPTS W, PIPELEERS DG. Identification of islet cells binding to circulating cell-surface antibodies of diabetic patients. Diabetologia 1981; 21: 338 (A)

VERGANI D, JOHNSTON C, B-ABDULLAH N, BARNETT AH. Low serum C4 concentrations: an inherited predisposition to insulin dependent diabetes? Br Med J 1983; 286: 926-8

VINKE B, NAGELSMIT WF, von BUCHEM FSP, SMID W. Some statistical investigations in diabetes mellitus. Diabetes 1959; 8: 100-104

Von MEYENBURG H. Ueber "Insulitis" bei Diabetes. Schweiz Med Wochenschr 1940; 21: 554-7

Von MERING J, MINKOWSKI O. Diabetes Mellitus nach Pankreasextirpation. Arch Exper Path Pharmakol 1889; 26: 371-87

WADSWORTH MEJ, JARRETT RJ. Incidence of diabetes in the first 26 years of life. Lancet 1974; ii: 1172-4

WAGENER D, SACKS JM, LaPORTE R, MacGREGOR JM. The Pittsburgh Study of insulin-dependent diabetes mellitus: risk for diabetes among relatives of IDDM. Diabetes 1982; 31: 136-44

WAGENER DK, LAPORTE RE, ORCHARD TJ, CAVENDER D, KULLER LH, DRASH AL. Pittsburgh Diabetes Mellitus Study: 3: an increased prevalence with older maternal age. Diabetologia 1983; 25:82-85

WALKER A, CUDWORTH AG. Type 1 (insulin dependent) diabetic multiplex families: mode of genetic transmission. Diabetes 1980; 29: 1036-9

WARREN S, ROOT HF. The pathology of diabetes, with special reference to pancreatic regeneration. Am J Path 1925; 1: 415-430

WARREN S. The pathology of diabetes in children. J Am Med Ass 1927; 88: 99-101

WEICHSELBAUM A. Über die Veränderungen des Pankreas bei Diabetes mellitus. Sitzungsberichte der Kaiserlichen Akademie der Wissenschaften (Mathematisch-Naturwissenschaftliche Klasse) 1910; 119: 73-281

WEITKAMP LR. HLA segregation ratios. Lancet 1979; ii: 745.

WEITKAMP LR, BARBOSA J, GUTTORMSEN SA, JOHNSON S. Insulin-dependent diabetes mellitus and properdin factor B. Lancet 1979; ii: 369-70

WELCH SG, BOUCHER BJ. A rapid micro-scale method for the measurement of HbA1(a+b+c). Diabetologia 1978; 14:

209-211

WEST KM. Response to cortisone in prediabetes. Diabetes 1960; 9: 379-385

WEST WH, PAYNE SM, WEESE JL, HERBERMAN RB. Human T-lymphocyte subpopulations: correlation between E-rosette -forming affinity and expression of the Fc receptor. J Immunol 1977; 119: 548-554

WEST KM. Epidemiology of diabetes and its vascular lesions. Elsevier, New York, 1980

WEST WH, BOOZER RB, HERBERMAN RB. Low affinity E-rosette formation by the human K cell. J Immunol 1978; 120: 90-95

WEST R, BELMONTE MM, COLLE E, CREPEAU MP, WILKINS J, POIRIER R. Epidemiologic survey of juvenile-onset diabetes in Montreal. Diabetes 1979; 28: 690-3

WHEAT LJ. Infection and diabetes mellitus. Diabetes Care 1980; 3: 187-97

WHITTINGHAM S, MATHEWS JD, MACKAY IR, STOCKS AE, UNGAR B, MARTIN FIR. Diabetes mellitus, autoimmunity and ageing. Lancet 1971; i: 763-7

WHO EXPERT COMMITTEE ON DIABETES MELLITUS. Second Report. WHO Tech Rep Ser 1980; 646

WILSON GL, BELLOMO SC, CRAIGHEAD JE. Effect of interferon on encephalomyocarditis virus infection of cultured mouse pancreatic B cells. Diabetologia 1983; 24: 38-41

WOOLF B. On estimating the relation between blood group and disease. Ann Human Genet 1955; 19: 251-253

WORKING PARTY OF THE COLLEGE OF GENERAL PRACTITIONERS. A diabetes survey. Br Med J 1962; i: 1497-1503

WORKING PARTY OF THE COLLEGE OF GENERAL PRACTITIONERS. The family history of diabetes. Br Med J 1965; i: 960-962

WRENSHALL GA, BOGOCH A, RITCHIE RC. Extractable insulin of the pancreas: correlation with pathological and clinical findings in diabetic and non-diabetic cases. Diabetes 1952; 1: 87-105

YASUDA N. Estimation of haplotype frequency and linkage disequilibrium parameter in the HLA system. Tissue Antigens 1978; 12: 315-32

YOON JW, AUSTIN M, ONODERA T, NOTKINS AL. Virus-induced diabetes mellitus: isolation of a virus from the pancreas of a child with diabetic ketoacidosis. N Engl J Med 1979; 300: 1173-1179

YOON JW, SELVAGGIO S, ONODERA T, WHEELER J, JENSON AB. Infection of cultured human pancreatic B cells with reovirus type 3. Diabetologia 1981; 20: 462-467

ZIMMET P. Type 2 (non-insulin-dependent) diabetes - an epidemiological overview. Diabetologia 1982; 22: 399-411

PART XIII — APPENDICES

CONTENTS

1 Diabetic probands in the BWFS
- A Age, sex, chromosome 6 marker genotype 352
- B Clinical features and family history of diabetes 355

2 Barts-Windsor Family Study: genetic analyses in probands of BWFS, London control subjects and family control haplotypes
- 1 HLA-A 358
- 2 HLA-B 359
- 3 HLA-Cw
  - a Before correction for homozygosity and incomplete Cw6,7,8 typing 360
  - b After correction for homozygosity and incomplete Cw6,7,8 typing 361

3 Newly-diagnosed Type 1 diabetic cases ascertained between August 1978 and July 1980 but not included in the Barts-Windsor Family Study before 31st October 1980 362

4 Barts-Windsor Family Study: availability of data 363

5 Computer programs by the author, used in analyses for this dissertation: contents and introduction 364
- a JDAT3. Julian date/calendar date processing 366
- b SPK2X2. Two-by-two contingency table analysis 370
- c SPKRXC. R(ows)-by-C(olumns) table analysis 374
- d SPKRCL. Three-dimensional table analysis 376
- e HARDY1. Hardy-Weinberg equilibrium 378
- f TOMBOD. Haplotype sharing by sib pairs: dominant and recessive models (goodness of fit) 379

6 East Berks, 1978-1980: demographic and epidemiological data
- 1 Incidence of Type 1 diabetes, East Berkshire, August 1978 - July 1980 382
- 2 Type 1 diabetes: prevalence in East Berks districts, excluding Bracknell, adjusted to age distribution of BWFS siblings in Oct 1980 383
- 3 Absolute risk of Type 1 diabetes in HLA-genotyped siblings of affected probands, East Berkshire, 1978-1980 . 384

7 Papers by the author forming part of this dissertation . . . . . 385 (appended inside back cover)

XIII APPENDIX 1 DIABETIC PROBANDS IN THE BWFS
A: Age, sex, chromosome 6 marker genotype

IDF Note (i)	Sex (ii)	Age at entry (y) (iii)	Paternal haplotype HLA A (iv)	B	Cw (v)	Bf	C2	GLO (vi)	Maternal haplotype HLA A	B	Cw	Bf	C2	GLO	No. in Sibship: DM	total
1	1	9	24	39	-		C	2	1	8	7		C	1	1	2
2	1	18	2	62	3		C	2	1	50	6		C	2	1	2
3	2	15	3	65	8		C	1	2	62	3		C	1	1	3
4	1	18	1	8	7	S	C	3	25	44	5	S	C	4	1	2
5	1	13	30	57	6			2	2	44	5			2	1	6
6	2	14	2	8	7		C	2	3	60	3		C	1	1	3
7	2	14	2	62	3		C	2	29	44	=		C	2	1	2
8	2	9	24	44	5			2	26	64	8			2	1	5
9	2	14	33	65	=	S			29	44	=	F1			1	2
10	2	17	23	8	4			2	29	7	-			2	1	3
11	1	12	1	8	7		C	1	2	7	7		C	1	1	2
12	2	12	1	8	7	S	C	2	3	39	7	F	C	1	1	2
13	2	10	1	44	5	S		2	2	18	5	F1		2	1	3
14	2	13	11	35	4			1	11	51	=			1	1	3
15	2	14	2	18	5	F1		2	3	7	7	F		2	1	2
16	2	16	2	7	=	S		1	30	18	5	F1		2	1	3
17	1	18	24	60	3	S	C	1	31	38	=	S	C	2	1	4
18	1	13	24	62	3	S	C	2	2	44	5	S	C	2	1	3
19	1	17	2	44	5	S	C	1	1	8	7	S	C	1	1	3
20	1	24	24	60	2	S	C	2	2	44	5	S	C	2	1	4
21	1	15	3	18	5	F1		1	1	8	7	S		2	1	3
22	1	19	1	8	7	S	C	2	2	44	5	S	C	1	1	5
23	1	13	2	62	3	S	C	3	2	62	3	S	C	4	1	3
24	1	8	1	8	7		C	2	1	8	7		C	2	1	2
25	2	10	2	62	3	S	C	2	3	62	3	S	C	2	1	4
26	2	15	24	39	=	S		2	2	18	-	S		2	1	5
27	2	11	2	56	1			3	2	7	7		C	4	1	4
28	1	19	3	62	3	S	C	2	1	8	7	S	C	2	1	3
29	1	14	1	8	=				2	62	3				1	3
30	2	9	1	8	7		D	2	24	62	3		E	2	2	4
31	1	14	2	60	3			2	2	62	3			1	1	4
32	1	13	26	57	6		C	2	3	62	3		C	2	1	2
33	2	17	2	27	3			1	2	44	=			1	1	2
34	2	11	23	49	=	F1		2	2	51	-	S		1	1	2
35	2	9	31	60	3		C	2	2	39	=		C	2	1	3
36	1	15	1	27	2	S		2	24	7	7	F		2	1	3
37	1	10	31	62	3		B	1	3	7	7		C	1	1	2
38	2	15	24	8	7		C	2	2	63	-		C	2	1	2
39	1	14	3	27	2	S	C	2	2	62	3	S	C	2	1	4
40	1	11	3	14	=		C	2	2	62	3		C	1	2	4
41	1	16	2	62	3		C	2	2	62	3		C	2	1	4
42	2	14	1	62	3	F	B	1	1	8	7	S	C	2	3	4
43	2	15	3	60	3		C	2	1	8	-		C	1	1	5
44	2	15	1	8	7		C	2	2	62	3		C	1	1	2
45	1	16	2	60	3	S	C	1	28	8	7	S	C	1	1	3
46	1	10	1	62	3		C	1	1	8	-		C	1	1	3
47	1	18	1	8	=		C	2	2	62	3		C	1	1	2
48	1	7	2	47	6		C	2	1	60	3		C	2	1	2
49	2	14	11	51	-		C	1	1	8	7		C	2	1	2
50	2	5	2	57	6	S		3	24	39	7	S		4	1	2
51	2	13	11	35	4		C	2	1	8	=		C	1	1	3
52	2	19	11	55	3			2	2	62	3			1	1	3
53	1	13	2	60	3		D	1	2	50	6		E	1	1	3
54	1	14	24	39	=	S		2	2	44	5	S		1	1	3
55	1	24	1	62	3		C	2	1	8	7		C	2	1	4
56	1	12	1	8	=	S	C	2	24	39	=	S	C	2	1	2
57	2	16	3	60	3		C	2	2	51	=		C	1	1	3

Continued on next page

XIII APPENDIX 1 DIABETIC PROBANDS IN THE BWFS
A: Age, sex, chromosome 6 marker genotype

IDF	Sex	Age at entry	Paternal haplotype HLA A	B	Cw	Bf	C2	GLO	Maternal haplotype HLA A	B	Cw	Bf	C2	GLO	No. in Sibship: DM	total
58	1	16	3	62	3		C	2	11	44	5		C	2	1	2
59	2	11	1	8	7		D	1	3	60	3		E	1	1	4
60	1	20	1	57	6			1	31	44	4			2	1	3
61	2	25	25	18	=	F		2	2	8	=	S		2	1	6
62	2	16	2	39	7		C	1	1	8	7		C	2	1	5
63	1	9	2	62	3		C	2	2	62	3		C	2	1	2
64	1	11	2	62	3		C	2	1	8	7		C	1	1	3
65	1	16	24	8	7		C	2	1	7	7		C	1	1	3
66	1	15	2	39	5		C	2	1	8	7		C	1	1	4
67	2	15	11	44	5		C	1	1	8	7		C	2	1	3
68	2	16	1	7	=			1	29	44	=			1	1	3
69	1	19	1	8	=	S	C	2	2	8	7	S	C	2	1	3
70	1	20	23	8	7		C	1	30	35	4		C	2	1	8
71	1	7	2	13	=	S	C	2	2	8	=	S	C	2	1	2
72	2	18	2	7	7		C	1	28	35	4		C	2	1	2
73	1	17	24	18	5	F1		2	26	44	2	S		2	1	3
74	1	7	1	62	3		C	1	2	35	4		C	2	1	3
75	2	18	28	14	=		C	2	2	62	4		C	1	1	2
76	1	14	1	51	4		C	2	1	8	7		C	1	1	2
77	1	10	2	44	5			1	28	35	4			2	1	2
78	1	11	3	7	=			1	3	44	5			2	1	3
79	1	16	2	8	=		C	1	3	60	3		C	1	1	5
80	2	9	3	62	=	S	C	2	1	8	=	S	C	1	1	3
81	2	28	2	62	3		B	2	2	8	7		C	1	2	11
82	1	21	2	8	=		C	1	25	8	=		C	2	1	5
83	1	22	2	7	=	S	C	2	1	8	=	S	C	2	1	6
84	2	19	1	7	=			2	24	35	4			1	1	4
85	2	9	3	18	5	F1	C	1	2	51	=	S	C	2	1	3
86	2	14	1	8	=		C	2	2	44	5		C	1	1	2
87	1	11	3	35	4		C	2	1	8	=		C	2	1	3
88	2	22	2	8	7	S	C	1	24	39	7	S	C	1	1	2
89	1	19	2	39	=		C	2	1	8	=		C	2	1	2
90	1	14	28	39	=		C	1	2	8	=		C	1	1	2
91	1	11	1	8	=		C	3	1	13	=		C	4	1	2
92	1	21	2	8	=			1	2	51	2			2	1	3
93	1	7	11	60	3		C	2	1	41	1		C	2	1	3
94	2	21	2	62	3	S	C	1	1	8	7	S	C	1	1	4
95	1	20	2	50	=		C	2	1	8	=		C	2	1	5
96	1	19	26	37	6		C	2	2	62	3		C	1	1	3
97	1	16	24	27	1	S	C	1	1	8	7	S	C	2	1	4
98	1	9	2	44	5				32	7	=				1	2
99	2	22	1	8	=		C	3	2	62	3		C	4	1	3
100	2	15	11	55	3		C	1	3	44	5		C	1	1	3
101	1	5	2	27	1	S		2	3	18	-	S		2	1	2
102	1	20	1	8	7		D	2	2	62	3		E	2	1	2
103	1	10	3	62	3		C	1	1	8	7		C	2	1	3
104	2	13	1	8	=		C	1	2	8	=		C	1	1	2
105	1	11	1	8	=			2	31	60	3			1	1	2
106	2	29	1	8	=			1	2	27	2			1	1	3
107	2	15	25	44	5		C	2	2	62	3		C	2	1	3
108	2	19	3	44	2		C	2	2	62	3		C	2	1	3
109	1	11	11	27	1		C	2	2	62	3		C	2	1	3
110	2	13	3	51	3	S	C	1	1	8	=	S	C	1	1	3
111	1	22	2	44	-	F		2	24	55	3	S		2	1	2
112	1	12	1	8	7		C	1	11	60	3		C	2	1	3
113	2	7	11	7	7		C	1	1	8	7		C	1	1	2
114	2	11	29	18	5	F1		1	30	18	5	F1		1	1	2
115	1	9	1	8	7	S		1	2	62	3	S		1	1	4

(vii) (viii)

Continued on next page

XIII APPENDIX 1 DIABETIC PROBANDS IN THE BWFS
A: Age, sex, chromosome 6 marker genotype

IDF	Sex	Age at entry	Paternal haplotype						Maternal haplotype						No. in Sibship:	
			HLA			Bf	C2	GLO	HLA			Bf	C2	GLO		
			A	B	Cw				A	B	Cw				DM	total
116	1	9	1	8	=		C	1	2	62	3		C	2	1	2
117	2	17	32	60	3			1	1	8	=			2	1	2
118	2	13	2	14	-		C	2	2	60	3		C	1	1	2
119	1	9	1	44	7	F	C	3	29	44	-	S	C	4	1	2
120	1	5	24	8	7		C	2	2	60	3		C	2	1	2
121	1	16	2	50	6			2	29	44	-			2	1	3
122	2	19	24	62	1		C	1	1	8	=		C	2	1	3
123	2	12	1	8	=			1	1	62	3			1	1	2
124	1	11	1	62	3		D	2	1	8	=		E	2	1	2
125	1	19	2	7	=		C	2	32	60	3		C	2	1	2
126	2	10	2	8	=		C	1	28	60	3			2	1	2
127	1	18	3	44	-			3	23	60	3			4	1	4
128	2	4	24	39	7			1	32	14	=			2	1	3
129	1	10	30	8	=				28	55	3				1	2
130	1	12	24	44	2			1	2	44	5			1	1	2
131	2	17	2	44	5		C	1	29	8	=		C	1	1	2
132	1	14	1	8	7	S	C	2	24	39	-	S	C	1	1	3
133	2	12	1	8	7			2	3	35	4			2	1	4
134	1	9	3	8	=				2	60	3				1	2
135	1	12	2	62	3		C	1	3	7	7		C	2	1	2
136	2	14	31	51	=			1	26	7	5			2	1	2
137	2	9	31	7	=			1	24	35	=			1	1	3
138	2	7	24	62	3	S	C	2	30	18	8	F1	C	2	1	3
139	1	15	2	39	=				1	8	=				1	3
140	2	13	26	18	7	F		1	2	35	4	F		1	1	4
141	2	12	1	8	=	S	C	2	24	39	-	S	C	2	1	3
142	2	16	2	35	4	F		1	1	35	4	F		1	1	2
143	2	15	3	7	7	F	C	2	1	8	7	S	C	1	1	3
144	2	15	28	18	=	S	C	2	2	60	3	S	C	2	1	3
145	2	6	2	8	=		C	1	1	8	=		C	1	1	3
146	1	12	24	27	=	S	C	1	2	60	3	S	C	1	1	2
147	2	9	24	35	=	S		1	1	8	=	S		1	1	2
148	1	12	2	44	5	S		1	25	18	5	F1		2	1	2
149	1	16	1	7	=			2	28	71	7			2	1	2
150	1	9	31	18	5	F1	C	2	2	62	3	S	C	2	1	2
151	1	10	1	8	7		C	1	24	62	3		C	2	1	3
152	1	14	24	60	3		C	1	32	8	7		C	1	1	3
153	1	10	2	62	3		C	2	1	8	7		C	1	1	2
154	2	7	2	8	=	S	C	1	26	44	=	S	C	1	1	2
155	2	15	11	27	1	S		1	25	8	=	S		1	1	3
156	1	14	1	8	=		C	2	29	44	=		C	2	1	3
157	1	17	2	62	3		C	2	3	51	2		C	2	1	3
158	1	14	28	60	3		C	1	1	8	7		C	2	1	2
159	2	13	2	60	3			1	28	60	3			1	1	2
160	1	9	24	27	2			2	3	8	7			2	1	2
161	2	11	3	8	=		D	1	1	8	=		E	2	1	2
162	2	4	26	8	=			1	1	8	=			2	1	2
191	1	11	3	62	3				1	37	=				2	2
192	1	14	1	8	=				28	39	=				2	2
193	2	6	1	8	=			2	24	39	=			2	2	2
194	2	27	1	8	7		D	3	19	27	2		E	4	3	5

NOTES: (i) IDF is identity code number for family. (ii) 1 = male. (iii) Age is calculated to nearest birthday. (iv) HLA-A and -B specificites up to Bw63 were regularly determined; some of the newer 'splits' could be resolved only at times, or not at all, in which case parent specificities are given (see Figure III-4). (v) In cases where sera determining Cw 6, 7 and 8 were not available a 'blank' is coded ' ='; where such sera were used, blanks are coded '-'. (vi) C2 B/C with C/B, and GLO 2/1 with 1/2, indicate heterozygosity with equivocal haplotype assignment. (vii) These two probands are first cousins.

B: Clinical features and family history of diabetes

IDF	At diagnosis:			Duration of DM at entry to study (y) (ix)	Maximum ketones recorded (x)	DM in relatives (xi)			
Note (i)	Symptoms (wks) (viii)	Age (y) (ii)	Calendar month			Sibs	Parents	2nd degree	3rd degree
1	3	4	8	5	4	0	0	1	0
2	3	2	10	16	3	0	0	0	0
3	2	3	9	3		0	0	1	0
4	6	3	9	15	3	0	0	2	0
5	50	13	2	2/12		0	0	0	0
6		4	4	10	4	0	0	2	0
7	8	13	3	1	2	0	0	1,2	0
8	8	3	11	6	4	2	0	0	0
9	12	14	1	3/12		0	0	1	0
10	1	12	4	5	3	0	0	1	0
11	6	10	1	1		0	0	0	0
12		10	12	3	3	0	0	0	0
13	2	5	7	5	4	0	0	0	0
14	4	4	1	8		0	1	0	0
15	4	11	3	3	3	0	2	2	0
16	50	13	1	3	3	0	0	0	0
17	1	15	1	3		0	0	0	0
18	2	2	11	10	2	0	1	1	0
19	1	12	2	5	2	0	2	0	0
20	1	16	4	8	9	0	0	0	2
21	2	15	2	3/12	3	0	1	0	1
22	4	19	1	4/12	3	0	0	0	0
23	4	6	3	8	0	0	0	0	0
24	1	6	10	7/12	0	0	1	2	1
25	1	3	5	7		0	0	1	0
26	16	12	5	3	4	0	0	0	0
27	1	3	1	7		0	0	0	0
28	3	15	11	5	1	0	1	0	0
29	4	10	4	4		0	0	0	1
30	1	3	5	6	9	1	0	1	2
31	13	13	3	1	4	0	0	0	0
32	1	4	5	9	1	0	0	0	2
33		12	1	5	3	0	0	0	0
34		5	11	6	3	0	0	0	0
35	26	8	6	1	2	0	0	0	0
36	2	15	4	2/12	4	0	0	0	0
37	1	9	10	8/12	4	0	1	0	0
38	4	1	6	14	4	0	0	0	2
39	1	12	3	2		0	0	0	0
40	1	5	12	6	4	1	0	0	0
41	50	4	11	12	4	0	0	0	0
42		9	1	5	9	1	0	0	0
43	15	13	4	2	4	0	0	0	0
44	6	10	7	5	2	0	0	0	0
45	1	3	12	13	4	0	0	0	0
46	4	8	4	2	3	0	0	0	0
47	10	18	4	2/12	2	0	0	0	3
48	4	3	2	4	3	0	0	0	0
49	4	12	9	3	3	0	0	2	3
50	4	4	10	2	3	0	0	0	0
51		11	5	2	4	0	0	0	0
52	4	7	2	12	4	0	0	0	2
53	2	11	11	2		0	0	0	1
54	4	13	11	8/12	3	0	0	0	0
55	1	3	12	11		1	0	0	0
56	1	3	6	9	3	0	0	0	2
57	10	13	3	3	3	0	0	0	1

Continued on next page

B: Clinical features and family history of diabetes

	IDF	At diagnosis:			Duration	Maximum	DM in relatives			
		Symptoms (wks)	Age (y)	Calendar month	of DM at entry (y)	ketones recorded	Sibs	Parents	2nd degree	3rd degree
	58	20	16	5	2/12	3	0	0	0	0
	59	8	4	9	7		0	0	0	0
	60	1	15	4	5	4	0	0	0	0
	61	1	13	2	12	4	0	0	0	0
	62	1	10	12	7	3	0	0	1	1
	63	11	3	7	6		0	0	0	0
	64	6	10	5	1	4	0	0	0	1
	65	13	5	11	12	4	0	0	0	0
	66	5	11	8	4	0	0	0	0	0
	67	90	4	3	6		0	0	0	0
	68	12	8	11	7	4	0	0	1	0
	69		1	2	17	4	0	0	0	1
	70	4	4	1	16	4	0	0	0	0
	71	8	3	7	4	4	0	2	0	0
	72	4	14	8	4	4	0	0	0	2,1
	73		14	9	3	3	0	0	2	2
	74	7	3	11	4		0	0	0	1,2
	75	39	3	7	15	3	0	0	0	1,2
	76	34	8	3	5		0	0	0	0
	77	30	9	7	1	4	0	0	1	0
	78	2	5	1	7	4	0	0	0	1,2
	79	4	13	3	2		0	0	0	0
	80	1	7	11	2	3	0	0	0	2
	81	100	16	4	11	3	1	0	0	1
	82	2	14	11	7	4	0	0	0	0
	83	1	10	4	12	4	0	0	2	2
	84	50	9	12	11	3	0	0	0	0
	85	16	4	11	5		0	0	0	2
	86	4	7	12	6	4	0	0	0	0
	87	12	5	10	6	4	0	0	0	0
	88	1	3	10	19		1	0	0	2
	89	10	8	10	11	3	0	0	0	2
	90	2	11	1	4	3	0	0	0	0
	91	2	8	3	2	3	0	0	0	2
	92	30	16	3	5	3	0	0	0	3
(vii)	93	3	5	6	1	4	0	0	0	1
	94	4	18	2	4	4	0	0	0	0
	95	2	16	1	5		0	0	0	0
	96	3	13	3	6	4	0	0	0	2
	97	0	5	6	11	0	0	0	1	0
	98	6	5	8	4	4	0	0	1,2	0
	99		4	12	18		0	1	1	1
	100	8	13	9	2	3	0	0	0	1
	101	5	2	7	2	4	0	0	0	0
	102	3	13	2	8		0	0	0	0
	103	6	9	9	1		0	0	1	2
	104	50	10	4	3	2	0	0	0	0
	105	4	8	7	3	3	0	0	1	1,2
	106	8	12	5	17		0	1	1,2	2
	107	1	10	8	5	4	0	0	2	0
	108	25	15	9	4	3	0	0	0	0
	109	1	7	12	5	4	0	0	0	2
	110	6	13	11	0	2	0	0	0	0
	111	2	6	12	16	4	0	0	0	0
(vii)	112	22	6	12	6		0	0	0	1
	113	1	5	3	3	3	0	0	0	0
	114	25	7	8	4	4	0	0	0	3
	115	25	4	9	5	4	0	0	0	1

Continued on next page

B: Clinical features and family history of diabetes

IDF	At diagnosis:			Duration	Maximum	DM in relatives			
	Symptoms (wks)	Age (y)	Calendar month	of DM at entry (y)	ketones recorded	Sibs	Parents	2nd degree	3rd degree
116	25	6	3	3	4	0	0	0	0
117	4	4	5	13	4	1	0	2	2
118	1	8	9	5	4	0	0	2	1,3
119	1	1	11	8	4	0	0	0	0
120	4	4	5	1	4	0	0	0	2
121	22	4	2	12		0	0	1,2	3
122	1	5	9	14		0	0	0	0
123	4	12	1	11/12	4	0	0	0	0
124	1	5	3	6	1	0	0	0	0
125	50	15	8	4	4	0	0	0	1,2
126	6	5	3	5	4	0	0	0	0
127		18	12	0		0	0	0	0
128		4	11	0	4	0	1	2	2
129	1	7	12	3	4	0	0	0	1
130		11	3	9	0	0	0	0	0
131	4	10	2	7	3	0	0	0	0
132	8	9	6	6	3	0	0	1	0
133	26	5	1	7	3	0	0	0	0
134	5	8	12	1	2	0	0	0	0
135		12	10	3/12	2	0	0	0	1
136		11	7	3		0	0	1	2
137		9	12	0	3	0	0	1	2
138	26	4	6	4	1	0	0	0	0
139		9	3	6	3	0	0	0	2
140	20	8	10	5	3	0	1	1	1
141	26	7	2	5	3	0	0	0	0
142	4	2	4	14	4	0	0	0	0
143	4	9	4	7	3	0	0	0	0
144	1	13	12	2	3	0	0	0	0
145	1	2	7	4	1	0	0	0	1
146	3	11	5	9		0	0	0	2
147	8	7	6	2	3	0	0	2	3
148	2	11	3	11	4	0	0	0	0
149		9	3	7	3	0	0	0	0
150	3	0	1	0		0	0	0	0
151	1	10	1	0	4	0	0	0	2
152	3	7	9	7	3	0	0	0	1,2
153	2	2	7	8	4	0	0	0	0
154	2	6	6	2		0	0	0	0
155	4	12	6	3	3	0	0	0	0
156	2	13	4	1	2	0	0	2	0
157	6	17	2	0		0	0	0	1
158	3	14	12	4/12	4	0	3	3	3
159	2	13	2	0	3	0	0	2	1,2
160	10	9	3	0	3	0	0	2	2
161	17	11	4	0	3	0	0	0	0
162	8	4	4	0	4	0	0	0	2
191	1		8		3	1	0	0	0
192	8		11		4	1	0	0	0
193	2		5			1	0	0	0
194	0		8			1	0	0	0

Note: (i - vii) see Appendix 1-A. (viii) Time (in weeks) from first of symptoms clearly leading up to diagnosis. 0 = asymptomatic. (ix) Time from diagnosis to ascertainment, to nearest full year or, if less than 1y, to nearest month. (x) Maximum degree of ketonuria recorded in hospital records: 0 = no ketonuria (records from first admission inspected); 1 = +, 2 = ++, 3 = +++, 4 = ketoacidosis (plasma bicarbonate < 18mmol/l), blank = no record of tests for ketonuria found, or records from first admission missing and later records negative (in many of these cases parents recalled ketosis at diagnosis). (xi) Family history: 0 = no diabetes, 1 = Type 1, 2 = Type 2, 3 = type not known.

BARTS-WINDSOR FAMILY STUDY: GENETIC ANALYSES

1. HLA-A IN PROBANDS OF BWFS, LONDON CONTROL SUBJECTS AND FAMILY CONTROL HAPLOTYPES

	PHENOTYPES					GENE COUNTS corrected for homoz'y				
HLA	Probands positive		Control * subjects positive		Fisher exact test:	Probands' haplotypes positive		Control * haplotypes positive		Fisher exact test:
	No.	%	No.	%	2p	No.	Freq.	No.	Freq	2p
n=	165		170			330		307		
A 1	81	49	58	34	0.008	89	0.2697	48[e]	0.1564	<0.001
A 2	92	56	93	55	NS	102	0.3091	84[f]	0.2736	NS
A 3	34	21	37	22	NS	35	0.1061	44[g]	0.1433	NS
A 11	11	7	17	10	NS	12	0.0364	19	0.0619	NS
Aw19[a]	1	1	0	0		1	0.0030	0	0.0000	
Aw23	4	2	6	4		4	0.0121	5	0.0163	
Aw24	32	19	23	14	NS	13	0.0970	26[h]	0.0847	NS
A 25	6	4	5	3		0	0.0182	7	0.0228	NS
A 26	8	5	7	4	NS	2	0.0242	11	0.0358	NS
A 28	12	7	11	6	NS	2	0.0364	19[i]	0.0619	NS
A 29	9	6	16	9	NS	0	0.0273	16	0.0521	NS
Aw30	6	4	6	4		2	0.0182	7[j]	0.0228	NS
Aw31	8	5	9	5	NS	2	0.0242	9	0.0293	NS
Aw32	5	3	12	7	NS	0	0.0152	9	0.0293	NS
Aw33	1	1	3	2		0	0.0030	3	0.0098	
Aw34	0	0	0	0		0	0.0000	0	0.0000	
Aw36	0	0	0	0		0	0.0000	0	0.0000	
Aw43	0	0	0	0		0	0.0000	0	0.0000	
A'-'[b]	0	0	-	-		0	0.0000	0	0.0000	
A'x'[c]	20	12	37	22		-		-		
TOTAL	330	202	340	201		90	1.0001	307	1.0000	
A'x'[d]	-	-	-	-		0	0.0000	26	0.0850	

NOTES

For relative risks and population attributable risks see Table VIII-1.

* Controls for phenotype and gene counts are from different populations.

a Available sera could not distinguish between A32 and A33 in one family; hence categorization as A19 (the broad antigen concerned).

b True serological blank confirmed in at least two family members with different HLA-A phenotypes.

c Other blanks (phenotype counts). In probands, all these were confirmed to be due to homozygosity.

d Other blanks (gene counts): counted as a second (homozygous) occurrence of the specificity on the paired haplotype, as detailed below:

e Including 10 'x' paired with A1; h Including 1 'x' paired with A 24

f Including 8 'x' paired with A2; i Including 1 'x' paired with A 28

g Including 5 'x' paired with A3; j Including 1 'x' paired with A 30

BARTS-WINDSOR FAMILY STUDY: GENETIC ANALYSES

2. HLA-B IN PROBANDS OF BWFS, LONDON CONTROL SUBJECTS AND FAMILY CONTROL HAPLOTYPES

	PHENOTYPES					GENE COUNTS corrected for homoz'y				
HLA	Probands positive		Control subjects positive		Fisher exact test:	Probands' haplotypes positive		Control haplotypes positive		Fisher exact test:
	No.	%	No.	%	2p	No.	Freq.	No.	Freq	2p
n=	165		170			330		307		
B 7	21	13	45	27	0.002	21	0.0638	50[b]	0.1629[b]	10^{-4}
B 8	89	54	46	27	10^{-6}	96	0.2909	37[b]	0.1205	$<10^{-6}$
B 13	2	1	8	5	NS	2	0.0061	7	0.0228	
B 14	7	4	13	8	NS	7	0.0212	12[b]	0.0391[b]	NS
B 18	15	9	15	9	NS	16	0.0485	12	0.0391	NS
B 27	11	7	14	8	NS	11	0.0333	17	0.0554	NS
Bw35	14	8	25	15	NS	15	0.0455	22	0.0717	NS
B 37	2	1	4	2		2	0.0061	5	0.0163	
Bw38	1	1	6	4		1	0.0030	2	0.0065	
Bw39	18	11	7	4	0.029	18	0.0547	5	0.0163	0.019
Bw41	1	1	0	0		1	0.0030	2	0.0065	
Bw42	0	0	0	0		0	0.0000	0	0.0000	
Bw44	33	20	50	30	(0.061)	35	0.1061	56[b]	0.1824[b]	0.008
Bw45	0	0	1	1		0	0.0000	3	0.0098	
Bw46	0	0	0	0		0	0.0000	0	0.0000	
Bw47	1	1	1	1		1	0.0030	2	0.0065	
Bw48	0	0	0	0		0	0.0000	0	0.0000	
Bw49	1	1	3	2		1	0.0030	4	0.0130	
Bw50	4	2	6	4		4	0.0121	2	0.0065	
Bw51	10	6	15	9	NS	10	0.0303	15	0.0489	NS
Bw52	0	0	1	1		0	0.0000	4	0.0130	
Bw53	0	0	2	1		0	0.0000	0	0.0000	
Bw54	0	0	0	0		0	0.0000	0	0.0000	
Bw55	4	2	4	2		4	0.0121	4	0.0130	
Bw56	1	1	0	0		1	0.0030	1	0.0033	
Bw57	4	2	13	8	(0.0504)	4	0.0121	12	0.0391	0.045
Bw58	0	0	2	1		0	0.0000	3	0.0098	
Bw59	0	0	0	0		0	0.0000	0	0.0000	
Bw60	26	16	23	14	NS	27	0.0818	11	0.0358	0.028
Bw61	0	0	6	4		0	0.0000	2	0.0065	
Bw62	47	28	13	8	10^{-6}	51	0.1545	13	0.0423	$<10^{-6}$
Bw63	1	1	0	0		1	0.0030	1	0.0033	
Bw71	1	1	1	1		1	0.0030	3	0.0098	
Bw72	0	0	1	1		0	0.0000	0	0.0000	
B'-'	0	0	-	-		0	0.0000	0	0.0000	
B'x'	16	10	15	9	NS	-		-		
TOTAL	330	201	340	206		330	1.0001	307	1.0001	

NOTES (SEE ALSO NOTES TO APPENDIX 2-1)

ab Probable but unconfirmed homozygous HLA-B 'blanks' are included in the above figures since no confirmed blanks were seen:

a One haplotype in a male proband and father, both typing B8,x;

b Ten 'control' haplotypes in parents typing as follows (fathers, mothers) B7,x (2, 0); B8,x (3, 3); B14,x (1,0); B44,x (0,1).

BARTS-WINDSOR FAMILY STUDY: GENETIC ANALYSES

3(a). HLA-Cw IN PROBANDS OF BWFS, LONDON CONTROL SUBJECTS AND FAMILY CONTROL HAPLOTYPES: OBSERVATIONS BEFORE CORRECTION FOR HOMOZYGOSITY AND INCOMPLETE Cw6,7,8 TYPING

	PHENOTYPES				GENE COUNTS			
HLA	Probands positive		Control subjects positive		Probands' haplotypes positive		Control haplotypes positive	
	No.	%	No.	%	No.	Freq.	No.	Freq
N	165		170		330		306	
Cw 1	7	4	7	4	7	0.0212	13	0.0425
Cw 2	12	7	16	9	12	0.0364	14	0.0458
Cw 3	74	45	41	24	81	0.2455	27	0.0882
Cw 4	15	9	29	17	16	0.0485	26	0.0850
Cw 5	30	18	27	16	32	0.0970	34	0.1111
Cw"="[a]	40	24	-	-	52	0.1575	58	0.1895
Cw 6	11		32	19	11		19	
Cw 7	82		93	55	94		51	
Cw 8	5	72	0	0	5	0.3939	5	0.4379
Cw"-"[b]	15		-	-	15		39	
Cw"x"[c]	5		95	56	5		20	
Homzgs[d]	34	21	-	-				
TOTAL	330	200	340	200	330	1.0000	3 306	1.0000

NOTES (SEE ALSO NOTES TO APPENDIX 2-1)
A proportion of probands and families in the BWFS were not tested for Cw6-8 (see text, VIII-1).

a Blanks, confirmed in other family members, but not typed for Cw6-8
b As a, but typed for Cw6-8
c 'Probable-homozygous' blanks, unconfirmed in other family members
d Confirmed homozygous 'blanks' (whether ot not typed for Cw6-8).

BARTS-WINDSOR FAMILY STUDY: GENETIC ANALYSES

3(b). HLA-Cw IN PROBANDS OF BWFS, LONDON CONTROL SUBJECTS AND FAMILY CONTROL HAPLOTYPES: FREQUENCIES AFTER CORRECTION FOR HOMOZYGOSITY[a] AND INCOMPLETE Cw6,7,8 TYPING[b]

	PHENOTYPES					GENE COUNTS				
HLA	Probands +ve (assuming HWE)[c]		Control subjects positive		Fisher exact test:	Probands' haplotypes positive		Control haplotypes positive		Fisher exact test:
	No.	%	No.	%	2p	No.	Freq.	No.	Freq	2p
N	165		170			330		306		
N' [b,c]	117.5[c]		170			234.9		210.5		
Cw 1	7	4	7	4	NS	7.0	0.0212	13.0	0.0425	NS
Cw 2	12	7	16	9	NS	12.0	0.0364	14.0	0.0458	NS
Cw 3	71	43	41	24	10^{-4}	81.0	0.2455	27.0	0.0882	$<10^{-6}$
Cw 4	16	9	29	17	0.045	16.0	0.0485	26.0	0.0850	(0.090)
Cw 5	31	19	27	16	NS	32.8	0.0994	35.6	0.1163	NS
Cw 6	11[c]	9	32	19	0.038	11.0[b]	0.0468	20.4[b]	0.0968	(0.060)
Cw 7	77[c]	66	93	55	(0.086)	97.7[b]	0.4159	61.1[b]	0.2902	0.008
Cw 8	5[c]	4	0	0		5.0[b]	0.0213	5.7[b]	0.0271	
Cw'-'	-	-	-	-		15.3	0.0651	43.8	0.2080	
Cw'x'	64	39	94	55	NS	-	-	-	-	
TOTAL	200		199			1.0001		0.9999		

NOTES (SEE ALSO NOTES TO APPENDIX 2-1)

a Correction of gene-counts for homozygosity: 'x' was distributed between paired specificity and '-' or '=' according to the overall distribution; see text, VIII-1-B-i.

b Correction for incomplete typing: for gene frequency, N' (the nominal total of haplotypes typed for Cw6-8) is calculated according to the formula

$$N' = \frac{N}{\left[1 + \frac{n_-}{n_- + n_6 + n_7 + n_8}\right]}$$

where n_a is the number of haplotypes counted positive for specificity 'a' after correction for homozygosity.

c Phenotype frequency F is calculated from gene frequency by assuming Hardy-Weinberg equilibrium (III-6-B-ii), the nominal total (N') of subjects typed for Cw6-8 is taken as half that for haplotypes, and the nominal count of subjects positive for the specificity concerned is derived indirectly as F.N'' for Cw6-8, or as F.N for Cw1-5.

NEWLY-DIAGNOSED TYPE 1 DIABETIC CASES ascertained between August 1978 and July 1980 but not included in the Barts-Windsor Family Study before 31st October 1980

IDF	Sex	Age	HLA phenotype						Calendar month	Diabetes in relatives:	
			A		B		Cw			1st degree	2nd degree
512	2	17	2	28	35	44	4	x	12	0	0
519	1	13	2	23	44	98	5	x	2	0	0
521	1	14	2	23	8	x	=	=	4	0	0
525	2	19	1	2	8	60	7	3	5	0	0
527	1	16	23	28	50	x	6	x	6	0	0
528	1	13	1	2	8	44	7	5	6	0	0
530	2	7	2	24	61	8	2	7	6	0	0
534	1	8	1	24	17	39	=	=	10	0	0
535	1	17	1	2	8	51	7	x	10	0	0
536	1	7	3	24	14	18	5	x	10	0	0
540	2	11	1	29	14	44	=	=	12	0	0
541	2	8	2	3	7	18	5	x	12	0	1
542	2	9	2	2	62	27	3	2	12	0	0
543	2	10	2	1	7	7	7	7	12	0	0
544	1	9	1	24	8	35	4	7	12	0	0
547	1	15	2	2	39	62	-	3	1	0	0
549	2	14	31	2	60	18	3	7	1	0	0
550	1	10	1	19	8	18	5	7	2	0	1
551	1	14	1	29	8	62	7	x	2	0	0
553	1	16	3	11	8	62	3	x	2	0	1
555	1	14	3	28	35	40	3	4	3		
556	1	10	24	29	39	44	=	=	3	1	0
558	1	14	1	3	8	7	7	7	4	0	0
562	1	14	1	3	8	62	3	x	5	0	0
563	1	8	1	2	8	62	3	x	5	0	0
564	2	9	24	x	40	x	3	x	5	0	0
566	1	8	1	3	8	40	3	x	6		
567	1	14	2	3	7	44	5	x	6	0	2
573	1	14	2	2	62	35	3	4	7	0	0
574	2	19	25	31	8	60	7	3	3	1	1

NOTES: See Appendix 1. x = blank (possible homozygote) not clarified by genotyping.

BARTS-WINDSOR FAMILY STUDY:

AVAILABILITY OF DATA

A requirement of the Data Protection Act (1984) means that the original data on which this work is based may not be published in any form which would reveal confidential information and would allow identification of families from, for example, sex and dates of birth of individuals. Partly for this reason, and also because of the sheer volume of information concerned (see VI-3-A-v; VI-3-B-vii-c), full details on individual families are not given here. (Non-identifiable details on diabetic probands are given in Appendices 1 and 3.)

Bona-fide research workers requesting access to appropriate classes of data may make formal application in writing to the Steering Committee of the Barts-Oxford Family Study (the expanded derivative of the Barts-Windsor study), care of:

Dr. Edwin Gale,
Dept of Diabetes and Immunogenetics,
St. Bartholomew's Hospital,
West Smithfield,
London EC1A 7BE

Telephone 01-606 2032

COMPUTER PROGRAMS

CONTENTS

XIII-5a JDAT3. Julian date/calendar date processing · · · · · · · · 366
Conversion between 'julian' date, i.e. number of days from an arbitrary zero (here 31:12:1600), and calendar date by the current Gregorian calendar, with full allowance for leap years including that of 2000 AD. Day of week. Date of Easter.

5b SPK2X2. Two-by-two contingency table analysis · · · · · · · 370
Approximate Relative Risk (RR), population attributable risk (δ), X^2 with and without Yates' correction, Fisher's exact test, and, if required, predictive value and associated statistics. Output option to file or printer as well as screen.

5c SPKRXC. R(ows)-by-C(olumns) table analysis · · · · · · · · · 374
X^2 for independence and for linear trend; Goodman and Kruskaul's Gamma statistic.

5d SPKRCL. Three-dimensional table analysis · · · · · · · · · · 376
Expected values, degrees of freedom and X^2 for mutual independence and partial independence models.

5e HARDY1. Hardy-Weinberg equilibrium · · · · · · · · · · · · · 378
Given number of subjects tested, phenotype count and gene count, computes gene frequencies and numbers of homozygotes and heterozygotes, both observed and as predicted by Hardy-Weinberg law. Tests for deviation from H-W equilibrium by X^2.

5f TOMBOD. Haplotype sharing by sib pairs: dominant & recessive models (goodness of fit) · · · · · · · · 379
Thomson and Bodmer (1977) method to compute goodness-of-fit X^2 at various gene frequencies for dominant and recessive models given numbers of affected (e.g. diabetic) sib-pairs concordant for both, one or neither (HLA) haplotype. Numerical output may be directed to screen, file or printer. Optional graphical output to screen and printer.

COMPUTER PROGRAMS

ABOUT THESE PROGRAMS

All are presented as BASIC programs, written in the Microsoft GW-Basic dialect which is almost identical to IBM BASIC-A. All are self-explanatory in use. All should run on any IBM-PC compatible personal computer, although certain types of display adaptors will not show the graphical part of output from TOMBOD (a CGA or compatible adaptor is required for this; a Hercules monochrome graphics adaptor may also be used, with a CGA emulator program). Adaptation to run on different machines with different dialects of BASIC should not be difficult for a moderately experienced programmer, and the program listings are sufficiently well-documented to permit this.

Detailed instructions for running the programs will vary according to the configuration of hardware in use. A copy of GWBASIC or BASICA is necessary. If printouts of the graphs produced by TOMBOD are required, the program GRAPHICS.COM or GRAPHICS.EXE (supplied with MS-DOS or PC-DOS) should first be loaded, and TOMBOD is best run direct from a DOS command line of the form "GWBASIC TOMBOD". Otherwise the GWBASIC screen editor menu will appear on the printed graphs.

These programs may be freely copied, used and distributed, but they may not be sold for gain and remain the copyright of the author. In no event will the author be liable for any damages arising out of use or attempted use of the programs.

The author will on request, and so far as is feasible, supply copies on 5¼" floppy diskettes (double-sided, double-density, 48 tracks per inch, 9 sectors - a standard IBM-PC format) in exchange for reasonable costs to cover media, postage and packaging.

COMPUTER PROGRAM JDAT3

```
1   REM:*********************************************
2   REM:**                  JDAT3                    **
3   REM:** JULIAN DATE / CALENDAR DATE PROCESSING **
4   REM:**                                           **
5   REM:**          Written in Casio BASIC           **
6   REM:**     for Casio FX-702P pocket computer     **
7   REM:**                April 1983                **
8   REM:**                                           **
9   REM:**    Translated to GWBASIC for IBM-PC       **
10  REM:**   compatible computers, February 1988     **
11  REM:**                                           **
12  '   **     JDAT3 (C) Andrew Gorsuch              **
13  REM:**                                           **
14  REM:**                                           **
15  REM:*********************************************
16  REM:
17  REM:
18  REM:    In the FX-702P the total program space
              may be divided into up in
19  REM:       to ten 'program areas', #0 - #9,
              each of which may be
20  REM:       used (as here) for a separate
              subroutine under the
21  REM:       control of one of the other program
              areas.
22  REM:    In this GWBASIC version these high-level
              subroutines start at
23  REM:     1000,  5000, 10000... 45000.
24  REM:    Here the master control program (0)
              includes a menu giving
25  REM:       four options, coded in areas beginning
              5000, 10000, 15000 and 40000
26  REM:
27  REM:       1) Conversion of calendar date to
              Julian date (here
28  REM:       arbitrary zero of 31 December 1600),
              and also day of
29  REM:       week.
30  REM:       2) Conversion of Julian date,
              defined as in 1), to
31  REM:       calendar date and day of week.  The
              result is verified
32  REM:       automatically by back-checking
              through 1).
33  REM:
34  REM:       3) Calculation of interval between
              two calendar dates
35  REM:       (result given in days and in weeks
              + days).
36  REM:
37  REM:       4) Calculation of the date of Easter.
38  REM:
39  REM:    Areas beginning 20000 - 35000 and 45000
              are used for subroutines under
40  REM:       the control of one or more areas
              beginning 5000 - 15000 and 40000
41  REM:
42  REM:    Full allowance is made for leap years
              according to the
43  REM:       Gregorian calendar as currently
              defined. This came into
44  REM:       use in some European countries late
              in the 16th century,
45  REM:       and in UK in September 1752.  Dates
              earlier than 1/1/1601
46  REM:       are not permitted here.
47  REM:    To facilitate translation into
              different BASIC dialects and
48  REM:       other languages, REM statements are
              included (although
49  REM:       not permissible in the FX-702P),
              dimensions of variables
50  REM:       are defined and their status with
              respect to subroutines
51  REM:       is stated.
52  REM:    All lines including REM: may be removed
              to speed processing.
53  REM:
54  REM:    VARIABLES:
55  REM:    REAL NUMBER VARIABLES R(n)
              (Double-precision in GWBASIC)
56  REM:    (n) refers to number of digits required
              in formatting
57  REM:       display; internal computations may
58  REM:       involve larger integers or decimals
59  REM:       accurate to six places.
60  REM:    DIMENS. NAME  DEFINITION
61  REM:    R(6)    A#    Julian date relative to
                          arbitrary zero
62  REM:    R(6)    B#    Previous value of A# (used
                          in interval
63  REM:                  calculations and
                          approximations)
64  REM:    R       C#    No. of previous century
                          years (eg 1700) since
65  REM:                  arbitrary zero
66  REM:    R       E#    Day of year
67  REM:    R(6)    G#    Difference between two
                          Julian dates
68  REM:    R       H#    Day of year (stored value
                          of E#)
69  REM:    R       L#    No. of previous years
                          divisible by 4 since
70  REM:                  arbitrary zero
71  REM:    R       N#    Derivative or precursor of
                          M, in various
72  REM:                  calculations; eg in #5,
                        N# = no. of completed
73  REM:                  2-month periods in
                          current year.
74  REM:    R       S#    Mean solar year in days,
                          i.e. 365.2422(0)
75  REM:                  (constant)
76  REM:    R(4)    Y#    Year - either AD or
                          relative to 1600
77  REM:
78  REM:    INTEGER VARIABLES (I) - value <32,000
79  REM:
80  REM:    DIM     NAME  DEFINITION
81  REM:    I(2)    D%    Date of month
82  REM:    I       F%    Day of year (as E# but in
                          different subroutine)
83  REM:    I(2)    M%    Month of year
84  REM:    I       P%    No. of days in most recent
                          odd-numbered month
85  REM:                  of current year, ie 31
                          (Jan-July) or 30
86  REM:                  (Sep, Nov)
```

COMPUTER PROGRAM JDAT3

```
87  REM:   I       R%    Pointer from menu to
                           program area pointers
88  REM:
89  REM:   I       T%    Used in leap-year
                           calculations (value 1 in
90  REM:                   leap years, 2 otherwise)
91  REM:   I       V%    Remainder when A# divided
                           by 7 (used for day of
92  REM:                   week calculation in #9)
93  REM:   I(5)    W%    No. of completed weeks in
                           an interval of time
94  REM:   I       Z%    Used as common variable
                           for year in leap-year
95  REM:                   calculations (#7)
96  REM:
97  REM:   CHARACTER VARIABLES (STR = string, B =
                           single-byte)
98  REM:   B       R$    Character version of R%
                           (entered from menu)
99  REM:   STR(9)  U$    Day of week (eg
                           "Wednesday")
100 REM:   B       Q$    Used for Yes/No answers
                           ("Y" or "N")
910 REM:
990 REM:   1000   MASTER CONTROL PROGRAM and Menu
991 REM:
992 REM:   VARIABLES
993 REM:        INITIALIZED S#
994 REM:        READ/INPUT  Q$
995 REM:        INTERNAL    Q$ R%
996 REM:        COMMON - to higher levels  (NONE)
997 REM:        COMMON - to lower levels   S#
998 REM:
1000      'MASTER CONTROL PROGRAM
1010  CLEAR:DIM A%(20),B%(20):S#=365.2422#
1020  R%=0:CLS:PRINT:PRINT
        "JDAT3 (Date conversion and manipulation)":
        PRINT
1028  REM:   Menu
1030  PRINT "      OPTIONS AVAILABLE":PRINT
1031  PRINT"1.   Conversion from dd mm yy to julian
        form, giving day of week"
1032  PRINT"2.   Conversion from julian form to dd
        mm yyyy"
1033  PRINT"3.   Calculation of interval between
        two dates"
1034  PRINT"4.   Date of Easter in any year"
1035  PRINT"5.   Stop"
1036  PRINT:PRINT:PRINT"OPTION? "
1040  R$=""
1042  R$=INKEY$: IF R$="" THEN 1042
1048  REM:   SUBROUTINE Check validity of R$
1050  IF R$ <> "1" AND R$ <> "2" AND R$ <> "3" AND
        R$ <> "4" AND R$ <> "5" THEN GOSUB 2100:
        GOTO 1040
1055  R%=VAL(R$)
1060  CLS:ON R% GOSUB 5000,10000,15000,40000,4000
1067  REM:   Repeat option (Y/N using INKEY$ allows
        single-key answer);
1070  Q$=""
1080  PRINT:PRINT "SAME AGAIN?"
1090  Q$=INKEY$: IF Q$="" THEN 1090
1100  IF Q$="N" OR Q$="n" THEN 1010
1110  IF Q$ <> "Y" AND Q$ <> "y" THEN GOSUB 2100:
        GOTO 1070
1120  GOTO 1060
2099  REM:   SUBROUTINE Message after invalid entry
2100  PRINT "TRY AGAIN"
2120  RETURN
4000  END
5000  '      SUBROUTINE #1  CONTROLLING CONVERSION
        FROM CALENDAR DATE  D% M% Y#(AD) TO A#,
        i.e. JD(1600)
5001  REM:   VARIABLES
5002  REM:        INITIALIZED HERE (none)
5003  REM:        READ/INPUT HERE  (none)
5004  REM:        DISPLAYED HERE   A# Y# D% M U$
5005  REM:        INTERNAL         (none)
5006  REM:        COMMON - to higher levels A# Y#
        D% M% U$
5007  REM:        COMMON - to lower levels  A# Y#
        D% M% U$
5008  REM:
5009  REM:   Enter D% M% Y#
5010  GOSUB 30000: Y#=Y#-1600
5019  REM:   Compute JD(1600)
5020  GOSUB 20000
5029  REM:   Compute day of week
5030  GOSUB 45000
5039  REM:   Display
5050  PRINT:PRINT "JD(1600) =";A#
5060  PRINT:PRINT U$;D%;M%;Y#+1600
5070  RETURN
5099  REM:
10000 '      SUBROUTINE #2 CONTROLLING CONVERSION
        FROM JD(1600) TO CALENDAR DATE (AD)
10001 REM:
10002 REM:   VARIABLES
10003 REM:        INITIALIZED HERE none
10004 REM:        READ/INPUT HERE  A#
10005 REM:        DISPLAYED HERE   Y# D% M% U$
10006 REM:        INTERNAL         none
10007 REM:        COMMON - to higher levels  S# Y#
        D% M% U$
10008 REM:        COMMON - to lower levels   A# S#
        Y# D% M% U$
10009 REM:   Enter JD(1600)
10010 PRINT:INPUT "JD(1600)";A#
10015 IF A# <1 THEN PRINT"TRY AGAIN: ":GOTO 10010
10020 IF A#<55406!  THEN PRINT "(NOT IN UK).."
10029 REM:   Compute D% M% Y# and back-check result
10030 GOSUB 25000
10039 REM:   Compute day of week
10040 GOSUB 45000
10049 REM:   Convert Y#(1600) to Y#(AD) and display
        results
10050 Y#=Y#+1600
10060 PRINT:PRINT U$;D%;M%;Y#
10070 RETURN
10098 REM:
10099 REM:
15000 '      SUBROUTINE #3 INTERVALS
15001 REM:   VARIABLES
15002 REM:        INITIALIZED HERE none
15003 REM:        READ/INPUT HERE  none
15004 REM:        DISPLAYED HERE   G# W%
15005 REM:        INTERNAL         B#
```

COMPUTER PROGRAM JDAT3

```
15006 REM:        COMMON - to higher levels  G# W%
15007 REM:        COMMON - to lower levels   A# Y#
       D% M%
15008 REM:
15009 REM:   Enter first date, compute JD(1600)
15010 PRINT:PRINT "1ST DATE:"
15020 GOSUB 30000: Y#=Y#-1600: GOSUB 20000
15030 B#=A#
15049 REM:   Enter 2nd date, compute JD(1600)
15050 PRINT:PRINT "2ND DATE:"
15070 GOSUB 30000: Y#=Y#-1600: GOSUB 20000
15098 REM:   Calculate interval (the use of
       ABS(olute) and SGN allow
15099 REM:        for negative intervals)
15100 G#=A#-B#: W%=(INT (ABS(G#)/7))*SGN(G#)
15110 PRINT:PRINT "Interval =";G#;" days"
15120 PRINT W%; " weeks and"; G#-7*W%;" days"
15130 RETURN
19999 REM:
20000 '      SUBROUTINE #4 COMPUTATION OF JD(1600)
       FROM D% M% Y#(1600)
20001 REM:   VARIABLES
20002 REM:        INITIALIZED HERE none
20003 REM:        READ/INPUT HERE  none
20004 REM:        DISPLAYED HERE   none
20005 REM:        INTERNAL         F% N#
20006 REM:        COMMON - to higher levels  A# Y#
       D% M%
20007 REM:        COMMON - to lower levels   C# L#
       T% Z%
20008 REM:
20009 REM:   Compute F% = day of current year,
       correcting for length of February if
       appropriate
20070 F%=0
20100 IF M% >8 THEN N#=M%+1: GOTO 20120
20109 REM:  WHAT WAS THIS LINE:  999: RET
20110 N#=M%
20120 F%=D%+INT(N#/2)
20130 F%=F%+30*(M%-1)
20140 Z%=Y#: GOSUB 35000
20150 IF M% <3 THEN 20200
20160 F%=F%-T%
20198 REM:   Compute no. of leap years since
       reference date (31/12/1600)
20199 REM:        and calculate JD(1600)
20200 IF Z%=0 THEN 20500
20210 Z%=Y#-1: GOSUB 35000
20220 L#=L#-C#+INT(C#/4)
20500 A#=L#+F%+365*Z%
20520 RETURN
24998 REM:
24999 REM:
25000 '    SUBROUTINE #5  COMPUTATION OF D% M%
       Y#(1600) FROM JD(1600)
25001 REM:   VARIABLES
25002 REM:        INITIALIZED HERE (none)
25003 REM:        READ/INPUT HERE  (none)
25004 REM:        DISPLAYED HERE   (none)
25005 REM:        INTERNAL         B# E# H# N# P%
25006 REM:        COMMON - to higher levels  A# S#
       Y# D% M%
25007 REM:        COMMON - to lower levels   A# Y#
       D% M% Z% T%
25008 REM:
25009 REM:   Store JD(1600) for later back-check,
        compute year and day of year E# (first
       approximation), using mean solar year
       constant S#
25020 B#=A#
25030 Y#=A#/S#
25040 E#=(Y#-INT(Y#))*S#
25050 Y#=INT(Y#)+1
25059 REM:   Correct E# for phase of 4-year
       leap-year cycle
25060 Z%=Y#-1: GOSUB 35000
25070 IF T%=2 THEN E#=INT(E#)+1: GOTO 25200
25080 E#=INT(E#)
25090 IF E#=0 THEN E#=367-T%:Y#=Y#-1
25197 REM:
25198 REM:   Modify E# according to length of
       February of current year, if
 25199 REM:        appropriate, and calculate D% and
       M%
25200 Z%=Y#: H#=E#: GOSUB 35000
25210 IF E#+T% >61 THEN E#=E#+T%
25220 IF E#>244 THEN E#=E#-1: P%=30: GOTO 25230
25225 P%=31
25230 N#=2*INT((E#-1)/61)
25240 D%=E#-61*N#/2
25250 IF D%<P% THEN M%=N#+1: GOTO 25270
25260 M%=N#+2: D%=D%-P%: GOTO 25270
25270 IF M%>12 THEN Y#=Y#+1: E#=E#-367+T%: GOTO
       25220
25298 REM:   Back-check by calculating JD(1600)
       obtained using computed
25299 REM:        values of D% M% Y#
25300 GOSUB 20000
25310 IF A#=B# THEN 25400
25311 REM:   Correct E# accordingly and
       re-calculate D% M% (Y#)
25320 E#=H#-A#+B#: GOTO 25090
25400 RETURN
29999 REM:
30000 '   SUBROUTINE #6  INPUT OF CALENDAR DATE,
       AND VALIDITY CHECK
30001 REM:   VARIABLES
30002 REM:        INITIALIZED HERE (none)
30003 REM:        READ/INPUT HERE  Y# D% M%
30004 REM:        DISPLAYED HERE   (Y# D% M% in
       error message only)
30005 REM:        INTERNAL         (none)
30006 REM:        COMMON - to higher levels  Y# D%
       M%
30007 REM:        COMMON - to lower levels   Z% T%
30008 REM:
30010 INPUT "Date of month";D%
30020 INPUT "Month";M%
30030 INPUT "Year";Y#
30031 REM:
30038 REM:   Check validity
30040 IF D%>0 THEN IF D%<31 THEN GOTO 30070
30050 IF D%=31 THEN GOTO 30500
30060 GOSUB 31300: GOTO 30010
30070 IF M%>0 THEN IF M%<13 THEN GOTO 30100
30080 GOSUB 31300: GOTO 30020
30100 IF Y#<100 THEN Y#=Y#+1900: GOTO 30200
30120 IF Y#>1600 THEN 30140
```

COMPUTER PROGRAM JDAT3

```
30130 GOSUB 31300: GOSUB 31200: GOTO 30030
30140 IF Y#>1752 THEN 30200
30150 GOSUB 31100
30199 REM:   (inequality here is "not equal to"
30200 IF M%<>2 THEN 30400
30210 Z%=Y#-1600: GOSUB 35000
30220 IF D%+T%>30 THEN GOSUB 31300: GOTO 30010
30400 RETURN
30499 REM:   Check month M% has indeed 31 days
30500 N#=M%*1.13
30510 IF INT(N#) MOD 2 = 0 THEN 30530
30520 GOTO 30070
30530 GOSUB 31300: GOTO 30010
30599 REM:
31099 '      Error message subroutines
31100 PRINT "UK: CALENDAR INVALID BEFORE 11-9-1752"
      : PRINT:RETURN
31200 CLS:PRINT "NOT APPLICABLE FOR DATES BEFORE
      1601":PRINT: RETURN
31300 CLS:PRINT D%;M%;Y#;" INVALID": RETURN
31399 REM:
34999 REM:
35000 '   SUBROUTINE #7  LEAP YEAR CALCULATIONS
35001 REM:   VARIABLES:   INITIALIZED, READ/INPUT,
      DISPLAYED HERE   (none)
      INTERNAL         (none)
35002        COMMON - TO HIGHER LEVELS  C# L# T% Z%
      COMMON - TO LOWER LEVELS   (NONE)
35004 REM:   Calculates status of year Z% (T%=1 for
      leap year, else T%=2), and variables L# and
      C# indicating number of previous leap years
35005 C#=0: L#=0
35010 L#=Z%/4
35020 C#=L#/25
35030 IF L# - INT(L#) <>0 THEN GOTO 35120
35040 IF C# - INT(C#) <>0 THEN GOTO 35110
35050 IF C#/4-INT(C#/4) =0 THEN GOTO 35110
35060 GOTO 35120
35110 T%=1: GOTO 35200
35120 T%=2: GOTO 35200
35200 C#=INT(C#): L#=INT(L#)
35210 RETURN
35299 REM:
39999 REM:
40000 RESTORE              'EASTER SUBROUTINE
40010 FOR I%=0 TO 18
40020    READ A%(I%),B%(I%)
40030 NEXT I%
40100 GOSUB 41100
40110 GN%= Y#+1-19*INT((Y#+1)/19)
40120 D%=A%(GN%):M%=B%(GN%)
40130 Y#=Y#-1600:GOSUB 20000  'Computes JD(1600) of
      Paschal full moon
40140 A#=7+7*INT(A#/7)        'First Sunday after
      Paschal full moon
40150 GOSUB 25000:Y#=Y#+1600
40200 PRINT:PRINT"Easter Day is ";D%;M%;Y#
40210 RETURN
40300 DATA 27,3
40301 DATA 14,4
40302 DATA  3,4
40303 DATA 23,3
40304 DATA 11,4
40305 DATA 31,3
40306 DATA 18,4
40307 DATA  8,4
40308 DATA 28,3
40309 DATA 16,4
40310 DATA  5,4
40311 DATA 25,3
40312 DATA 13,4
40313 DATA  2,4
40314 DATA 22,3
40315 DATA 10,4
40316 DATA 30,3
40317 DATA 17,4
40318 DATA  7,4
41100 INPUT "Year";Y#
41120 IF Y#>1600 THEN 41140
41130 GOSUB 42200: GOTO 41100
41140 IF Y#>1752 THEN 41200
41150 GOSUB 42100
41200 RETURN
42100 CLS:PRINT "UK: CALENDAR INVALID BEFORE
      11-9-1752": PRINT:RETURN
42200 CLS:PRINT "NOT APPLICABLE FOR DATES BEFORE
      1601"
42210 PRINT"TRY AGAIN":PRINT:RETURN
42300 CLS:PRINT D%;M%;Y#;" INVALID": RETURN
45000 '   SUBROUTINE #9 DAY OF WEEK
45001 REM:   VARIABLES
45002 REM:        INITIALIZED HERE (none)
45003 REM:        READ/INPUT HERE  (none)
45004 REM:        DISPLAYED HERE   (none)
45005 REM:        INTERNAL         V%
45006 REM:        COMMON - to higher levels  A# U$
45007 REM:        COMMON - to lower levels   (none)
45008 REM:
45009 REM:   31/12/1600 was a Sunday by Gregorian
      calendar
45010 V%=A#-7*INT (A#/7)
45020 ON V%+1 GOSUB
      45100,45101,45102,45103,45104,45105,45106
45030 RETURN
45039 REM:
45100 U$="SUNDAY": RETURN
45101 U$="MONDAY": RETURN
45102 U$="TUESDAY": RETURN
45103 U$="WEDNESDAY": RETURN
45104 U$="THURSDAY": RETURN
45105 U$="FRIDAY": RETURN
45106 U$="SATURDAY": RETURN
```

COMPUTER PROGRAM SPK2X2

```
1       REM     SPK2x2.BAS
2       '       2 x 2 contingency table: Chi^2, RR, PAR, Fisher, pred.value
3       '       By Andrew Gorsuch (C)
4       '       Feb. 1982 (Casio FX702P); Jan 1987 (GWBASIC); Aug 1987 (Pred. value)
10      CLEAR:CLS
20      DIM A#(20),D#(10)
25      PRINT "STATPACK 2X2 TABLE":PRINT
26      GOSUB 1600:GOSUB 1800
27      IF F$<>"" THEN GOSUB 1700
28      IF PV$="Y" THEN GOSUB 1840:GOTO 50
30      FOR I=1 TO 4
40              PRINT "Data";I;:INPUT D#(I)
43              PRINT #1, D#(I);
45              IF D#(I)=0 THEN D#(I)=1E-10
48      NEXT I
50      PRINT #1,"":PRINT #1,""
59      REM                                     'Marginal totals and RR
60      D#(5)=D#(1)+D#(2):D#(6)=D#(3)+D#(4)
70      D#(7)=D#(1)+D#(3):D#(8)=D#(2)+D#(4):D#(0)=D#(5)+D#(6)
80      D#(9)=1
90      FOR I=1 TO 4
100             D#(9)=D#(9)*D#(I+4):V=V+1/(D#(I)+.5)
105             IF D#(9)=0 THEN 1100            'Data error: marginal total = 0
110     NEXT I
112     FOR I=0 TO 9
114             A#(I)=D#(I)
116     NEXT I
120     Y=LOG((A#(1)+.5)/(A#(2)+.5)*(A#(4)+.5)/(A#(3)+.5))
130     L=1.96*SQR(V):M=Y-L:N=Y+L
140     GOSUB 1200                              'Print RR and related statistics
150     IF F$<>"" THEN GOSUB 1900
170     E=A#(3)/A#(6):D=(A#(1)/A#(5)-E)/(1-E)   'Pop. attr. risk
180     GOSUB 1300
190     IF F$<>"" THEN GOSUB 2000
200     A#(10)=A#(1)*A#(4)-A#(2)*A#(3)          'Chi^2 test
210     T=A#(0)*A#(10)^2/A#(9)
220     U=A#(0)*(ABS (A#(10))-.5*A#(0))^2/A#(9)
230     GOSUB 1400
240     IF F$<>"" THEN GOSUB 2100
300     A#(10)=SGN(A#(10))                      'Fisher's exact test
310     P#=.5:IF A#(10)=0 THEN 550                  'No association
320     P#=0
330     FOR I=0 TO 4                                'Probability of table
340             GOSUB 1000                              'Log denominator
350             A#(15)=A#(15)+A#(17)
360     NEXT I
370     FOR I=5 TO 8                                   'Log enumerator
380             GOSUB 1000
390             A#(16)=A#(16)+A#(17)
400     NEXT I
410     Q#=A#(16)-A#(15)
420     Q#=EXP(Q#)                                      'Result
430     IF Q#<=.0000001 THEN 550                        'End of iteration
440     IF P#/Q#>=1000000# THEN 550
450     P#=P#+Q#                                    'Cusum of probabilities
460     FOR I=1 TO 4                                'Next table
470             A#(I+10)=A#(I)
480     NEXT I
490     A#(1)=A#(1)+A#(10):A#(4)=A#(4)+A#(10)
500     A#(2)=A#(2)-A#(10):A#(3)=A#(3)-A#(10)
510     A#(18)=A#(1)*A#(2)*A#(3)*A#(4)
520     IF A#(18)<0 THEN 550                        'Last table
530     IF A#(10)<0 THEN Q#=A#(11)*A#(14)/A#(2)/A#(3)*Q#:GOTO 430     'Reiterate
540     Q#=A#(12)*A#(13)/A#(1)/A#(4)*Q#:GOTO 430
```

COMPUTER PROGRAM SPK2X2

```
550     GOSUB 1500
560     IF F$<>"" THEN GOSUB 2200
570     IF PV$ = "N" OR PV$ = "n" THEN 990
598     REM
599     REM     Predictive value calculations
600     PR#=D#(5)/D#(0):QR#=1-PR#
610     SE#=D#(1)/D#(5):SP#=D#(4)/D#(6)
620     LP#=SE#/(1-SP#)
630             IF LP#=0 THEN LP#=.0000001
640     LN#=SP#/(1-SE#)
650             IF LN#=0 THEN LN#=.0000001
660     VL#=1/D#(1)+1/D#(3)+1/D#(5)+1/D#(6):VM#=1/D#(2)+1/D#(4)+1/D#(5)+1/D#(6)
670     LL#=EXP(LOG(LP#)-VL#)                           'CONF.LIM. LP#
680             IF LL#=0 THEN LL#=.0000001
690     UL#=EXP(LOG(LP#)+VL#)
700             IF UL#=0 THEN UL#=.0000001
710     LM#=EXP(LOG(LN#)-VM#)                           'CONF.LIM. LN#
720             IF LM#=0 THEN LM#=.0000001
730     UM#=EXP(LOG(LN#)+VM#)
740             IF UM#=0 THEN UM#=.0000001
750     PP#=PR#/(PR#+QR#/LP#):PN#=QR#/(QR#+PR#/LN#)
760     LU#=PR#/(PR#+QR#/LL#):UU#=PR#/(PR#+QR#/UL#)     'CONF.LIM. PP#
770     LV#=QR#/(QR#+PR#/LM#):UV#=QR#/(QR#+PR#/UM#)     'CONF.LIM. PN#
780     EF#=PR#*SE#+QR#*SP#
790     GOSUB 1550
800     IF F$<>"" THEN GOSUB 2250
810     GOTO 990
990     GOSUB 2500
992     IF Q$="Y" THEN GOSUB 2550:CLS:GOTO 27
995     END
1000    IF A#(I)<=10 THEN 1030                  'Factorials
1010            A#(17)=LOG(SQR(2*3.1415926#))+(A#(I)+.5)*LOG(A#(I))-A#(I)+1/12/A#(I)
1020    RETURN
1030            IF A#(I)=0 THEN A#(15)=1
1040            K=1
1050            FOR J=1 TO A#(I)
1060                    K=K*J
1070            NEXT J
1080            A#(17)=LOG(K)
1090    RETURN
1100    PRINT "Data error - one or more marginal totals = zero":PRINT
1110            CLEAR
1120    GOTO 30
1199    REM                                     'Print RR results
1200    PRINT"Total numbers: '1+2' =";A#(5);"  '3+4' =";A#(6);"  '1+2+3+4' =";A#(0)
1210    PRINT: PRINT "Relative risk (Woolf, Haldane):"
1220    PRINT USING "     & ####.##     & ###.## & ###.# &";"x =";EXP(Y);"(95% conf. limits ";EXP(M),"
          to ";EXP(N);")"
1230    PRINT USING "     & ##.####     & ##.####";"y =";Y;"v =";V
1240    RETURN
1299    REM                                     'Print PAR results
1300    PRINT :PRINT USING "& ####.##";"Population attributable risk =";D
1310    RETURN
1399    REM                                     'Print Chi^2 results
1400    PRINT:PRINT "Chi-squared test (1 degree of freedom):"
1410    PRINT USING "     \                                 \ ###.####  ";"Uncorrected,
          Chi^2 =";T
1420    PRINT USING "     \                                 \ ###.####  ";"With Yates' correction,
          Chi^2 =";U
1430    RETURN
1499    REM                                     'Print Fisher results
1500    PRINT:PRINT "Fisher Exact Test:":PRINT USING "     & #.######";"One-tailed p =";P#;
1510            PRINT USING "     & #.######  &";"         2p =";2*P#;"
1520    RETURN
```

```
1550    PRINT:PRINT "Predictive value calculations (L = likelihood ratio, PV = pred. value):":PRINT
          USING "     & #.###     & #.###";"Sensitivity =";SE#;"Specificity =";SP#
1560    IF LP# > 1000000! THEN LP$="Infinite":GOTO 1564
1562    LP#=(INT(10*LP#))/10:LL#=(INT(10*LL#))/10:UL#=(INT(10*UL#))/10
1563    LP$=STR$(LP#):LL$=STR$(LL#):UL$=STR$(UL#)
1564    IF LN# > 1000000! THEN LN$="Infinite":GOTO 1568
1566    LN#=(INT(10*LN#))/10:LM#=(INT(10*LM#))/10:UM#=(INT(10*UM#))/10
1567    LN$=STR$(LN#):LM$=STR$(LM#):UM$=STR$(UM#)
1568    PRINT USING "     & & && &&&     & & && &&&";"Lpos  =";LP$;"(95% c.l.";LL$;"-";UL$;")";
          "Lneg  =";LN$;"(";LM$;"-";UM$;")"
1570    PRINT USING "     & #.### &#.### & #.####    & #.### &#.### & #.####";"PVpos =";PP#;"(";
          LU#;"-";UU#;")";"PVneg =";PN#;"(";LV#;"-";UV#;")"
1580    PRINT USING "     & #.###";"Efficiency =";EF#
1590    RETURN
1600    INPUT "Outfile? (if none needed just press <ENTER>) ",F$
1610          IF F$ <> "" THEN 1640
1620          OPEN "\DEV\NUL" FOR OUTPUT AS 1
1630    RETURN
1640          OPEN F$ FOR APPEND AS 1
1650          PRINT #1, "STATPACK 2X2 TABLE",F$:PRINT #1,""
1660    RETURN
1700    INPUT "Table name";S$
1710          PRINT #1, "Table name:  ";S$
1720          PRINT #1, "Data: ";
1730    RETURN
1800    PV$="":INPUT "Predictive value calculations";PV$
1810    IF PV$ = "y" THEN PV$ = "Y"
1820    IF PV$ = "n" THEN PV$ = "N"
1830    IF (PV$ <> "N" AND PV$ <> "Y") THEN 1800
1835    RETURN
1840    INPUT "Enter D+T+, D+T-, D-T+, D-T-:  ",D#(1),D#(2),D#(3),D#(4)
1850          PRINT #1, "       T+     T-"
1852          PRINT #1, "     D+   ";D#(1);"  ";D#(2)
1854          PRINT #1, "     D-   ";D#(3);"  ";D#(4)
1860    FOR I=1 TO 4
1870          IF D#(I)=0 THEN D#(I)=1E-10
1880    NEXT I
1890    RETURN
1900    PRINT#1,"Total numbers: '1+2' =";A#(5);"  '3+4' =";A#(6);"  '1+2+3+4' =";A#(0)
1910    PRINT#1,"":PRINT#1,"Relative risk (Woolf, Haldane):"
1920          PRINT #1, USING "     & ####.##     & ###.## & ###.# &";"x =";EXP(Y);
          "(95% conf. limits ";EXP(M)," to ";EXP(N);")"
1930          PRINT #1, USING "     & ##.####     & ##.####";"y =";Y;"v =";V
1940    RETURN
1999    REM                                     'LPrint PAR results
2000    PRINT #1,"":PRINT #1, USING "& ####.##";"Population attributable risk =";D
2010    RETURN
2099    REM                                     'Print Chi^2 results
2100    PRINT #1,"":PRINT #1, "Chi-squared test (1 degree of freedom):"
2110    PRINT #1, USING "     \                                 \ ###.####  ";"Uncorrected,
          Chi^2 =";T
2120    PRINT #1, USING "     \                                 \ ###.####  ";"With Yates' correction,
          Chi^2 =";U
2130    RETURN
2199    REM                                     'Print Fisher results
2200    PRINT #1,"":PRINT #1, "Fisher Exact Test:":PRINT #1, USING "     & #.######";"One-tailed p =";
          P#
2210          PRINT #1, USING "     & #.######  &";"         2p =";2*P#;
          " (valid if samples comparable)"
2220    RETURN
2250    PRINT #1,:PRINT #1,"Predictive value calculations:":PRINT #1,USING "     & #.###     & #.###";
          "Sensitivity =";SE#;"Specificity =";SP#
2268    PRINT #1,USING "     & & && &&&     & & && &&&";"Lpos  =";LP$;"(95% c.l.";LL$;"-";UL$;")";
          "Lneg  =";LN$;"(";LM$;"-";UM$;")"
```

```
2270    PRINT #1,USING "     & #.### &#.### & #.####    & #.### &#.### & #.####&";"PVpos =";PP#;"(";LU#;
          "-";UU#;")";"PVneg =";PN#;"(";LV#;"-";UV#;")"
2280    PRINT #1,USING "     & #.###";"Efficiency =";EF#
2282    PRINT #1,:PRINT #1,
2290    RETURN
2500    Q$="":INPUT "More";Q$
2510    IF Q$ = "y" THEN Q$="Y"
2520    IF Q$ = "n" THEN Q$="N"
2530    IF (Q$ <> "N" AND Q$ <> "Y") THEN 2500
2540    RETURN
2550    FOR I=0 TO 9
2560            A#(I)=0:A#(10+I)=0:D#(I)=0
2570    NEXT I
2580    RETURN
```

EXAMPLE OF OUTPUT

```
STATPACK 2X2 TABLE          spk2x2.out

Table name:  Conv-ICA + at entry vs later diabetes
Data:        T+     T-
     D+     3      2
     D-     7      572

Total numbers: '1+2' = 5   '3+4' = 579   '1+2+3+4' = 584

Relative risk (Woolf, Haldane):
     x =  106.87     (95% conf. limits    3.46  to  %3297.9 )
     y =  4.6716     v =  3.0615

Population attributable risk =    0.60

Chi-squared test (1 degree of freedom):
     Uncorrected,            Chi^2 =     101.8056
     With Yates' correction, Chi^2 =      69.8700

Fisher Exact Test:
     One-tailed p = 0.000036
               2p = 0.000071   (valid if samples comparable)

Predictive value calculations:
     Sensitivity = 0.600     Specificity = 0.988
     Lpos  =  49.6 (95% c.l. 25.1 - 97.7)     Lneg  =  2.4 ( 1.2 - 4.9)
     PVpos = 0.300 (0.179 - 0.458)    PVneg = 0.997 (0.993 - 0.998)
     Efficiency = 0.985
```

COMPUTER PROGRAM SPKRXC

```
2     REM          ************* SPKRXC *************
4     REM          *   R-by-C Chi-squared analysis   *
6     REM          *   with linear trend analysis    *
8     REM          * and Goodman & Kruskaul's Gamma   *
10    REM          *       By Andrew Gorsuch  (C)     *
12    REM          *    July 1983 (Casio FX-702P)     *
14    REM          *    February 1987 (GW-Basic)      *
16    REM          ************************************
18    REM     Data array:
20    REM            Columns: j      1       2  .... C       Total
22    REM            Rows:    i
24    REM                     1      A(1,1)  A(1,2)  A(1,C)  A(1,0)
26    REM                     2      A(2,1)  A(2,2)  A(2,C)  A(2,0)
28    REM                     .      .       .       .       .
30    REM                     R      A(R,1)  A(R,2)  A(R,C)  A(R,0)
32    REM                 Total      A(0,1)  A(0,2)  A(0,C)  N
34    REM     Variables: (see Everitt, 1977)
36    REM            $       (messages)              N       n(..)
38    REM            A       Sigma B*N/n(i.)         O       Sigma2 n(ij)ij
40    REM            B       Sigma n(ij)^2/n(.j)     P       P for Tau stats
42    REM            C       No. of columns          Q       Q for Tau stats
44    REM            D       C(yx)                   R       No. rows
46    REM            E       C(xx)                   S       S
48    REM            F       C(yy)                   T       Sigma n(i.)i
50    REM            G       n(.j)j                  U       (not used)
52    REM            H       n(.j)j^2                V       nu (deg of frdm)
54    REM            I       i                       W       stored value V
56    REM            J       j                       X       Chi^2
58    REM            K       Deriv. of i             Y       Stored value X
60    REM            L       Deriv. of j             Z$      Y/N answers
62    REM            M       n(.i)i^2                A(i,j)  n(ij) for i,j >0
64    REM                                            A(0,j)/(i,0) n(.j)/(i,.)
2000  REM
2002  REM
2004  REM            Chi-squared subroutine
2006  REM
2010  CLEAR:CLS:INPUT "Enter number of rows, columns: ",R,C
2020  FOR I=1 TO R:FOR J=1 TO C
2030          PRINT "A(";I;",";J;")";:INPUT A(I,J)
2040          A(I,0)=A(I,0)+A(I,J)
2050          A(0,J)=A(0,J)+A(I,J)
2060          N=N+A(I,J)
2070  NEXT J:NEXT I
2080  FOR I=1 TO R
2090      B=0
2100      FOR J=1 TO C
2110          B=B+A(I,J)^2/A(0,J)
2120      NEXT J: A=A+B*N/A(I,0): NEXT I
2130  V=(R-1)*(C-1)
2140  X=A-N:W=V
2150  PRINT "":XQ$="Total":GOSUB 4000
2160  XQ$="Linear trend analysis?":GOSUB 2400
2180  IF Z$="N" THEN 2210
2190  IF Z$<>"Y" THEN GOSUB 2440:GOTO 2180
2200  GOSUB 3000
2210  XQ$="Goodman & Kruskaul's Gamma?":GOSUB 2400
2220  IF Z$="N" THEN 2250
2230  IF Z$<>"Y"THEN GOSUB 2440:GOTO 2220
2240  GOSUB 5000
2250  PRINT "BYE"
2260  END
2400  Z$="":PRINT XQ$
2410  Z$=INKEY$:IF LEN(Z$)=0 THEN 2410
2420  RETURN
```

COMPUTER PROGRAM SPKRXC

```
2440  PRINT "Y or N?";:Z$="":GOTO 2410
3000  REM
3001  REM
3002  REM              Linear trend analysis subroutine
3004  REM
3100  FOR I=1 TO R
3110      T=T+A(I,0)*I
3120      M=M+A(I,0)*I^2
3130      FOR J=1 TO C
3140          O=O+A(I,J)*I*J
3150      NEXT J
3160  NEXT I
3200  FOR J=1 TO C
3210      G=G+A(0,J)*J
3220      H=H+A(0,J)*J^2
3230  NEXT J
3300  D=O-T*G/N
3310  E=H-G^2/N
3320  F=M-T^2/N
3330  Y=D^2/E*N/F
3510  XQ$="Linear trend":X=Y:W=1:GOSUB 4000
3520  XQ$="Remainder":X=A-N-Y:W=V-1:GOSUB 4000
3540  RETURN
4000  REM
4002  REM
4004  REM              Display result subroutine
4006  REM
4010  PRINT XQ$;"  Chi^2 (";W;"degrees of freedom) =";
4020  PRINT USING "####.####";X:PRINT ""
4030  RETURN
5000  REM
5002  REM
5004  REM              Goodman & Kruskaul's Gamma subroutine
5006  REM
5020  FOR I=1 TO R-1
5030      FOR J=1 TO C-1
5040          FOR K=1 TO R
5050              FOR L=J+1 TO C
5060                  P=P+A(I,J)*A(K,L)
5070              NEXT L
5080          NEXT K
5090      NEXT J
5130      FOR J=2 TO C
5140          FOR K=I+1 TO R
5150              FOR L=J-1 TO 1 STEP -1
5160                  Q=Q+A(I,J)*A(K,L)
5170              NEXT L
5180          NEXT K
5190      NEXT J
5200  NEXT I
5210  S=P-Q
5220  U=S/(P+Q)
5230  PRINT USING "&  #.###";"Gamma =";U:PRINT "":PRINT ""
5240  RETURN
```

XIII APPENDIX 5d

COMPUTER PROGRAM SPKRCL

```
1     REM         ********************************************
2     REM         **                SPKRCL                  **
3     REM         **                                        **
4     REM         **    TRIDIMENSIONAL CONTINGENCY TABLE    **
5     REM         **                                        **
6     '           **        (C)  Andrew Gorsuch             **
7     REM         **                                        **
8     REM         **     February 1985 (Casio FX702P)       **
9     REM         **       February 1988 (GW BASIC)         **
10    REM         **                                        **
11    REM         ********************************************
12    REM
13    REM
14    REM All lines including " REM " may be removed for faster processing.
15    REM
16    REM FUNCTION.  Set up a three-dimensional array in RAM to accommodate the desired numbers of rows,
17    REM     columns and layers.  Enter data from keyboard, with prompts of coordinates.  Compute
18    REM     marginal totals over one, two and three variables.  Compute expected values, degrees of
19    REM     freedom and Chi^2 for the following hypotheses: mutual independence; partial independence
20    REM     (row variables independent of comlumn and layer variables, and correspondingly for column
21    REM     and layer variables).  Hypotheses of conditional independence and measures of association
22    REM     are not covered, and require log-linear modelling (e.g. GLIM).  At present, storage of
23    REM     data files on disk is not supported, but this could readily be added.
24    REM REFERENCE.  Everitt, BS.  The Analysis of Contingency Tables.  Chapman & Hall, London, 1977.
25    REM
26    REM STRUCTURE OF TABLE.
27    REM     The number of observations in a cell of the table is given by N(ijk), where i is the row
28    REM     number (values 1 to R), j is the column number (1 to C) and k is the layer number (1 to
29    REM     L).  The number expected by a given hypothesis is E(ijk).
30    REM
31    REM     Since the computer for which this program was originally written does not support three-
32    REM     dimensional arrays, the algorithm uses a virtual three-dimensional array within a two-
33    REM     dimensional array.
34    REM
35    REM SIZE OF TABLE ACCOMMODATED.
36    REM     The data array is dynamically dimensioned, and its maximum size is very large on any
37    REM     IBM-PC-compatible computer capable of running GW-Basic.
38    REM
39    REM Master control program starts at 100
40    REM     and establishes size of table
41    REM Major subroutine starts at 3000 (ijk loops, expected values, Chi^2).
42    REM 2nd-level subroutines start at:
43    REM     4000 (data entry and marginal totals),
44    REM     5000 (expected values: mutual independence model)
45    REM     6000 (expected values: rows independent of columns and layers)
46    REM     7000 (expected values: columns independent of rows and layers)
47    REM     8000 (expected values: layers independent of rows and columns)
48    REM     9000 (display result)
49    REM
99    REM MASTER CONTROL PROGRAM P1
100   CLEAR:CLS:PRINT:PRINT:PRINT"            TRIDIMENSIONAL CONTINGENCY TABLE":PRINT
120   INPUT "Enter no. Rows, Columns, Layers (separated by commas): ",R%,C%,L%:T%=2*L%+1
130   DY%=R%*(C%+1):DX%=3*L%+2:DIM A(DX%,DY%)
140   REM   (this and following lines may be used if any constraint on table size)
200   Q%=4:F$="N":GOSUB 3000                                  'Enter data
299   REM Compute 'expected' numbers and Chi^2 values under the several null hypotheses.
300   F$=" "
310   FOR Q%=5 TO 8
320       GOSUB 3000
330   NEXT Q%
340   CLS:PRINT"SPKRCL Finished":END
2998  REM
2999  REM SUBROUTINE P3: IJK loops, expected values, cumulation of Chi^2
3000  X#=0:CLS
```

COMPUTER PROGRAM SPKRCL

```
3005 ON Q% GOSUB ,,,3240,3250,3260,3270,3280
3007 PRINT:PRINT:PRINT S$:PRINT:PRINT"  Cell        Observed  Expected  Difference"
3010 FOR I%=1 TO R%:FOR J%=1 TO C%
3020         H%=J%+(C%+1)*(I%-1)
3030         FOR K%=1 TO L%
3040             PRINT F$;I%;J%;K%;
3050             ON Q% GOSUB ,,,4000,5000,6000,7000,8000,9000
3060             IF Q%=4 THEN 3080
3065             PRINT USING "   #####     #####.##  #####.##";A(K%,H%);A(T%+K%,H%);A(K%,H%)-
                    A(T%+K%,H%)
3070             X#=X#+(A(K%,H%)-A(T%+K%,H%))^2/A(T%+K%,H%)
3080 NEXT K%:NEXT J%:NEXT I%
3090 IF Q%=4 THEN 3200
3100 ON KEY(10) GOSUB 3300
3110 KEY(10) ON
3120    GOSUB 9000
3130    PRINT:PRINT"Press key F10 to continue..."
3140    GOTO 3140
3190 KEY(10) OFF
3200 RETURN
3240 S$="Enter data"
3250 S$="Mutual independence":D%=R%*C%*L%-R%-C%-L%+2:RETURN
3260 S$="Independent Row variable:":D%=R%*C%*L%-C%*L%-R%+1:RETURN
3270 S$="Independent Column variable:":D%=R%*C%*L%-R%*L%-C%+1:RETURN
3280 S$="Independent Layer variable:":D%=R%*C%*L%-R%*C%-L%+1:RETURN
3300 RETURN 3190
3998 REM
3999 SUBROUTINE P4: COMPUTE MARGINAL TOTALS DURING DATA entry
4000 INPUT B:A(K%,H%)=B
4020 A(0,H%)=A(0,H%)+B
4030 A(K%,H%-J%)=A(K%,H%-J%)+B
4040 A(K%+L%,J%)=A(K%+L%,J%)+B
4050 A(0,H%-J%)=A(0,H%-J%)+B
4060 A(T%,J%)=A(T%,J%)+B
4070 A(K%+L%,0)=A(K%+L%,0)+B
4080 N#=N#+B
4090 RETURN
4998 REM
4999 REM SUBROUTINE P5: 'Expecteds' under mutual independence
5000 A(T%+K%,H%)=A(0,H%-J%)*A(T%,J%)*A(L%+K%,0)/N#^2:RETURN
5998 REM
5999 REM SUBROUTINE P6: 'Expecteds, R% independent of C% and L%
6000 A(T%+K%,H%)=A(0,H%-J%)*A(L%+K%,J)/N#:RETURN
6998 REM
6999 REM SUBROUTINE P7: 'Expecteds, C% independent of R% and L%
7000 A(T%+K%,H%)=A(K%,H%-J%)*A(T%,J%)/N#:RETURN
7998 REM
7999 REM SUBROUTINE P8: 'Expecteds, L% independent of R% and C%
8000 A(T%+K%,H%)=A(0,H%)*A(L%+K%,0)/N#:RETURN
8998 REM
8999 REM SUBROUTINE P9: Display results
9000 PRINT:PRINT
9020 PRINT USING "\                              \ \    \##\ \####.####";S$;"Chi^2(";D%;") =";X#
9040 RETURN
```

COMPUTER PROGRAM HARDY1

```
1       REM HARDY1 by Andrew Gorsuch 1983
2       REM To redirect output to printer, change "SCRN:" to "LPT1:" in Lines 10 and 50
3       REM To redirect output to file, change "SCRN:" to "filename" in and change OUTPUT to APPEND in
lines 10 and 50.
4       REM To exit, press CTRL+BREAK or CTRL+C
10      OPEN "SCRN:" FOR OUTPUT AS #1
20      CLS:PRINT#1,"Hardy-Weinberg Equilibrium":PRINT#1,"":PRINT#1,"":CLOSE
30      PRINT:INPUT "Enter number of subjects tested (N) ",N
40      PRINT:PRINT:INPUT "Enter Specificity, Phenotype count, Genotype count (S$,F,G) ",S$,F,G
50      OPEN "SCRN:" FOR OUTPUT AS #1
60      P=G/2/N:Q=1-SQR(1-F/N)
70      M=G-F:T=F-M:V=N-F
80      O=P^2*N:U=2*P*(1-P)*N:W=(1-P)^2*N
90      X=(M-O)^2/O+(T-U)^2/U+(V-W)^2/W
100     PRINT#1,"":PRINT#1,S$:PRINT#1,"":PRINT#1,USING"& #.####  & #.####";"Gene frequency: Observed";
P;"  Calculated";Q
110     PRINT#1,"":PRINT#1,USING"& ####  & ####.####";"Homozygous +:  Observed";M;"  Expected";O
120     PRINT#1,USING"& ####  & ####.####";"Heterozygous:  Observed";T;"  Expected";U
130     PRINT#1,USING"& ####  & ####.####";"Homozygous -:  Observed";V;"  Expected";W
140     PRINT#1,"":PRINT#1,USING"& ####.#### &";"Chi^2 =";X;" (2 degrees of freedom)":PRINT#1,"":
PRINT#1,""
150     CLOSE:GOTO 40
```

EXAMPLE OF OUTPUT

```
Hardy-Weinberg Equilibrium


HLA-A1

Gene frequency: Observed 0.2697    Calculated 0.2865

Homozygous +:  Observed     8     Expected    12.0015
Heterozygous:  Observed    73     Expected    64.9970
Homozygous -:  Observed    84     Expected    88.0015

Chi^2 =     2.5015  (2 degrees of freedom)


HLA-B8

Gene frequency: Observed 0.2909    Calculated 0.3213

Homozygous +:  Observed     7     Expected    13.9636
Heterozygous:  Observed    82     Expected    68.0727
Homozygous -:  Observed    76     Expected    82.9637

Chi^2 =     6.9067  (2 degrees of freedom)
```

COMPUTER PROGRAM TOMBOD

```
10  REM     ************************************************
20  REM     **                                            **
30  REM     **                   TOMBOD                   **
40  REM     **     Haplotype sharing by affected sibs:    **
50  REM     **           Goodness-of-fit Chi^2            **
60  REM     **                                            **
70  REM     **          (C) Andrew Gorsuch 1986           **
80  REM     **                                            **
90  REM     **       Written in Microsoft GWBASIC         **
100 REM     ************************************************
110 '       To plot goodness-of-fit Chi^2
120 '       versus gene frequency p
130 '       for distribution of affected
140 '       sib-pairs according to HLA
150 '       haplotype concordance,
160 '       assuming a single disease
170 '       susceptibility gene,
180 '       Mendelian recessive or
190 '       dominant inheritance and
200 '       no recombination.
210 '       From Thomson & Bodmer (1977).
220 '       SCREEN GRAPHICS (lines 1480 - 1770) work with CGA adaptor
230 '       on IBM-PC-compatible computers, and should work with EGA
240 '       adaptor.  To print out graphics, first load GRAPHICS program
250 '       from DOS, and press <SHIFT> + <PrtSc> keys when graphics are
260 '       displayed.  An IBM Graphics, Epson or compatible printer is
270 '       needed.
280 '
1000 SCREEN 2: CLEAR:CLS:ON ERROR GOTO 2230
1010 PRINT "TOMBOD"
1020 PRINT:PRINT "Output to (p)rinter, to (f)ile or to (s)creen?":PRINT
1030 Q$="":Q$=INKEY$:IF Q$ ="" THEN GOTO 1030:PRINT
1040 IF Q$ <> "P" AND Q$ <> "p" AND Q$ <> "F" AND Q$<>"f" AND Q$ <> "S" AND Q$ <>"s" THEN GOSUB 2220:
     GOTO 1020
1050 Q% = 1:IF Q$="F" OR Q$="f"THEN Q%=2:IF Q$="S" OR Q$="s" THEN Q%=3
1060 IF Q$="S" OR Q$="s" THEN Q%=3
1070 ON Q% GOSUB 1080,1090,1100:GOTO 1110
1080 OPEN "LPT1:" FOR OUTPUT AS #2:RETURN
1090 INPUT "Filename";F$:OPEN F$ FOR OUTPUT AS #2:RETURN
1100 OPEN "SCRN:" FOR OUTPUT AS #2:RETURN
1110 OPEN "SCRN:" FOR OUTPUT AS #1
1120 DIM R(4,21),D(4,21),X(21),Y(21),A(3),P(21)
1130 FOR I=1 TO 21
1140         READ R(1,I),R(2,I),R(3,I),R(4,I)
1150 NEXT I
1160 FOR I=1 TO 21
1170         READ D(1,I),D(2,I),D(3,I),D(4,I)
1180 NEXT I
1190 INPUT "Enter No's H2, H1, H0: ",A(1),A(2),A(3)
1200 LET N = A(1)+A(2)+A(3)
1210 PRINT#2,"":PRINT#2, "Recessive and Dominant":PRINT#2,""
1220 GOTO 1230
1230 FOR I=1 TO 21
1240         X(I) = 0:P(I) =R(1,I)
1250         Y(I) = 0
1260         FOR J=1 TO 3
1270                 X(I) = X(I)+ (N*R(J+1,I)-A(J))^2/(N*R(J+1,I))
1280                 Y(I) = Y(I)+ (N*D(J+1,I)-A(J))^2/(N*D(J+1,I))
1290 NEXT J:NEXT I
1300 M$ = "RECESSIVE"
1310 GOSUB 1380
1320 ON KEY (1) GOSUB 1480:ON KEY(2) GOSUB 1780:ON KEY (3) GOSUB 1790:KEY(1)ON:KEY(2)ON:KEY (3) ON
1330 PRINT#1,"":PRINT#1,"To display graph press key F1"
1340 PRINT#1,"To print graph press SHIFT PrtSc (if Graphics program has been loaded from DOS)":PRINT#1,
```

```
     "To enter more data after displaying graph press key F2":PRINT#1,"":PRINT#1,"To finish press key
     F3"
1350 GOTO 1350
1360 CLS:SCREEN 2:GOTO 1190
1370 KEY (1)OFF:KEY(2)OFF:KEY(3)OFF:CLS:SCREEN 2:PRINT#1,"Leaving TOMBOD":END
1380 PRINT#2, "GOODNESS-OF-FIT CHI-SQUARED":PRINT#2,"":PRINT#2,"H2 =";A(1);"  H1 =";A(2);"  H0 =";A(3):
     PRINT#2,"N =";N:PRINT#2,""
1390 PRINT#2,USING "\         \";"GENE FREQ.";"RECESSIVE";"DOMINANT":PRINT#2,""
1400 PRINT#2,USING "\         \";"";CHR$(0);CHR$(0)
1410 FOR I=1 TO 21
1420         IF X(I)>999.999 THEN LET X(I)=999.999#
1430         IF Y(I)>999.999 THEN LET Y(I)=999.999#
1440         PRINT#2,USING "   #.##    ###.###    ###.###";P(I);X(I);Y(I)
1450 NEXT I
1460 PRINT#2,"":PRINT#2,""
1470 RETURN
1480 CLS:SCREEN 1                    'Graphics subroutine
1490 LINE (68,4)-(68,164)
1500 LINE (68,164)-(228,164)
1510 FOR I=0 TO 10 STEP 1
1520 PRESET (66,164)
1530 PSET STEP (0,-16*I):LOCATE 21-2*I,5:PRINT#1,2*I
1540 NEXT I
1550 FOR I=0 TO 10
1560 PRESET (68,166)
1570 PSET STEP (16*I,0)
1580 NEXT I
1590 LOCATE 22,9:PRINT#1,"0":LOCATE 22,18:PRINT#1,"0.5":LOCATE 22,29:PRINT#1,"1":LOCATE 23,12:PRINT#1,
     "Gene frequency":LOCATE 2,1:PRINT#1,"Chi2":LOCATE 4,1:PRINT#1,"(2df)"
1600 FOR I=1 TO 40
1610         PRESET (68,116):PSET STEP (4*I,0)
1620 NEXT I
1630 LOCATE 15,31:PRINT#1,"P=0.05":LOCATE 19,31:PRINT"Rec've":CIRCLE(308,148),3:PAINT (308,148):LOCATE
     17,31:PRINT"Domin't":CIRCLE (308,132),2
1640 LX=68:LY=-999
1650 FOR I=1 TO 21
1660         IF X(I)>50 THEN 1690
1670         PX=68+P(I)*160:PY=164-X(I)*8:LINE (LX,LY)-(PX,PY)
1680         CIRCLE (PX,PY),3:PAINT (PX-1,PY-1):PAINT (PX+1,PY+1):LX=PX:LY=PY
1690 NEXT I
1700 LX=68:LY=-999
1710 FOR I=1 TO 21
1720         IF Y(I)>50 THEN 1750
1730         PX=68+P(I)*160:PY=164-Y(I)*8:LINE (LX,LY)-(PX,PY)
1740         CIRCLE (PX,PY),2:LX=PX:LY=PY
1750 NEXT I
1760 LOCATE 2,31:PRINT#1,"H2 =";A(1):LOCATE 4,31:PRINT#1,"H1 =";A(2):LOCATE 6,31:PRINT#1,"H0 =";A(3)
1770 GOTO 1770
1780 RETURN 1360
1790 RETURN 1370
1800 DATA .00,1.00,.00001,.00001     'Data from Thomson & Bodmer (1977):
1810 DATA .05,.907,.091,.002         'gene freq. and expected proportions H2, H1 and H0 respectively.
1820 DATA .10,.827,.165,.008
1830 DATA .15,.756,.227,.017
1840 DATA .20,.694,.278,.028
1850 DATA .25,.640,.320,.040
1860 DATA .30,.592,.355,.053
1870 DATA .35,.549,.384,.067
1880 DATA .40,.510,.408,.082
1890 DATA .45,.476,.428,.096
1900 DATA .50,.445,.444,.111
1910 DATA .55,.416,.458,.126
1920 DATA .60,.391,.469,.140
1930 DATA .65,.367,.478,.155
```

```
1940 DATA .70,.346,.484,.170
1950 DATA .75,.326,.490,.184
1960 DATA .80,.309,.494,.198
1970 DATA .85,.292,.497,.211
1980 DATA .90,.277,.499,.224
1990 DATA .95,.263,.500,.237
2000 DATA 1.0,.250,.500,.250
2010 DATA .00,.500,.500,.00001
2020 DATA .05,.460,.495,.045
2030 DATA .10,.428,.491,.081
2040 DATA .15,.401,.488,.111
2050 DATA .20,.377,.487,.136
2060 DATA .25,.358,.486,.156
2070 DATA .30,.341,.485,.174
2080 DATA .35,.326,.485,.189
2090 DATA .40,.313,.486,.201
2100 DATA .45,.302,.487,.211
2110 DATA .50,.293,.488,.219
2120 DATA .55,.284,.489,.227
2130 DATA .60,.277,.490,.233
2140 DATA .65,.271,.492,.237
2150 DATA .70,.265,.494,.241
2160 DATA .75,.261,.495,.244
2170 DATA .80,.257,.497,.246
2180 DATA .85,.254,.498,.248
2190 DATA .90,.252,.499,.249
2200 DATA .95,.250,.500,.250
2210 DATA 1.0,.250,.500,.250
2220 PRINT:PRINT"Please enter p, f or s":RETURN
2230 PRINT:IF ERR = 24 OR ERR = 25 THEN IF ERL = 1210 THEN PRINT"Is printer on line?":PRINT:PRINT
     "Press key F5 to continue":PRINT"":BEEP:STOP:RESUME  1210
2240 PRINT:BEEP:ON ERROR GOTO 0
```

EXAMPLE OF ALPHA-NUMERIC OUTPUT (See Figure VIII-2 for graphical output)

Recessive and Dominant
GOODNESS-OF-FIT CHI-SQUARED

H2 = 182 H1 = 128 H0 = 18
N = 328

GENE FREQ.	RECESSIVE	DOMINANT
0.00	999.999	999.999
0.05	826.160	14.401
0.10	220.324	21.882
0.15	83.737	35.098
0.20	32.475	49.705
0.25	10.586	63.201
0.30	1.933	76.821
0.35	0.773	89.997
0.40	4.491	102.339
0.45	11.157	113.648
0.50	20.341	123.538
0.55	31.663	134.092
0.60	43.842	142.758
0.65	58.045	150.344
0.70	72.888	158.301
0.75	89.088	163.886
0.80	105.254	169.470
0.85	123.036	173.877
0.90	141.089	176.815
0.95	160.054	179.805
1.00	179.805	179.805

'Comment:
'999.999 represents any $\chi^2 > 999.998$

XIII APPENDIX 6

DIABETES IN EAST BERKS, 1978-1980

XIII-6-1. INCIDENCE OF TYPE 1 DIABETES EAST BERKSHIRE, AUGUST 1978 - JULY 1980

A. CALCULATION OF AGE ADJUSTMENT FACTORS

Demographic information from Office of Population Censuses & Surveys Census 1981: County Report - Berkshire, Part 1. HMSO 1982

DISTRICT	TOTAL POPULATION			WHITE CAUCASOID (est'd by subtracting proportion of New Commonwealth origin)										
				POPULATION BY AGE										
	All	New C'wth	White Cauc'd	Age group 0-4 yr (interval y=5yr)			Age 5 - 9yr (y = 5)			Age 10 -15yr (y = 6)			Age 16 -19yr (y = 4)			TOTAL AGED 0-19yr (y=20)
		(%)		No.	n/t (%)	Factor	No.	n/t (%)	Factor	No.	n/t (%)	Factor	No.	n/t (%)	Factor	
	T	c	T-Tc/100 t	n		f	n		f	n		f	n		f	U
Bracknell	81225	2.0	79601	5970	7.5	1.0533	6288	7.9	1.0000	7482	9.4	1.0085	5412	6.8	0.9295	25152
Slough	96715	20.9	76502	5355	7.0	1.0643	5508	7.2	1.0347	6579	8.6	1.0395	5355	7.0	0.8514	22797
Windsor & Maidenhead	129245	3.6	124592	7226	5.8	1.1854	8846	7.1	0.9683	10715	8.6	0.9593	7476	6.0	0.9166	34263
E.Berks total	307185	8.6	280694	18551	6.6	1.1079	20642	7.4	0.9957	24776	8.8	0.9955	18243	6.5	0.9013	82212
Sl,Windsor and M'head	225960	11.0	201094	12581	6.3	1.1339	14354	7.1	0.9938	17294	8.6	0.9898	12831	6.4	0.8894	57060

B. NEW CASES OF DIABETES AND CALCULATION OF AGE-ADJUSTED INCIDENCE RATES.

DISTRICT	AGE GROUP								2-YEAR TOTAL		INCIDENCE		
	0 - 4 yr		5 - 9 yr		10 -15 yr		16 -19 yr		Ages 0 -19		Cases $.10^{-5}.y^{-1}$		
	No. cases		No. cases		No. cases		No. cases				Crude	Adjusted	
	Obs. O	Adj. O.f	Obs.	Adj.	Obs.	Adj.	Obs.	Adj.	Obs.	Adj.		Mean	95% conf.lim.
Bracknell	0	0.00	0	0.00	2	2.02	0	0.00	2	2.02	3.98	4.01	2.04- 5.98
Slough	0	0.00	3	3.10	3	3.12	4	3.41	10	9.63	21.93	21.12	16.82-25.42
Windsor and Maidenhead	1	1.19	3	2.90	4	3.84	3	2.75	11	10.68	16.05	15.58	11.05-20.11
E.Berks total	1	1.11	6	5.97	9	8.96	7	6.31	23	22.35	13.99	13.59	7.04-20.14
Sl.,Windsor and M'head	1	1.13	6	5.96	7	6.93	7	6.23	21	20.25	18.40	17.75	11.51-23.98

DIABETES IN EAST BERKS, 1978-1980

XIII-6-2. TYPE 1 DIABETES: PREVALENCE IN EAST BERKS DISTRICTS, excluding Bracknell, ADJUSTED to age distribution of BWFS siblings in Oct 1980

CONSTANTS	SYMBOL	NAME	VALUE
Population aged 0-19, excluding ''New Commonwealth: (1981 Census)	p	POPEB20	57060
Mean age-adjusted annual incidence in districts:	i	INCEB	1.775E-4

AGE	No.SIBS in BWFS	RCUFREQa Cum.freq. DM reltv to c(14) From lit.	CUFREQa Cum.freq. DM in dist calc. by: r.14.i	EXPEBa Cum.no.DM in distct calc.by: c.p/20	AdjEXPEBa Adj. of e to age di of BWFS s e.n.34/28
a	n	r (%)	c	e	d
0	0	0.7	0.000017	0.0496	0.0000
1	1	3.6	0.000089	0.2552	0.0301
2	1	7.3	0.000181	0.5175	0.0611
3	3	11.7	0.000291	0.8295	0.2938
4	0	16.7	0.000415	1.1840	0.0000
5	5	22.3	0.000554	1.5810	0.9332
6	5	28.3	0.000703	2.0064	1.1843
7	10	34.5	0.000857	2.4459	2.8876
8	12	41.8	0.001039	2.9635	4.1983
9	14	49.7	0.001235	3.5236	5.8237
10	8	59.1	0.001469	4.1900	3.9572
11	14	68.8	0.001710	4.8777	8.0618
12	14	79.3	0.001971	5.6221	9.2921
13	21	90.7	0.002254	6.4304	15.9419
14	18	100.0	0.002485	7.0897	15.0656
15	22	106.0	0.002634	7.5151	19.5184
16	16	112.0	0.002783	7.9405	14.9987
17	17	119.0	0.002957	8.4367	16.9321
18	16	123.4	0.003066	8.7487	16.5253
19	18	128.4	0.003191	9.1032	19.3443
20	12	135.9	0.003377	9.6349	13.6495
21	13	141.8	0.003524	10.0532	15.4289
22	8	147.3	0.003660	10.4431	9.8630
23	8	152.4	0.003787	10.8047	10.2044
24	7	157.2	0.003906	11.1450	9.2101
25	4	162.3	0.004033	11.5066	5.4337
26	2	166.9	0.004147	11.8327	2.7938
27	6	172.7	0.004292	12.2439	8.6728
28	3	178.1	0.004426	12.6268	4.4720
29	3	184.0	0.004572	13.0451	4.6201
30	0	189.0	0.004697	13.3995	0.0000
31	2	195.5	0.004858	13.8604	3.2726
32	3	199.7	0.004963	14.1581	5.0143
33	2	205.0	0.005094	14.5339	3.4316
Total	288			254.5984	251.1163

DIABETES IN EAST BERKS, 1978-1980

XIII-6-3. ABSOLUTE RISK OF TYPE 1 DIABETES IN HLA-GENOTYPED SIBLINGS OF AFFECTED PROBANDS EAST BERKSHIRE, 1978-1980

AGE	CUFREQa Cum.freq. DM in dist	ABSOLUTE RISK (%) Estimate, with lower and upper conf.limits (lcl, ucl)								
		H2 SIBS			H1 SIBS			ALL SIBS		
		Est.	lcl	ucl	Est.	lcl	ucl	Est.	lcl	ucl
		40.c	13.c	121.c	10.c	2.c	40.c	12.c	7.c	21.c
a	c									
0	0.000017	0	0	0	0	0	0	0	0	0
1	0.000089	0	0	1	0	0	0	0	0	0
2	0.000181	1	0	2	0	0	1	0	0	0
3	0.000291	1	0	4	0	0	1	0	0	1
4	0.000415	2	1	5	0	0	2	0	0	1
5	0.000554	2	1	7	1	0	2	1	0	1
6	0.000703	3	1	9	1	0	3	1	0	1
7	0.000857	3	1	10	1	0	3	1	1	2
8	0.001039	4	1	13	1	0	4	1	1	2
9	0.001235	5	2	15	1	0	5	1	1	3
10	0.001469	6	2	18	1	0	6	2	1	3
11	0.001710	7	2	21	2	0	7	2	1	4
12	0.001971	8	3	24	2	0	8	2	1	4
13	0.002254	9	3	27	2	0	9	3	2	5
14	0.002485	10	3	30	2	0	10	3	2	5
15	0.002634	11	3	32	3	1	11	3	2	6
16	0.002783	11	4	34	3	1	11	3	2	6
17	0.002957	12	4	36	3	1	12	4	2	6
18	0.003066	12	4	37	3	1	12	4	2	6
19	0.003191	13	4	39	3	1	13	4	2	7
20	0.003377	14	4	41	3	1	14	4	2	7
21	0.003524	14	5	43	4	1	14	4	2	7
22	0.003660	15	5	44	4	1	15	4	3	8
23	0.003787	15	5	46	4	1	15	5	3	8
24	0.003906	16	5	47	4	1	16	5	3	8
25	0.004033	16	5	49	4	1	16	5	3	8
26	0.004147	17	5	50	4	1	17	5	3	9
27	0.004292	17	6	52	4	1	17	5	3	9
28	0.004426	18	6	54	4	1	18	5	3	9
29	0.004572	18	6	55	5	1	18	5	3	10
30	0.004697	19	6	57	5	1	19	6	3	10

XIII APPENDIX 7

PUBLISHED PAPERS FORMING PART OF DISSERTATION

The following papers written by the author form part of this Dissertation, and are appended inside the back cover:

7-1. GORSUCH AN, DEAN BM, BOTTAZZO GF, LISTER J, CUDWORTH AG. Evidence that Type 1 diabetes and thyrogastric autoimmunity have different genetic determinants. Br Med J 1980; 280: 145-7

7-2. GORSUCH AN, SPENCER KM, LISTER J, McNALLY JM, DEAN BM, BOTTAZZO GF, CUDWORTH AG. Evidence for a long prediabetic period in Type 1 (insulin-dependent) diabetes mellitus. Lancet 1981; ii: 1363-5

7-3. GORSUCH AN, SPENCER KM, LISTER J, WOLF E, BOTTAZZO GF, CUDWORTH AG. Can future Type 1 diabetes be predicted? A study in families of affected children. Diabetes 1982b; 31: 862-866

7-4. GORSUCH AN, SPENCER KMM, WOLF E, CUDWORTH AG. HLA and family studies. In: Köbberling J, Tattersall R (eds), The genetics of diabetes mellitus. Proc Serono Symposia, Vol 47 Academic Press, London. 1982c 43-53

Evidence that type I diabetes and thyrogastric autoimmunity have different genetic determinants

A N GORSUCH, B M DEAN, G F BOTTAZZO, J LISTER, A G CUDWORTH

Summary and conclusions

The prevalences of autoimmune endocrine disease and relevant organ-specific autoantibodies were determined in 141 patients with type I (insulin-dependent) diabetes and their families. All available members of the families were genotyped for HLA. Islet-cell antibody was found in 10 (4%) out of 248 unaffected siblings, all of whom were genetically potential cases of diabetes. One developed classical symptoms six months later. In contrast, thyroid and gastric parietal-cell antibodies occurred independent of the HLA-linked susceptibility to diabetes.

These results suggest that different genes control the production of these autoantibodies and the susceptibility to type I diabetes.

Department of Diabetes, Medical Unit, St Bartholomew's Hospital, London EC1A 7BE

A N GORSUCH, MRCP, honorary lecturer in medicine
B M DEAN, PHD, lecturer
A G CUDWORTH, MD, FRCP, consultant physician and senior lecturer

The Middlesex Hospital, London W1N 8AA

G.F BOTTAZZO, MD, lecturer in immunology

King Edward VII Hospital, Windsor

J LISTER, MD, FRCP, consultant physician

COPYRIGHT © 1980. ALL RIGHTS OF REPRODUCTION OF THIS REPRINT ARE RESERVED IN ALL COUNTRIES OF THE WORLD

BMJ/45/80

Introduction

Both hereditary and environmental factors are concerned in the aetiology of type I (insulin-dependent) diabetes mellitus (DM). Probably the major genetic susceptibility to DM is determined by a gene in the HLA regions of chromosomes 6.[1] In Caucasians this gene (which we term S_1) is in linkage disequilibrium with the HLA haplotype DW3-DRW3-B8-A1. A different gene (S_2), in linkage disequilibrium with DR4, DRW4, B15, CW3, and A2, possibly increases the risk of DM when present with S_1.[2 3] Family studies (as opposed to population studies) show a close linkage between these genes and the HLA regions of chromosomes 6 (the HLA haplotypes) transmitted in each family whether or not these haplotypes include the specificities mentioned.[4] Thus defining HLA phenotypes in all members of a diabetic's family, and hence deducing the HLA genotypes, permits haplotype-segregation analysis to determine which of the unaffected siblings may be susceptible to DM.

Several diseases, particularly the autoimmune endocrinopathies, are associated with HLA-B8,[5 6] and there is an increased prevalence both of these diseases and of organ-specific humoral autoimmunity in type I diabetics and their relatives.[7 8] The detection of pancreatic islet-cell antibody (ICA) in diabetes[9] apparently supported the classification of type I DM as an autoimmune endocrinopathy, but this may be an oversimplification. Thus in contrast to other organ-specific autoantibodies ICA is transient in most diabetics, being almost invariably detectable at diagnosis but persisting beyond two to three years in only 15-20% of cases.[10] In polyendocrine disease, however, ICA generally persists for many years, and the few patients with type I DM alone who have persistent ICA apparently form a distinct, homogeneous subgroup of diabetics characterised by a comparatively late onset of disease, a preponderance of female patients, increased prevalence of organ-specific autoimmunity, and a strong family history of autoimmune endocrinopathy.[11] This subgroup may represent patients with primary autoimmune or type IB DM, whereas juvenile-onset or type IA diabetes (for which transient ICA is a marker) may have a different aetiology, possibly viral.

We must discover whether susceptibility to thyroid and other organ-specific autoimmunity is determined by the same gene as susceptibility to type I DM. Bottazzo *et al*[12] tried to resolve this by analysing HLA phenotypes in a group of patients with coexistent type I DM and autoimmune thyroid disease. Their findings—namely, that the relative risk in respect of B8 was roughly additive compared with that in type I DM and Graves's disease when analysed separately—suggested that different HLA-linked genes were responsible for the two diseases. A more powerful method of investigation is by HLA-haplotype-segregation analysis. If the B8-linked "diabetogenic" and "thyroid-autoimmunity" genes were identical, then auto-

immunity would be expected to occur preferentially in subjects with a diabetogenic haplotype. We present strong evidence to the contrary.

Subjects and methods

Population studied—The Barts-Windsor Prospective Family Study was begun in 1978 to investigate genetic, immunological, and environmental factors in the aetiology of type I DM. The 158 Caucasian families who participated were selected at random with respect to autoimmune endocrinopathy, but each contained a classical type I diabetic proband with at least one available sibling aged under 20. HLA haplotypes were unequivocally distinguished in 141 families, the others being excluded from the present analysis because of expaternity of the proband (six families) or HLA-A, B, C homozygosity in a parent. Half-siblings were also excluded. Our analysis was therefore conducted on 141 diabetic probands (mean age 14·0 years, range 4-29; male:female ratio 1·2:1; mean duration of diabetes 5·6 years), 255 siblings (mean age 14·2 years, range 1-32), 136 fathers, and 140 mothers (mean ages 43·4 and 40·3 years respectively).

HLA genotypes—HLA-A, B, and C typing was performed on all available members of the families by the microlymphocytotoxicity technique with 100 fully standardised antisera covering all known specificities in Caucasian populations.[13]

Autoantibodies—ICA and gastric parietal-cell antibody were sought by the standard immunofluorescence sandwich technique using undiluted sera on fresh blood-group-O pancreas obtained at necropsy and group-O stomach obtained at partial gastrectomy respectively.[9] Weakly positive reactions were confirmed on at least two different substrates. Thyroglobulin and thyroid microsomal antibody titres were estimated with commercial haemagglutination kits (Wellcome Reagents Ltd) of formolised red blood cells coated with the appropriate human antigens. We took titres of at least 1/20 and 1/400 respectively as positive.

Statistical significance was assessed by the χ^2 test with Yates's correction.

Results

CLINICAL AUTOIMMUNE DISEASE

Type I DM was present in seven siblings, seven fathers, and two mothers. There was no excess of type II DM (0·7% of first-degree relatives). Autoimmune thyroid disease was reported in first-degree or second-degree relatives of 37 probands (26%); they included 15 mothers and three fathers (11% and 2% respectively; $p < 0{\cdot}01$). Most had a history of thyrotoxicosis, but three mothers had primary myxoedema and two euthyroid goitres. Other organ-specific autoimmune diseases were less common, including pernicious anaemia (three families), vitiligo (five), and coeliac disease (three). Nineteen probands (13%) had first- or second-degree relatives with a history of rheumatoid arthritis, which is the first time that such an association has been noted. Fifty-seven families (40%) reported one or more of these conditions, excluding diabetes, in first- or second-degree relatives.

TABLE I—*Prevalence of autoantibodies in non-diabetic siblings of probands*

		HLA haplotypes in common with proband							
		Both		One		Neither		Total	
No (%) of siblings with ICA	M	1 (5)	4 (9)	4 (5)	6 (4)	0	0	5 (4)	10 (4)
	F	3 (13)		2 (4)		0		5 (4)	
No (%) of siblings with TGA	M	2 (10)	7 (16)	5 (6)	12 (9)	4 (13)	10 (15)	11 (8)	29 (12)
	F	5 (21)		7 (13)		6 (17)		18 (16)	
No examined	M	21	45	82	136	32	67	135 (54)	248
	F	24		54		35		113 (46)	

ICA = Islet-cell antibody.
TGA = Thyroid (microsomal or thyroglobulin) or gastric parietal-cell antibody.

ICA was found in five fathers and three mothers (a prevalence in parents of 3%) and in 10 siblings (4%). Interestingly, all the non-diabetic siblings with ICA were potentially susceptible to diabetes, since they had at least one HLA haplotype in common with their probands (table I). Thyroid or gastric parietal-cell antibody (TGA) was present in 56 mothers and 26 fathers (41% and 20% respectively; $p < 0{\cdot}001$) and in 29 siblings (12%; female excess not significant). In contrast to the distribution of ICA, the distribution of TGA among the non-diabetic siblings (table I) was independent of the degree of HLA identity with the proband. There were no significant differences between these groups in age or sex distribution. The same pattern was observed when thyroid and gastric parietal-cell antibodies were analysed separately.

Table II shows the prevalence of autoantibodies in non-diabetic parents and siblings selected for the presence of a single haplotype carrying HLA-B8. In two-thirds of cases the proband also had this haplotype, which was termed the "diabetogenic" B8 haplotype, whereas in the other cases it was termed "non-diabetogenic." The prevalences of TGA in these two groups were almost identical and the age and sex distributions comparable.

TABLE II—*Prevalence of autoantibodies in non-diabetic first-degree relatives with one HLA-B8-positive haplotype*

	"Diabetogenic" HLA-B8 haplotype	"Non-diabetogenic" HLA-B8 haplotype
No (%) with ICA	6 (6)	2 (4)
No (%) with TGA	26 (25)	14 (27)
No examined	105	51

ICA = Islet-cell antibody.
TGA = Thyroid (microsomal or thyroglobulin) or gastric parietal-cell antibody.

Discussion

Islet-cell antibody may serve as a predictive marker for insulin-dependent diabetes in older patients initially treated with oral hypoglycaemic agents.[14] Our results show that ICA may occur in children before the onset of symptomatic insulin-dependent diabetes. Furthermore, ICA was detected only in those who were genetically potentially susceptible to DM. Interestingly, one of the siblings with ICA, who had both HLA haplotypes in common with the proband, abruptly developed symptoms of diabetes several months after ICA was detected and after a period of asymptomatic, persistent postprandial hyperglycaemia.

If thyrogastric autoimmunity were determined by the same S_1 gene as type I DM, then non-diabetic siblings who had both haplotypes in common with their probands would have been expected to show a greater prevalence of TGA than those with neither of them. That this was not so (table I) provides evidence for different genes determining the susceptibilities to type I DM and to other autoimmune endocrinopathies. Though

relatively few siblings with TGA are shown in each group in table I, this hypothesis was supported by the results of analysing autoantibodies in relation to HLA-B8-haplotype segregation (table II). By considering only those first-degree relatives with a single B8-positive haplotype we could incorporate the data on parents as well as siblings, which showed no direct association between the occurrence of TGA and the likely presence of an S_1 diabetogenic gene (table II). Alternatively, genes outside the HLA region may control the production of organ-specific autoantibodies.

A history suggesting rheumatoid arthritis in first-degree or second-degree relatives was obtained in 13% of families. Although the diagnosis could not always be proved, this finding is interesting because of increasing evidence for a positive association between rheumatoid arthritis and HLA-DRW4,[15] which is also a susceptibility factor in type I diabetes.

We have presented strong evidence for genetic heterogeneity within the range of autoimmune endocrinopathies, and our findings reinforce the view that the presence of ICA in the serum of children genetically susceptible to type I DM may be a predictive marker for the disease at a relatively early stage of pathogenesis. This would have important implications should it prove possible to prevent irreversible β-cell damage in such children.

This study was supported by the Medical Research Council (ANG), the British Diabetic Association, the Juvenile Diabetes Foundation, and the Joint Research Board of St Bartholomew's Hospital. We thank Miss B Watson, Miss V Drummond, Miss J Pritchard, and Dr E Wolf for their accurate typing of HLA, and Miss M Shattock for help with the characterisation of autoantibodies. We also thank Drs W B Thomson, D H Garrow, C H Cheetham, J Scott, and other physicians for allowing us to study their patients.

References

[1] Cudworth AG. Type I diabetes mellitus. *Diabetologia* 1978;**14**:281-91.
[2] Cudworth AG, Festenstein H. HLA genetic heterogeneity in diabetes mellitus. *Br Med Bull* 1978;**34**:285-9.
[3] Sachs J, Cudworth AG, Jaraquemada D, Gorsuch AN, Festenstein H. Type I diabetes and the HLA-D locus. *Diabetologia* (in press).
[4] Cudworth AG, Woodrow JC. Evidence for HLA-linked genes in "juvenile" diabetes mellitus. *Br Med J* 1975;iii:133-5.
[5] Nerup J, Cathelineau C, Seignalet J, Thomsen M. In: Dausset J, Svejgaard A, eds. *HLA and diseases.* Copenhagen:Munksgaard, 1977;149-67.
[6] Cudworth AG. The HLA system, autoimmune endocrinopathy and diabetes mellitus. *Eur J Clin Invest* 1978;**8**:355-7.
[7] Irvine WJ, Clarke BF, Scarth L, Cullen DR, Duncan LJP. Thyroid and gastric autoimmunity in patients with diabetes mellitus. *Lancet* 1970; ii:163-8.
[8] Bottazzo GF, Mann JI, Thorogood M, Baum JD, Doniach D. Autoimmunity in juvenile diabetics and their families. *Br Med J* 1978; ii:165-8.
[9] Bottazzo GF, Florin-Christensen A, Doniach D. Islet cell antibodies in diabetes mellitus with autoimmune polyendocrine deficiencies. *Lancet* 1974;ii:1279-82.

[10] Lendrum R, Walker G, Cudworth AG, *et al.* Islet cell antibodies in diabetes mellitus. *Lancet* 1976;ii:1273-6.
[11] Cudworth AG, Bottazzo GF, Doniach D. Genetic and immunological factors in type I diabetes mellitus. In: Irvine WJ, ed. *Immunology of diabetes*. Edinburgh:Teviot Scientific Publications (in press).
[12] Bottazzo GF, Cudworth AG, Moul DJ, Doniach D, Festenstein H. Evidence for a primary autoimmune type of diabetes mellitus. *Br Med J* 1978;ii:1253-5.
[13] Joysey VC, Wolf E. HLA-A, -B and -C antigens. Their serology and cross reactions. *Br Med Bull* 1978;**34**:217-22.
[14] Irvine WJ, Gray RS, McCallum CJ. Pancreatic islet-cell antibody as a marker for asymptomatic and latent diabetes and pre-diabetes. *Lancet* 1976;ii:1097-1102.
[15] Panayi GS, Wooley P, Batchelor JR. Genetic basis of rheumatoid disease: HLA antigens, disease manifestations, and toxic reactions to drugs. *Br Med J* 1978;ii:1326-8.

(*Accepted 17 October 1979*)

Printed in Great Britain by Qualitex Printing Limited, Cardiff

EVIDENCE FOR A LONG PREDIABETIC PERIOD IN TYPE I (INSULIN-DEPENDENT) DIABETES MELLITUS

A. N. GORSUCH, K. M. SPENCER, J. LISTER, J. M. MCNALLY, B.M. DEAN, G.F. BOTTAZZO, A.G. CUDWORTH

Medical Unit, St Bartholomew's Hospital, London; Department of Immunology, Middlesex Hospital, London; and King Edward VII Hospital, Windsor

Summary In a prospective investigation of the prediabetic period before onset of type 1 (insulin-dependent) diabetes, HLA genotypes were determined in 582 healthy parents and siblings of 160 affected children. Islet cell antibody was sought by both the conventional (ICA-IgG) and the complement fixation (CF-ICA) techniques during regular prospective observation over a mean period of 2·0 years. 4 siblings and 2 parents became diabetic; the interval before detection of any biochemical abnormality exceeded a year in 4 of these (range 3–30 months), and in all cases ICA-IgG was positive from the outset, CF-ICA being positive in 5. These observations suggest that the initiation of pathogenesis may precede the abrupt clinical onset of diabetes by several years, even in children. This has important implications, both for research and for possible future prophylaxis.

Introduction

THE aetiology of type 1 (insulin-dependent) diabetes mellitus involves both an inherited susceptibility and environmental factors.[1] The exact details of these interacting factors and the mechanism of pathogenesis are as yet poorly understood. One reason for this is that by the time a child is clinically diabetic destruction of pancreatic islet B cells is nearly complete,[2] whereas it is the early stages of the process which are of particular interest.

In order to investigate events before onset of type 1 diabetes, the Barts-Windsor Prospective Family Study was started in 1978. Two concepts form the rationale for this project. The first is well established: susceptibility to this disease shows linkage to HLA haplotype inheritance within families, whatever the actual HLA phenotypes may be,[3] so that a child who is HLA-identical to his diabetic sibling, sharing both haplotypes, is himself likely to be susceptible. The second concept is that islet cell antibody (ICA), especially when complement-fixing (CF-ICA), may act as a marker of islet B cell damage before progression to insulin insufficiency.[4] These antibodies are present in most diabetic children at the time of diagnosis.[5]

During the first 3 years of this prospective study diabetes has developed in 6 family members. Here we present unexpected observations concerning the prediabetic period.

Subjects and Methods

Ascertainment

One hundred and sixty families were ascertained through type 1 diabetic probands attending diabetic clinics in East Berkshire and surrounding districts. Other criteria for inclusion were availability of at least 1 sibling aged less than 20 years, and informed consent. Any half-siblings were excluded. Table I gives details of the subjects included.

Follow-up

Four families were not followed up because all children were diabetic. The remaining one hundred and fifty-six were reviewed and tested for ICA every 1–4 months. All family members were asked to test for glycosuria at least once a month with 'Clinitest' or 'Diastix' (Ames) after a meal. A record was kept of the results and of any intercurrent infections. Glycosylated haemoglobins (HbA_1) were assayed at least twice during 1979 in all subjects. When a non-diabetic was found to be ICA-positive, by previous agreement this was not disclosed, but random venous blood glucose and HbA_1 values were checked at most visits.

By late 1980 sixteen families had withdrawn from the full protocol, but occasional contact was maintained for at least 2 years. For the remaining one hundred and forty families the mean follow-up period at the time of this analysis is 2·0 years (range 1·1–2·5).

HLA Typing

HLA genotypes were determined in all families. The standard microlymphocytotoxicity technique was employed, with sera covering all the Caucasoid HLA-A, B, and Cw specificities recognised in 1978. More recently many families have been HLA-DR genotyped but this work is not yet complete.

Islet Cell Antibodies

Conventional ICA (ICA-IgG) and CF-ICA were determined as previously described.[4] Weakly positive reactions were confirmed on a different substrate pancreas by independent observers.

Glycosylated Haemoglobin

Total HbA_1 was measured by microcolumn chromatography as described elsewhere,[6] after storage of the samples for 3–10 days at 4°C. Values in ICA-negative normal subjects in the present study were 8·7±SD 0·9%.

Blood Glucose

Venous blood was collected into fluoride-oxalate and stored overnight at 4°C before assay by a glucose oxidase/peroxidase 'Auto- Analyzer' technique.

Islet Cell Function

In some subjects, B cell secretory capacity was assessed by measurement of plasma glucose, insulin, and C-peptide responses to a standard breakfast (65 g carbohydrates, 20 g fats, 17 g protein) after an overnight fast. A control panel of healthy individuals of similar age was available for comparison. Details of the radio-immunoassay are described elsewhere;[7,8] the sensitivities and coefficients of variation are, for insulin, 1 mU/l and 6·2%, and for C-peptide 0·05 nmol/l and 7·7%.

Results

Of the non-diabetic subjects in table I, serum from 20 parents (7%) and 34 siblings (12%) was positive for conventional ICA. CF-ICA was found in 6 parents (2%) and 12 siblings (4%). CF-ICA positive sera were invariably also positive by the conventional technique employing fluoresceinated anti-human-IgG. 4 siblings and 2 parents have become diabetic during the prospective study. In all cases ICA was present throughout the period of observation.

TABLE I—SUBJECTS STUDIED IN 160 FAMILIES

—	Parents		Probands		Siblings*	
	Fathers	Mothers	Male	Female	Brothers	Sisters
Total no.	154	159	86	74	154	134
Age at entry (yr):						
Mean	43.6	40.4	13.8	13.9	14.0	14.6
Range	27–62	26–59	5–24	4–29	1–29	0–32
No. not diabetic	147	157	—	—	152	126

*A further 5 brothers and 10 sisters were unavailable for study. Half-siblings were excluded.

TABLE II—DEVELOPMENT OF DIABETES IN 6 SUBJECTS UNDER PROSPECTIVE OBSERVATION

Case no.	Sex	Relationship to proband	Entry to study: Age (yr)	Entry to study: Date	Islet cell antibody: ICA-IgG	Islet cell antibody: CF-ICA	Onset of diabetes: Date of first abnormality	Onset of diabetes: Date started insulin	Period ICA positive before first metabolic abnormality (mo)
1	M	Brother	3	June 18, 1978	+	+	Oct. 9, 1980	Oct. 9, 1978	28
2	F	Sister	17	July 11, 1978	+	+	Oct. 10, 1978	Jan. 27, 1979	3
3	F	Sister	19	Sept. 5, 1978	+	+	Feb. 22, 1981	—	30
4	M	Brother	18	Nov. 26, 1978	+	−	Oct. 13, 1980	Oct. 19, 1980	23
5	F	Mother	44	Jan. 28, 1979	+	+	June. 16, 1980	Dec. 16, 1980	16
6	M	Father	47	April 11, 1979	+	+	Aug. 19, 1979	—	4

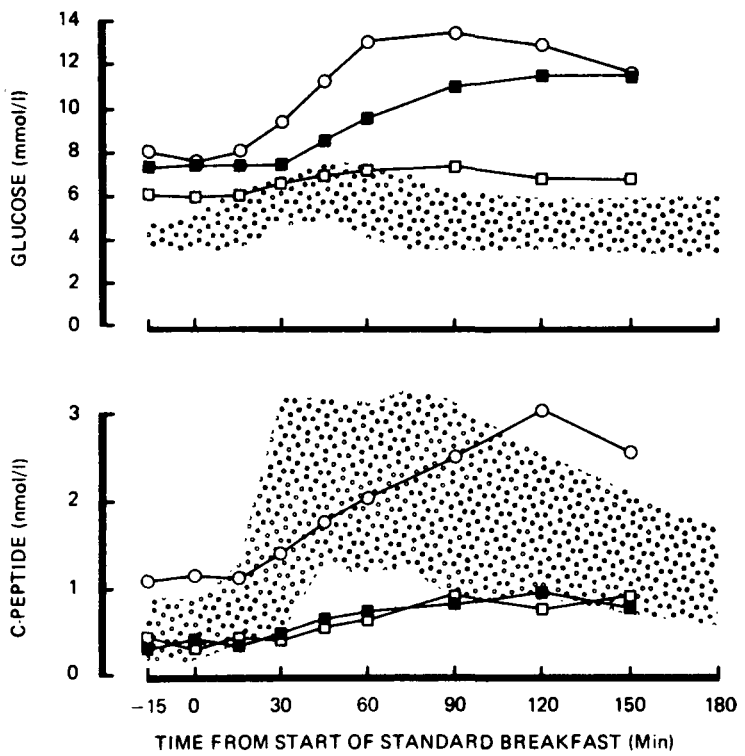
Case reports are presented below, and some features are summarised in table II.

Case 1

This child shares only the maternal HLA haplotype with his 2 diabetic brothers, one of whom was diagnosed long before, and one shortly after his own birth in 1975. Other relatives have vitiligo, myxoedema, or atopy, but not diabetes. CF-ICA was strongly positive when first tested for at entry to the study, and remained so throughout the next 28 months. During this period he had fourteen recognised upper respiratory infections as well as one attack of mild gastroenteritis and one of conjunctivitis. Up till Sept. 23, 1980, urine tests (33) showed no glycosuria, postprandial blood glucose ranged from 4·9 to 6·4 mmol/l (1 mmol/l = 18·02 mg/dl), and HbA_1 ranged from 7·7 to 9·4%, the latest blood test being in March, 1980. Excess thirst began on Oct. 2, 1980, six days after onset of an upper respiratory tract infection and nine days after the last (negative) urine test. On Oct. 9 there was 2% glycosuria but no ketonuria, and blood glucose was 15·1 mmol/l. Insulin treatment was started at once.

Case 2

This girl was born 3½ years after the diagnosis of diabetes in her elder brother, to whom she is HLA-identical. No other relative is know to be diabetic; the mother has myxoedema and a maternal aunt had thyrotoxicosis. At entry to the study ICA was strongly positive by both techniques. 1 month later the postprandial blood glucose was 6·7 mmol/l, but 3 months after entry it was 10·6 mmol/l, although 3 urine tests had shown no glycosuria and there had been no recent infections. On Oct. 23, 1978, islet B cell function studies were abnormal (see figure). At the next two monthly visits postprandial blood glucose levels were 9·2 and 15·2 mmol/l, respectively, and frequent urinalysis showed intermittent but increasing glycosuria. CF-ICA remained strongly positive. Not until the end of January, 1979, did she have diabetic symptoms with moderate ketosis, and thus require insulin therapy.



Islet B cell function studies.

Plasma concentrations of glucose (upper) and C-peptide (lower) in 3 subjects after standard breakfast at time 0; ■———■ = case 2, Oct 23, 1978; □———□ = case 5, June 16, 1980; ○———○ = case 6, Feb 1, 1980. Insulin levels, which parallel C-peptide levels, are omitted for clarity. Shaded areas represent normal ranges (mean ± 2 SD, n = 8).

Case 3

This girl was born a year after onset of diabetes in her HLA-identical sister. When first tested for ICA at age 18¾ yr, she was strongly positive by both techniques. Apart from a cold soon afterwards she reported no infections during the follow-up period. Between 1978 and late 1980 21 urinalyses were negative for sugar; nine random blood glucose estimations (3·6–6·0 mmol/l) and seven HbA_1 assays (7·8–10·1%) were normal, the last being in June, 1980. An oral glucose tolerance test showed diabetes in February, 1981, with a fasting venous blood glucose of 6·8 mmol/l, a peak value of 13·6 mmol/l at 60 min, and a 2-hour value of 13·0 mmol/l after a 75 g glucose load. 2 months later there was still no glycosuria and she remained symptom-free with a random blood glucose of 7·3 mmol/l. Both types of ICA remain strongly positive.

Case 4

This boy is haploidentical to his diabetic younger sister, diagnosed in 1966. At entry to the study he was weakly positive for conventional ICA but negative for CF-ICA. Over the next months there were three upper respiratory infections. HbA_1 was normal on two occasions (10·1 and 8·6%). Unfortunately no further blood tests were performed until 16 months later when he reported weight loss for 2 months, increasing thirst, and glycosuria. Blood glucose was 16·5 mmol/l and he was admitted for insulin treatment. ICA status was unchanged.

Case 5

This subject's mother has myxoedema and type 1 diabetes; in addition there is a family history of type 1 diabetes in an aunt, an uncle, and her younger daughter (the proband, diagnosed in 1973). She herself has asthma but does not take corticosteroid drugs. Throughout the study ICA has remained positive by both techniques. Despite transient glycosuria on three occasions during the 15 months before insulin dependence supervened, random blood glucose (3·3–5·9 mmol/l) and HbA_1 (7·9–10·1%) were normal at all ten visits up to August, 1980. B cell function studies on June 16, 1980, showed a subnormal C-peptide response (see figure). On Nov. 2, 1980, seven days after an upper respiratory infection, an oral glucose tolerance test showed a diabetic glucose curve and impaired insulin response, although urine tests were still negative. She later became symptomatic and was admitted for insulin treatment on Dec. 16; ketonuria was detected on one occasion.

Case 6

The proband in this family was diagnosed diabetic in December, 1978. A few days later the father tested his own urine and found 2% glycosuria, but several further tests were negative. There are no diabetic relatives apart from the proband; an aunt has pernicious

anaemia. On entering the prospective study this father was positive for CF-ICA. Thereafter monthly urinalysis was negative until Aug. 19, 1979, two weeks after an upper respiratory infection, when there was again glycosuria, the evening blood glucose was 13·6 mmol/l, and HbA_1 was 13·3%. Over the next two weeks there was frequent glycosuria and occasional ketonuria (up to ++ with Ames 'Acetest'). After a standard breakfast the peak blood glucose was 15·7 mmol/l at 90 min. He remained symptom-free. As he was overweight (116% ideal body weight), treatment was started with a reducing diet. Over the next 18 months there were three further brief episodes of repeated glycosuria but otherwise the urine tests remained negative with random postprandial blood glucose below 10 mmol/l. Conventional ICA remains positive but CF-ICA became undetectable in September, 1980.

Discussion

The most striking feature of these prospective observations is the strong evidence they provide for a long latency of type 1 diabetes—a condition usually considered to have a short time-course. In the typical childhood diabetes exemplified by the probands in this study there is very little information about the period before the first symptoms. Two retrospective studies[9,10] have suggested that symptomatic onset may not be abrupt. The observations reported here extend this concept; in every male subject who became diabetic during the follow-up period, ICA was already present at the first test, up to 30 months before any biochemical abnormality. If ICA is acceptable as a marker of prediabetes in these cases then the beginning of the pathogenetic process may well be several years before the diagnosis—even perhaps in infancy (case 1).

It is well established that a prolonged autoimmune process underlies certain other endocrine diseases such as primary myxoedema and Addison's disease and that the corresponding organ-specific antibodies may be present for many years before diagnosis.[11,12] There are a few reports[13,15] of diabetes developing after some years in ICA positive adults with such diseases. These are in keeping with the theory that the "autoimmune" or "type 1B" variant of diabetes has a slow time course.[16,17] A possible interpretation of our observations is that a similar prolonged process may play a role in juvenile type 1 diabetes. It remains to be explained, however, why ICA generally disappears so rapidly after diagnosis in this condition, in contrast to the behaviour of autoantibodies in type 1B diabetes and thyroid and other autoimmune diseases.

It might be argued that 2 of our cases do not represent type 1 diabetes, in that the patients are not insulin-dependent. However, in neither is there a close family history of type 2 diabetes. Furthermore, patient 3 has so far been recognised as diabetic for only a few months, whereas in patient 6 ketonuria has been detected several times. Both these subjects had CF-ICA, which is very rare in type 2 diabetes (unpublished observations).

If there is indeed a long prediabetic period in most type 1 diabetes, this may seem difficult to reconcile with the theory that viral infections occurring shortly before onset of clinical diabetes may have an aetiological role.[18,21] One possible explanation is that these agents merely have a non-specific "stress" effect, triggering off metabolic decompensation when islet B cell function is already compromised by a process beginning much earlier. An alternative possibility is that B cell destruction is a cumulative process, each stage being initiated by a new environmental event, the repeated episodes eventually destroying sufficient B cells to cause diabetes. In this respect it is interesting that 3 of our patients (cases 1, 5, and 6) had upper respiratory tract infections within two weeks before the first evidence of diabetes.

An important question raised by these findings is whether type 1 diabetes can reliably be predicted in relatives of affected children. This is to be considered in detail in another report. It is, however, noteworthy that standard biochemical tests (random blood glucose, HbA_1, regular urinalysis) were of little value as early warning signals of impending diabetes, unlike HLA genotyping and testing for CF-ICA.

Finally, there are implications for research aimed at identifying the factors initiating pathogenesis. If the latent period does indeed extend to several years, then in studies such as ours it will be necessary to start prospective surveillance of unaffected siblings earlier in childhood and continue longer than originally planned. On the other hand, there is the future possibility of intervention to arrest B cell destruction. If such prophylaxis should become available it would probably be facilitated by a long prediabetic period.

This project has been supported by the Medical Research Council, the British Diabetic Association, the Juvenile Diabetes Foundation, the Joint Research Board of St Bartholomew's Hospital, and Miles Laboratories Ltd. We gratefully acknowledge the help given by the physicians and paediatricians in the area of the study in allowing us to study their patients. We also thank Dr Sami Medbak for the C-peptide studies and Dr Simon Welch for the HbA_1 measurements.

Correspondence should be addressed to A. G. C., Medical Unit, St Bartholomew's Hospital, London EC1A 7BE.

Addendum

Type 1 diabetes developed in another child, age 9½ yr, in November, 1981. He is HLA-identical to his diabetic elder brother and had been strongly positive for CF-ICA since the first test in October, 1978.

REFERENCES

1. Cudworth AG. Current concepts in the aetiology of type 1 (insulin dependent) diabetes mellitus. In: Bellingham AJ, ed. Advanced medicine 16. London: Pitman Medical; 1980: 123–35.
2. Gepts W. Pathological anatomy of the pancreas in juvenile diabetes mellitus *Diabetes* 1965; **14:** 619–33.
3. Walker A, Cudworth AG. Type 1 (insulin dependent) multiplex families: mode of genetic transmission. *Diabetes* 1980; **29:** 1036–39.
4. Lendrum R, Walker G, Cudworth AG, Theophanides C, Pyke DA, Bloom A, Gamble DR. Islet cell antibody in diabetes mellitus. *Lancet* 1976; ii: 1273–76.
5. Bottazzo GF, Dean BM, Gorsuch AN, Cudworth AG, Doniach D. Complement fixing islet cell antibodies in type 1 diabetes: possible monitors of active beta-cell damage. *Lancet* 1980; i: 668–72.
6. Welch S, Boucher BJ. A rapid micro-scale method for the measurement of haemoglobin A1 (a+b+c). *Diabetologia* 1978; **14:** 209–11.
7. Morgan CR, Lazarow A. Immunoassay of plasma insulin levels of normal, sub-diabetic and diabetic rats. *Diabetes* 1963; **12:** 115–26.
8. Heding LG. Radioimmunological determination of human C-peptide in serum *Diabetologia* 1975; **11:** 541–48.
9. Hamilton DV, Mundia SS, Lister J. Mode of presentation of juvenile diabetes. *Br Med J* 1976; ii: 211–12.
10. Cudworth AG, Gamble DR, White G BBB, Lendrum R, Woodrow JC, Bloom A. Aetiology of juvenile onset diabetes: a prospective study *Lancet* 1977; i: 385–8.
11. Doniach D, Bottazzo GF. Polyendocrine autoimmunity. In: Franklin EC, ed. Clinical immunology update. New York: Elsevier, 1981: 95–121.
12. Gorsuch AN, Cudworth AG. Immunological aspects of endocrine disease in childhood. In: Soothill JF, Greenwood H, Hayward AR, Wood CBS, eds. Paediatric immunology Oxford: Blackwell (in press).
13. Bottazzo GF, Florin-Christensen A, Doniach D. Islet cell antibodies in diabetes with autoimmune polyendocrine deficiencies. *Lancet* 1974; ii: 1279–82.
14. Irvine WJ, Sawers JSA, Feck CM, Prescott AJ, Duncan LJP. The value of islet cell antibody in predicting secondary failure of oral hypoglycaemic treatment. *J Lab Immunol* 1979; **2:** 23–26.
15. Betterle C, Caretto A, Tiengo A, Trevison A. Complement-fixing islet cell antibodies in type 1 diabetes in susceptible patients with autoimmune diseases. *Lancet* 1980; i: 1418–9.
16. Bottazzo GF, Doniach D. Pancreatic autoimmunity and HLA antigens. *Lancet* 1976; ii: 800.
17. Bottazzo GF, Cudworth AG, Moul DJ, Doniach D, Festenstein H. Evidence for a primary autoimmune type of diabetes mellitus (type 1b). *Br Med J* 1978; ii: 1253–55.
18. Bloom A, Hayes TM, Gamble DR. Register of diabetic children. *Br Med J* 1975; iii: 580–83.
19. Durruty P, Ruiz F, Garcia de los Rios L. Age at diagnosis and seasonable variations in the onset of insulin dependent diabetes in Chile (Southern Hemisphere). *Diabetologia* 1979; **17:** 357–60.
20. Gamble DR. Relation of antecedent illness to development of diabetes in children, *Br Med J* 1980; **281:** 99–101.
21. Yoon JW, Austin M, Onodera T, Notkins AL. Isolation of a virus from the pancreas of a child with diabetic ketoacidosis. *N Engl J Med* 1979; **300:** 1173–79.

MANNOSE-RESISTANT HAEMAGGLUTINATION AND P ANTIGEN RECOGNITION ARE CHARACTERISTIC OF ESCHERICHIA COLI CAUSING PRIMARY PYELONEPHRITIS

VUOKKO VÄISÄNEN, JAAKKO ELO, LEIF G. TALLGREN, ANJA SIITONEN, P. HELENA MÄKELÄ, CATHARINA SVANBORG-EDÉN, GUNILLA KÄLLENIUS, STEFAN B. SVENSON, HANS HULTBERG, TIMO KORHONEN

Department of General Microbiology, University of Helsinki, Aurora Hospital, and Central Public Health Laboratory, Helsinki, Finland; and Department of Clinical Immunology, University of Göteborg, Department of Bacteriology, National Bacteriological Laboratory, and Department of Organic Chemistry, Arrhenius Laboratory, University of Stockholm, Sweden

Summary Thirty-two *Escherichia coli* strains from 30 children with pyelonephritis were examined for their haemagglutination patterns and O and K serotypes. 29 (91%) of the strains showed mannose-resistant haemagglutination (MRHA). By use of well-defined target cells, these MRHA$^+$ strains could be shown to recognise human cells either in a P-specific manner (recognising a specific galactosyl-galactose structure which is part of P blood group antigens) or in a separate, X-specific manner. Both recognition mechanisms could occur separately or together on the same bacteria, the frequencies of P and X specificity being 81 and 19%, respectively. Both MRHA and P specificity were significantly associated with the O antigens 01, 04, 06, 016, and 018, and the capsular antigen K1, which have previously been associated with pyelonephritis. However, the association of MRHA and P specificity with upper urinary tract infection in children is greater than that of any other laboratory-defined bacterial characteristic.

Introduction

ADHESION to epithelial cells is a characteristic of *Escherichia coli* strains isolated from patients with pyelonephritis as opposed to strains from other types of urinary tract infection (UTI) or from the faeces of healthy people.[1-3] Typically, the adhesion takes place through a reaction between the fimbriae on the bacterial cell surface[4,5] and specific receptor structures on the epithelial cells.[6,7] However, not all fimbriae are associated with uropathogenic properties. For example, the common, or type 1, fimbriae present on most enterobacterial strains[8] are found as often on strains causing UTI as on strains from other sources, and in fact may prevent adhesion to the epithelial cells of the urinary tract by trapping the bacteria in their covering mucus.[9]

Adhesion to epithelial cells is difficult to quantify, while agglutination of red blood cells is an easily measured reaction much used in clinical microbiology. The adhesion phenomena discussed here have been shown to correlate extremely well with agglutination of erythrocytes.[8] Thus, both adhesion and haemagglutination by type 1 fimbriae are mannose-sensitive (inhibited by α-methyl-D-mannopyranoside), indicating that a mannose-containing structure(s) on both erythrocytes and epithelial cells can serve as a receptor(s).[10-12] Pyelonephritis-associated *E. coli*, on the other hand, cause mannose-resistant hemagglutination (MHRA) of human red blood cells.[3,5,11,13,14] The minimal receptor structure for this recognition has been shown to be α-D-galactopyranosyl-(1→4)-β-D-galactopyranoside α-D-Gal*p*-(1→4)-β-D-Gal*p*.[6,15] This structure is part of the antigenic determinants P^K (trihexosyl ceramide) and P (globoside) of the human P blood group system, so that only individuals of the rare blood group $\bar{p}$ have cell membrane glycolipids devoid of the α-D-Gal*p*-(1→4)-β-D-Gal*p* structure.[16,17] Both trihexosyl ceramide[6] and globoside[7] have been proposed as receptors for the pyelonephritis-associated *E. coli*.

We have examined agglutination specificity with a new series of *E. coli* strains isolated from the urine of children with pyelonephritis, paying special attention to the receptor structure recognised by these strains. Our results show that both globoside and trihexosyl ceramide are recognised by the same strains in a P blood group-specific manner, inhibited by the α-D-Gal*p*-(1→4)-β-D-Gal*p* disaccharide. In addition, some pyelonephritis-associated strains can recognise, also in a mannose-resistant manner, a different receptor structure on human erythrocytes. These two types of recognition may be present alone or together on the same strains. A similar P-specific recognition in 91% of pyelonephritis strains isolated in Sweden has also been demonstrated.[18]

Patients, Materials, and Methods

Patients

Thirty-two *E. coli* strains were isolated during the acute phase of a pyelonephritic episode from midstream urine specimens from 30 children seen at the Aurora Hospital, Helsinki, between October, 1977, and December, 1979. Two strains came from 2 boys (ages 0·3 and 1·8 yr), and thirty strains came from 28 girls (mean age 4·4±3·6 yr). In the 2 cases in which two separate strains from the same patient were included, they came from episodes of infection at least 18 months apart. Neither of the boys had had a previous UTI whereas the mean number of previous UTI episodes among the girls was 2·3±3·0 (UTI defined as a urinary bacterial count of $\geqslant 10^5$/ml, or 10^4–10^5/ml on repeated culture). UTI was diagnosed as acute pyelonephritis by sudden onset of symptoms, a raised erythrocyte sedimentation rate (mean 47±21 mm/hr), and/or raised white blood cell count (mean 16·4±6·0×10^9/l).[19] In addition, 9/9 children had raised C-reactive protein levels, and 29 of the 32 episodes were associated with fever of 38°C or above. Combinations of these four symptoms and significant bacteriuria have previously been used to diagnose pyelonephritis in children.[1-3,18-20] Urography and ureterocystography were performed on 22 children (23 episodes, table I). Vesicoureteric reflux was evaluated, according to Winberg et al., as grades I or II with no evidence of upper urinary tract dilatation, or as grades III or IV with slight or gross dilatation of the upper urinary tract.[20]

Bacterial Cultures

The isolated *E. coli* strains were preserved in nutrient agar at room temperature for 1 to 3 yr until the time of testing. For haemagglutination assays they were subcultured twice on colonisation factor antigen (CFA) agar plates.[21]

Red Blood Cells

Human erythrocytes of types P_1, P_2^k, and $\bar{p}$ were supplied by Dr Anna Pirkola (Finnish Red Cross Blood Transfusion Service, Helsinki) and Dr Bertil Cedergren (Blood Bank, University Hospital, Umeå, Sweden). They were washed with phosphate buffered saline (PBS) at pH 7·1 and suspended to 2% in PBS for haemagglutination tests.

Horse erythrocytes were coated with trihexosyl ceramide or globoside (Supelco, Bellefonte, U.S.A.).[7] The structures of both glycolipids were determined by ^{1}H-nuclear magnetic resonance spectrometry using a Jeol instrument operated in the PFT mode. 50 μl of a 20% suspension of washed horse erythrocytes was added to 1 ml of a suspension of either glycolipid (200 μg/ml). The mixtures

Reprinted from DIABETES, VOL. 31, NO. 10, OCTOBER 1982
Copyright 1982 by THE JOURNAL OF THE AMERICAN DIABETES ASSOCIATION.

Can Future Type I Diabetes Be Predicted?
A Study in Families of Affected Children

A. N. GORSUCH, K. M. SPENCER, J. LISTER, E. WOLF, G. F. BOTTAZZO, AND A. G. CUDWORTH

SUMMARY

This article examines the risk of type I (insulin-dependent) diabetes in siblings of affected children, in relation to HLA genotypes. The 288 available siblings of 160 diabetic probands were grouped according to the number of HLA haplotypes in common with their probands. HLA-identical siblings (both haplotypes in common) have an approximately 100 times greater risk of developing the disease than that in the general population, and this risk is significantly higher than that in haplo-identical siblings (one haplotype in common) ($P = 0.008$). Thus, in Northern European populations, some 30% of HLA-identical siblings are expected to be diabetic by the age of 30 yr. The risk in nonidentical siblings (neither haplotype in common) is not significantly increased. These findings carry implications for genetic counseling and research. DIABETES *31*:862–866, October 1982.

It has long been recognized that diabetes shows a familial tendency.[1] Most early attempts to calculate the risk of recurrence (further cases) in the families of affected individuals failed to distinguish clearly between type I (insulin-dependent) and type II (non-insulin-dependent) diabetes, now known to be genetically distinct.[2] Recently, however, Gamble[3] has estimated from the Paediatric Register of the British Diabetic Association that 5.7% of the siblings of diabetic children are themselves affected by age 16 yr.

Hitherto, it has not been possible to predict at an early stage which siblings are destined to become diabetic. In recent years, however, two possible "risk factors" have been identified. One is HLA identity to the diabetic child, the rationale being that genetic susceptibility to this disease is principally HLA-linked.[4] The other is islet-cell antibody (ICA), which is present in around 80% of newly diagnosed diabetic children but only in about 1% of the normal population,[5] in some of whom it may be a "marker" of islet B-cell damage during the prediabetic period.

We have recently reported development of diabetes in seven first-degree relatives of type I diabetic probands while under prospective observation.[6] The presence of ICA in all seven supported the "marker" role of ICA (particularly when complement-fixing, CF-ICA[7]) and provided evidence for a long prediabetic period. Here we analyze further data from the same study to obtain an empirical estimate of risk in the siblings in relation to HLA identity to their probands, based on both new and preexisting cases.

SUBJECTS AND METHODS

Ascertainment. The subjects taking part in the Barts-Windsor Family Study (Table 1) are members of families ascertained through type I diabetic children and young adults attending diabetic clinics in East Berkshire, England, and surrounding districts. Most come from the central district, where the ascertainment rate is probably very high since the great majority of juvenile-onset and most adult-onset insulin-dependent diabetics have at some time attended one or both local diabetic clinics, the records of which are comprehensive and cover the past 20 yr. Criteria for inclusion in the study were availability of the proband and of at least one sibling aged under 20 yr, availability of both parents (or of one parent if there were three or more children), and informed consent. Selection was irrespective of how many children were diabetic. In the central area all but three of the eligible families known to us agreed to take part. Of the families entering the study during its first 2 yr, six were excluded because HLA genotypes revealed expaternity of the proband. All available first-degree relatives (but not half-siblings) of the remaining 160 probands are included in this analysis.

In "multiplex" families (those with more than one diabetic child) the first to have been diagnosed was defined as the proband. All eight such families at the time of ascertainment came from the central, East Berkshire district, with its very

From the Department of Diabetes and Immunogenetics, St. Bartholomew's Hospital Medical College, London, England.
Address reprint requests to Prof. A. G. Cudworth at the above address.
Received for publication 18 February 1982 and in revised form 24 May 1982.

TABLE 1
Subjects studied in 160 families

	Parents		Probands		Siblings*	
	Fathers	Mothers	Male	Female	Brothers	Sisters
Total number	154	159	86	74	154	134
Age at entry						
Mean (yr)	43.6	40.4	13.8	13.9	14.0	14.6
Range	27–62	26–59	5–24	4–29	1–29	0–32
Number not diabetic	147	157	—	—	152	126

* A further 5 brothers and 10 sisters were unavailable for study. Half-siblings are excluded.

high ascertainment rate. Thus, although this was not planned as an epidemiologic study, there is unlikely to be any significant bias in favor of multiplex families.

Prospective follow-up. Four families were not followed up because all children aged under 20 yr were diabetic. The remaining 156 were reviewed and tested for ICA every 1–4 mo. To facilitate early detection of new cases of diabetes all family members were asked to test for glycosuria at least once monthly using Clinitest or Diastix (Ames, Elkhart, Indiana) after meals. In addition, glycosylated hemoglobin (HbA_1) was assayed at least twice during the study.[8] When a nondiabetic was found to be ICA-positive, by prior agreement this was not disclosed, but random blood glucose and HbA_1 were checked at most visits. If any of these tests suggested hyperglycemia, the diagnosis of diabetes was confirmed or refuted according to WHO criteria.[9]

By late 1980 sixteen families had withdrawn from the full protocol, but occasional contact was maintained for at least 2 yr. For the remaining 140 families the mean follow-up period at the time of this analysis is 2.0 yr (range 1.1–2.5).

HLA typing. HLA genotypes were determined in all families. The standard microlymphocytotoxicity technique was employed, using sera covering all the Caucasoid HLA-A, -B, and -C specificities recognized in 1978. More recently many families have been HLA-DR genotyped, but this work is not yet complete.

Statistical analysis. Except where stated, differences between frequencies are tested for significance using X^2 (with Yates' correction where appropriate). Standard regression techniques are used for trend analyses.[10] The prevalence of diabetes in the siblings is compared with that in the general population using the "relative risk" (RR) or cross-product ratio, with corrections for bias and discontinuity.[11,12] This statistic has the advantage of a readily understood "medical meaning":[11] although more often used as a measure of association between a disease and a blood group or HLA phenotype in population studies, in the present context it is an estimate of the ratio of the risk of diabetes in the study population to that in the general population. It is used here because RR and its 95% confidence limits can be calculated from frequency data in control and study populations, and because the difference between two RR's can readily be tested for significance since $\log_e$ (RR) is approximately normally distributed and its variance is known.

Calculation of prevalence in the general population. There is no entirely appropriate control series of children and young adults without diabetic relatives for comparison with the siblings in this study. We have therefore used published data from two papers reporting occurrence of type I diabetes in the general population. In the first study there were 39 insulin-treated patients among 45,500 school children aged 5–16 yr in Northampton, England, in 1964, giving a prevalence of 0.09% at a mean age of 10.5 yr.[13] This figure was corrected for age by a simple mathematical procedure taking into account both the detailed age distribution of the siblings in the present study and the distribution of age at onset of the disease in published epidemiologic studies[14–17] (combined data by courtesy of Dr. D. R. Gamble).

A further correction was made to compensate for the "frequency reduction effect," that is, the apparent reduction in age-specific cumulative morbidity rate when calculated from prevalence data rather than prospectively from incidence data.[18] The correction factor used (1.1) was obtained by comparing Danish incidence and prevalence figures for the first three decades of life.[14,19] The corrected frequency is 66/45,500, a prevalence of 0.15%.

Similarly, data from an incidence study of insulin-dependent diabetes in Denmark[13] were age-adjusted: over a period of 5 yr 474 new cases were recorded from a population of 716,285 aged 0–29 yr. The frequency expected if this population had the same age distribution as the siblings in the present study would be 1629/716,285, or 0.23%.

RESULTS

Table 2 shows the siblings grouped according to whether they had two, one, or none of the parental HLA haplotypes in common with their probands, these groups being termed HLA-identical, haplo-identical, and nonidentical, respectively. When those with ambiguous assignment to these groups are excluded, the proportions of siblings in the three respective groups are 18%, 56%, and 26%. This distribution differs significantly from that of 25%, 50%, and 25% "expected" on the basis of random assortment of parental haplotypes in children ascertained in this way (see DISCUSSION). In the three cases of intra-HLA recombination so far recognized, siblings are grouped according to B-locus

TABLE 2
Siblings grouped according to HLA-A, -B, -C haplotype concordance with probands

Number of haplotypes in common with proband*	2	2/1†	1	1/0	0	Total
Number of siblings‡	50	8	151	8	71	288
Number diabetic at entry	5	0	4	0	1§	10
(%)	(10)		(3)		(1)	(3)

* Intra-HLA recombinants are grouped according to B-locus identity to proband because not all families have been HLA-DR genotyped (see text).
† 2/1: either two or one haplotype in common with proband, but ambiguous because the two haplotypes in one parent are indistinguishable (10 families).
‡ Distribution of 272 siblings between the three unambiguous groups, compared with the 68, 136, 68 "expected" (see text): $X^2_{(2)} = 6.55$; $P = 0.038$
§ This child has an HLA-B/DR recombinant haplotype and is thus identical to her proband for the DR region of that haplotype.

identity; DR-locus identity would be preferable, but it is not used because DR genotyping is not yet complete.

At the time of entry to the study 9 parents (2.9%) and 10 siblings (3.5%) were already diabetic and insulin dependent. The diabetic siblings were distributed as shown in Table 2; there was no significant sex difference (2 male, 8 female; P = 0.055 by exact binomial test).

During the follow-up period six subjects from these families developed diabetes. Two of these were parents, two were HLA-identical siblings, and two haplo-identical siblings. A further case occurred in an HLA-identical sibling after completion of this analysis and is not included (see DISCUSSION). Individual case reports are given elsewhere.[6]

Current prevalence of diabetes and relative risk in siblings. At the time of this analysis, 11 (3.5%) of the parents and 14 (5%) of the siblings have type I diabetes. The prevalence in the siblings (grouped according to HLA-ABC genotype identity to their probands) is shown in the upper section of Table 3. Also given in this table are the relative risks for diabetes calculated by comparing the frequencies in the siblings with the age-adjusted frequencies derived from the Northampton study (middle section of Table 3) and the Copenhagen study (bottom section) as low and high estimates of general population frequency. Whichever set of control figures is used, the RR for HLA-identical siblings is significantly higher than that for haplo-identical siblings (P = 0.011 and 0.008, respectively; one-tailed test assuming normal distribution of $\log_e$RR). In the case of nonidentical siblings, only one of whom is diabetic, the Woolf-Haldane method is inappropriate, but the binomial test shows that the frequency of diabetes is not significantly greater than that in controls.

DISCUSSION

These findings provide further support for the well-established concept that genetic susceptibility to this disease is mainly HLA-linked, in that the prevalence in the siblings is significantly associated with increasing HLA haplotype concordance with probands (Table 3). This point is underlined by the finding that the only nonidentical sibling who is diabetic has a B-DR recombinant haplotype, so that if DR-locus identity were the criterion used (as it would be if all families had been DR-typed), even she would be classed as haplo-identical.

The second and major point of this article is the risk of diabetes in siblings of affected children. An empirical approach to this problem, taking into account HLA haplotype concordance, is important. Others have used indirect calculations based on haplotype concordance frequencies observed in pairs of siblings selected because both members of each pair were diabetic.[20,21] They assume that 25%, 50%, and 25% of all siblings of diabetic children are, respectively, HLA-identical, haplo-identical, and nonidentical to them, these being the proportions predicted by Mendelian rules in sibships randomly ascertained with respect to HLA.

However, the present series, which is believed to be the largest of its kind, shows a significant difference from that distribution, with only 18% in the HLA-identical group (Table 2). This difference is not due to errors in HLA typing, which was repeated in case of any doubt; in fact, the majority of families have now been typed at least twice without any evidence of incorrect assignment to haplotype concordance groups on the first occasion. Furthermore, serologic problems with HLA typing should not lead to a systematic error of this kind. Nor can the discrepancy be explained by biased selection of families (see SUBJECTS AND METHODS).

We believe the true explanation to be as follows. Although the proband-sibling pairs in this (or any other) sample come from a population of families in which HLA haplotype segregation is (presumably) random, the selection process necessarily introduces a bias. This is a consequence of the fact that the probands have an HLA-linked disease, and can be predicted using simple probability,[22] as illustrated in the following example.

For simplicity, let us consider first a large, hypothetical series of sibships, each consisting of two healthy children. We assume that within each sibship there is random assortment of the four possible haplotype combinations. If one child from each sibship were chosen arbitrarily as the index case, then the proportions of all siblings of these index cases falling into the three haplotype-concordance classes would indeed be close to 25%, 50%, and 25%. Consider now, for the sake of argument, a series of similar families selected for presence of a specific HLA allele (Bw44, for example) in one paternal and one maternal haplotype; 25% of the children would be Bw44-homozygous, but 56.25% [i.e., $(1 - 0.25)^2 \times 100$] of the families would have no such child, while 6.25% would have two. If, now, families were ascertained from this series through Bw44-homozygous children (the probands) to a genetic study, 43.75% of all the families would be included. Since each proband has one sibling, and only 6.25% of the original series of families include two homozygotes, the expectation is that 6.25 × 100/43.75 = 14.29% of the siblings of the probands in the study would be HLA-identical to them—not the 25% that might have been anticipated.

A similar argument can be applied to diabetes:[22] because of the combination of ascertainment of families through affected children and HLA-linkage of the disease, the proportion of siblings expected to be HLA-identical to the pro-

TABLE 3
Current prevalence of diabetes in siblings grouped by HLA genotype identity to probands

Group	HLA-identical	Haplo-identical	Non-identical	All siblings*
Number of siblings	50	151	71	288
Number diabetic	7	6	1	14†
Prevalence (%)	14	4	1	5
Compared with N:‡ relative risk (95% confidence limits)	118 (52–265)	31 (13–69)	NS	36 (20–64)
P	$<10^{-10}$	$<10^{-10}$	0.09§	$<10^{-10}$
Compared with D:‡ relative risk (95% confidence limits)	76 (35–164)	20 (9–43)	NS	23 (14–39)
P	$<10^{-10}$	$<10^{-10}$	0.15§	$<10^{-10}$

* Including those with ambiguous HLA genotypes (see text).
† Distribution among nonambiguous groups compared with that expected if no linkage between HLA and diabetes: P = 0.0053. Analysis for quadratic trend: P < 0.005.
‡ Age-corrected frequencies of diabetes in general population samples (see text): N, Northampton, 1964:[13] 66/45,500 (0.15%); D, Denmark, 1977:[14] 1629/716,285 (0.23%).
§ One-tailed binomial test.

bands is less than 25%. The extent of this bias depends on the penetrance of the susceptibility genes (which, of course, is related to age and other factors), their mode of inheritance, the closeness of linkage to HLA, and family size.

Using the indirect method referred to above, Platz and co-workers[21] calculated that the prevalence expected in HLA-identical, haplo-identical, and nonidentical siblings in Denmark at age 16 yr would be about 12%, 4%, and 1%, respectively. Despite the assumption of HLA-identity in 25% of siblings, the similarity of these estimates to the prevalences actually observed here (Table 3) is striking, even though they are not strictly comparable because the populations concerned have different age distributions.

There is in theory a possibility of bias due to more assiduous screening for diabetes in our families than in the general population. In fact all but one of the 14 diabetic siblings developed characteristic symptoms and would undoubtedly have been diagnosed in any event. The single exception, who is HLA-identical to her proband, is not yet insulin dependent; but not included in this analysis is a seventh child, also HLA-identical to his proband, who has since developed symptomatic type I diabetes. Therefore, inclusion of the asymptomatic case is unlikely to cause any significant bias.

For the calculation of RR in the siblings in our study, it is not clear whether the estimate of general population frequency based on the Northampton data or that from the Copenhagen data is the more appropriate. Both have therefore been used. Gamble has argued that the Danish figures can be applied to the United Kingdom, and has estimated by a different method that the overall risk in siblings of diabetic children at age 16 yr is 27 times greater than that in the general population.[3] This is close to our estimate of 23 using the Copenhagen data, but is also well within the 95% confidence limits of that using the Northhampton data (Table 3). Whichever "control" figures are used, when siblings are grouped according to HLA haplotype concordance with the probands (Table 3), two features stand out. The RR for HLA-identical siblings is extremely high, and it is considerably and significantly higher than that for haplo-identical siblings. These results reinforce other evidence that HLA-linked susceptibility to type I diabetes is not determined by a single, dominant gene, and are compatible with theories involving recessive or intermediate inheritance or interaction of two or more genes.[4,21,23,24]

Finally, we return to the question of predicting recurrence of this disease in particular sibships. It seems reasonable to suggest that for each group of siblings the mean of the RRs calculated from the two sets of control data is a fair approximation to the true value. In that case, if HLA genotypes are disregarded, a sibling of an affected child carries approximately a 30-fold increased risk, and if the two are HLA-identical, the risk is increased around 100-fold. Thus some 30% of these siblings would be diabetic by age 30 yr, in Northern Europe, assuming that the general population prevalence at that age is 3 per 1000 (a conservative estimate from the Northampton study and from the combined age-at-onset data,[14–17] similar to the prevalence observed in Denmark[19]).

The finding of CF-ICA also increases the risk. That CF-ICA-positive nondiabetic subjects may later develop type I diabetes has been reported in a few instances, but many of these were adults with other autoimmune endocrine conditions,[25,26] whereas the remainder were single cases,[27] including one of six CF-ICA-positive siblings in a smaller family study. In the present prospective study, as already reported,[6] each of the six who later developed diabetes was already positive for conventional ICA when first tested, so that 11% of the 54 conventional-ICA-positive nondiabetic subjects have become diabetic compared with 0% of those who remained ICA-negative. This difference is highly significant ($2P = 10^{-6}$, Fisher's exact test). The equivalent figures for CF-ICA, which was detected in 18 initially nondiabetic subjects, are 28% and 0.2%, respectively ($2P < 10^{-6}$). When these data are expressed in risk terms, the Woolf-Haldane method gives an RR of 73 (95% confidence limits 11–483) for development of diabetes in the initially healthy parents and siblings who were CF-ICA positive, compared with those who remained CF-ICA negative throughout the follow-up period. This RR should be interpreted with caution because new cases appearing while under review are as yet very few for this type of analysis, and because only a restricted period in the life of each participant has been surveyed. Nevertheless, it is clear that the combination of a positive CF-ICA result and HLA-identity to a diabetic sibling must give a child a very high risk of diabetes, even though the combined risk cannot as yet be quantified accurately.

In conclusion, we believe this to be the first study providing sufficient HLA and (prospective) ICA data for estimation of empirical risk of developing type I diabetes in different groups of siblings. HLA genotyping proves to be a powerful tool enabling the definition of one group of siblings having a very high risk of diabetes, and another with a negligible risk either of developing the disease or, presumably, of transmitting genetic susceptibility to future generations. This has considerable importance for genetic counseling.

ACKNOWLEDGMENTS

We are grateful to Drs. R. D. M. Scott, W. B. Thompson, D. Garrow, C. H. Cheetham, J. Scott, T. Westcott, and other physicians for allowing and helping us to study their patients; to Bridget Watson and Varina Drummond for meticulous HLA typing; to Dr. B. M. Dean and J. M. McNally for many thousands of islet-cell antibody tests; and not least to all the families who so readily agreed to take part.

Support for this project has been provided by the Medical Research Council, the British Diabetic Association, the Juvenile Diabetes Foundation, Miles Laboratories Ltd., and the Joint Research Board of St. Bartholomew's Hospital.

REFERENCES

[1] Simpson, N. E.: The genetics of diabetes mellitus—a review of family data. *In* The Genetics of Diabetes Mellitus. Creutzfeldt, W., Kobberling, J., and Neel, J. V., Eds. Springer, Berlin, 1976, pp. 1–11.

[2] Cudworth, A. G.: Type I diabetes. Diabetologia *14*:281–91, 1978.

[3] Gamble, D. R.: An epidemiological study of childhood diabetes affecting two or more siblings. Diabetologia *19*:341–44, 1980.

[4] Walker, A., and Cudworth, A. G.: Type I (insulin-dependent) diabetic multiplex families: mode of genetic transmission. Diabetes *29*:1036–39, 1980.

[5] Lendrum, R., Walker, G., Cudworth, A. G , Theophanides, C., Pyke, D. A., Bloom, A., and Gamble, D. R.: Islet cell antibody in diabetes mellitus. Lancet *2*:1273–76, 1976.

[6] Gorsuch, A. N., Spencer, K. M., Lister, J., McNally, J. M., Dean, B. M., Bottazzo, G. F., and Cudworth, A. G.: Evidence for a long pre-diabetic period in Type I (insulin dependent) diabetes mellitus. Lancet *2*:1363–65, 1981.

[7] Bottazzo, G. F., Dean, B. M., Gorsuch, A. N., and Cudworth, A. G.: Complement fixing islet cell antibodies in Type I diabetes. Possible monitors of active B-cell damage. Lancet *1*:668–72, 1980.

[8] Welch, S. G., and Boucher, B. J.: A rapid micro-scale method for the measurement of $HbA_{1(a+b+c)}$. Diabetologia *14*:209–11, 1978.

[9] WHO Expert Committee on Diabetes Mellitus: Second report. WHO Technical Report Series No. 646, 1980.

[10] Everitt, B. S.: The analysis of contingency tables. London, Chapman & Hall, 1977.

[11] Woolf, B.: On estimating the relation between blood group and disease. Ann. Hum. Genet. *19*:251–53, 1955.

[12] Haldane, J. B. S.: The estimation and significance of the logarithm of a ratio of frequencies. Ann. Hum. Genet. *20*:309–11, 1955.

[13] Beardmore, M., and Reid, J. J. A.: Diabetic children. Br. Med. J. *2*:1383–84, 1966.

[14] Christau, B., Kromann, H., Andersen, O. O., et al.: Incidence, seasonal and geographical patterns of juvenile onset insulin dependent diabetes mellitus in Denmark. Diabetologia *13*:281–84, 1977.

[15] West, R., Belmonte, M. M., Crepeau, M. P., Wilkins, J., and Poirier, R.: Epidemiologic survey of juvenile-onset diabetes in Montreal. Diabetes *28*:690–93, 1979.

[16] Durruty, P., Ruiz, F., and de los Rios, M.: Age at diagnosis and seasonal variations in the onset of insulin dependent diabetes in Chile (Southern Hemisphere). Diabetologia *18*:29–34, 1980.

[17] LaPorte, R. E., Fishbein, H. A., Drash, A. L., et al.: The Pittsburgh insulin-dependent diabetes mellitus (IDDM) registry: the incidence of insulin-dependent diabetes mellitus in Allegheny County, Pennsylvania (1965–1976). Diabetes *30*:279–84, 1981.

[18] Falconer, D. S., Duncan, L. J. P., and Smith, C.: A statistical and genetical study of diabetes: I. Prevalence and morbidity. Ann Hum. Genet. *34*:347–69, 1971.

[19] Green, A., Hauge, M., Holm, N. V., and Rasch, L. L.: Epidemiological studies of diabetes in Denmark: II. A prevalence study based on insulin prescriptions. Diabetologia *20*:468–70, 1981.

[20] Thomson, G., and Bodmer, W. F.: The genetic analysis of HLA and disease associations. *In* HLA and Disease. Dausset, J., and Svejgaard, A., Eds. Munksgaard, Copenhagen, 1977.

[21] Platz, P., Jakobsen, B. K., Morling, N., et al.: HLA-D and -DR antigens in genetic analysis of insulin-dependent diabetes mellitus. Diabetologia *21*:108–15, 1981.

[22] Gorsuch, A. N., Spencer, K. M., Wolf, E., and Cudworth, A. G.: Insulin dependent diabetes mellitus (IDDM): HLA and family studies. *In* The Genetics of Diabetes Mellitus—Second International Congress. Kobberling, J., and Tattersall, R., Eds. In press.

[23] Rubinstein, P., Ginsberg-Fellner, F., and Falk, C.: Genetics of Type I diabetes mellitus: a single recessive predisposition gene mapping between HLA-B and GLO. Am. J. Hum. Genet. *33*:865–82, 1981.

[24] Spielman, R. S., Baker, L., and Zmijewski, C. M.: Gene dosage and susceptibility to insulin-dependent diabetes. Ann. Hum. Genet. *44*:135–50, 1980.

[25] Bottazzo, G. F., Florin-Christensen, A., and Doniach, D.: Islet cell antibodies in diabetes mellitus with autoimmune polyendocrine deficiencies. Lancet *2*:1271–82, 1874.

[26] Irvine, W. J., Gray, R. S., and McCallum, C. J.: Pancreatic-islet cell antibody as a marker for asymptomatic and latent diabetes and prediabetes. Lancet *2*:1097–1102, 1976.

[27] Ilonen, J., Mustonen, A., Åkerblom, H., and Huttunen, N.-P.: Complement-fixing islet cell antibodies before and after onset of insulin-dependent diabetes. Lancet *2*:805, 1980.

[28] Ginsberg-Fellner, F., Dobersen, M. J., Witt, M. E., Notkins, A. L., Rubinstein, P., and Rayfield, E. J.: HLA antigens, islet cell antibodies and carbohydrate metabolism in siblings of children with insulin-dependent diabetes mellitus. Diabetes *28*:396A, 1979.

HLA AND FAMILY STUDIES

A. N. Gorsuch, K. M. Spencer, E. Wolf and A. G. Cudworth

Department of Medicine, St Bartholomew's Hospital, West Smithfield, London, UK

INTRODUCTION

Investigation of HLA-genotyped families is making an important contribution to research into the aetiology of Type 1 diabetes. The earliest work concerned pairs of diabetic siblings, and linkage studies make up the first part of this chapter. Findings from these reinforce the evidence from population studies suggesting the existence, in the HLA region of chromosome 6, of genes determining the major part of susceptibility to the disease. The second topic examined here is the risk of diabetes in relation to HLA haplotype inheritance in the siblings of affected children, complementing the empirical risk figures of Dr Green (this volume, p. 21). Finally we mention briefly some ways in which family studies are helping to elucidate other problems, dealt with elsewhere in this volume.

HLA AND LINKAGE STUDIES

Background

By HLA typing a whole family, it is usually possible to work out which alleles at the various loci occur together on the same chromosome as a haplotype. There are, of course, four haplotypes in each nuclear family, examples being labelled *a* or *d* in Fig. 1. Any child has in theory an equal chance of inheriting either *a* or *b* from the mother.

Serono Symposium No. 47, "The Genetics of Diabetes Mellitus", edited by J. Köbberling and R. Tattersall, 1982. Academic Press, London and New York.

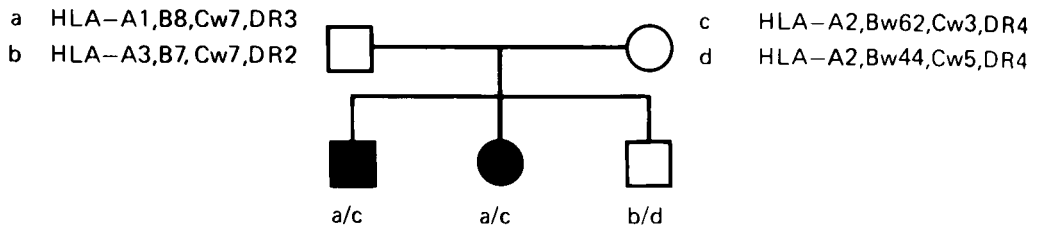


Fig. 1. HLA genotypes in a hypothetical family with two diabetic children. □ = male; ○ = female; ■ = diabetic. *a*, *b*, *c* and *d* are arbitrary symbols for the four HLA haplotypes in this particular family.

Assuming *no* linkage between diabetes and HLA, 25% of pairs of *diabetic* siblings would be expected to be *HLA-identical*, sharing both paternal and maternal haplotypes (for example, the pair in Fig. 1 shares *a* and *c*), 50% would share only *one* haplotype, which is generally termed *haplo-identity*, and 25% would be HLA *non-identical.* There has been some confusion in the literature over this terminology, and it is important to be sure what individual authors mean.

In the presence of linkage, we would expect an excess of HLA-identical and perhaps of haplo-identical pairs. The exact proportions would depend on several variables: the closeness of linkage, the mode of inheritance of susceptibility, the population gene frequency, and related to this the penetrance of the genes. Thus an accurate, empirical determination of the proportions of diabetic sib-pairs falling into the three categories of HLA haplotype identity would be important not only in demonstrating the existence of linkage, but perhaps also in helping to quantify these four variables.

Individual Linkage Studies

The first of these related to 17 families (Cudworth and Woodrow, 1975). Of 15 diabetic sib-pairs without further diabetic siblings, 10 (67%) were HLA identical and only one pair was non-identical. Each of the seven diabetic children in the other two families shared at least one haplotype with one or more affected siblings. These early findings thus confirmed the existence of linkage ($P = 0.001$).

Since 1977 there have been several similar studies (see below). The largest is the UK study (Walker and Cudworth, 1980) which was based on the Paediatric Register of the British Diabetic Association, to which physicians are invited to notify all new cases of diabetes occurring before the age of 16. Between 1972 and the end of 1978 there were 6600 notifications. From completed questionnaires relating to 4868 (74%) of these children, 175 families were found to have two affected children; eight further families had three and one had four diabetic children. An interesting observation is that in a significant excess (23%) of pairs, the interval between the development of diabetes in the first affected child and that in the second was less than one year, suggesting a role for an environmental factor determining onset of clinical disease (Gamble, 1980). Many of these families were HLA genotyped. Figure 2 incorporates results from these and some other Type 1 diabetic multiplex sibships typed in our laboratory. Of 135 pairs, 56% are HLA-identical, 38% are haplo-identical and 6% are non-identical. Fifteen pairs are from sibships with three or four affected children, the pair selected in each case being the first two children diagnosed as diabetic (see below).

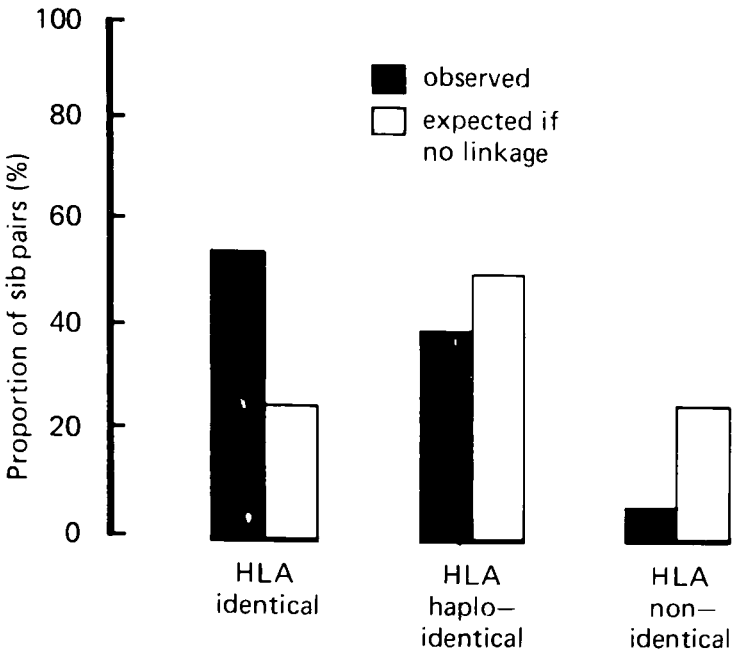


Fig. 2. **HLA haplotype concordance in 135 pairs of Type 1 diabetic siblings. Based on the UK Multiplex Family Study (see text).**

Linkage Studies Worldwide: Combined Data

In view of current interest in the mode of inheritance of susceptibility and the investigation of mathematical models based on various hypotheses, it seems appropriate to gather all available data from diabetic sib-pair studies. This is not as straightforward as it might at first appear. The principle problem is what to do with sibships with three or four diabetics. Some authors count all possible pairings, so that a sibship of four is represented no less than seven times – a dubious practice when assessing statistical significance of results from a small sample. If the aim is simply to establish linkage, the shared-haplotype scoring method of Green and Woodrow (1977) is preferable, but that is not the object of the present exercise. The method used here is to select one pair from each sibship, either at random or, where sufficient information is available, by taking the first two siblings diagnosed as diabetic. Multiple occurrence of susceptibility genes is theoretically more likely in these "triplex" (and "tetraplex") than in duplex families, and including them in mathematical models of inheritance mechanisms may unnecessarily increase complexity (R. Spielman, Personal Communication). Results from the larger sibships are therefore presented separately from those on the duplex sibships. This also facilitates comparison between the two types of families.

Sources

The sources of the sib pairs we have analysed are shown in Table I. We have excluded pairs where there is doubt about Caucasoid origin, uncertainty over haplotype sharing due to known or possible HLA homozygosity in a parent, inconsistency between text and tabulated results, or where diabetes does not fit the Type 1 pattern. The Danish contribution to the international Workshop figures has been excluded to avoid overlap with Platz *et al.* (1981), and so far as we can tell no family is included more than once. About half of the 376 pairs come from

Table I. Sources of combined data on Type 1 diabetic sib-pairs.

Location	Report		No. of pairs[a]	D/DR typing
International	Svejgaard *et al.*	(1980)	86[b]	some
Europe				
Austria	Ludwig *et al.*	(1977)	3	no
Denmark	Platz *et al.*	(1981)	51	yes
France	Deschamps *et al.*	(1980)	12	yes
West Germany	Bengsch *et al.*	(1978)	15	no
UK	Walker and Cudworth plus new data	(1980)	135	no
	Winearls and Bodmer	(1980)	2	?
USA				
Minnesota	Barbosa *et al.*	(1977, 1980a)	28	some
Maryland	Hsu *et al.*	(1977)	2	?
New York	Suciu-Foca *et al.*	(1979)	11	yes
Pennsylvania	Spielman *et al.*	(1980)	31	some

[a]The sib-pairs included are Type 1 diabetics of Caucasoid descent showing unambiguous HLA haplotypes; [b]Excluding the Danish families.

studies where HLA-D or -DR typing was performed, but the actual number of pairs successfully typed for these specificities is uncertain. A few authors have included an occasional diabetic sib-pair with an intra-HLA recombinant, classifying them by identity of the D/DR part of the affected haplotype, or where there is no informative D or DR result by using B-locus identity. Unfortunately, from the published literature it is impossible to state how many A-B and B-D recombinants there are in the combined data. In the report of Suciu-Foca *et al.* (1979) there appears to be some internal inconsistency, but it seems that four duplex and one triplex sibships out of a total of 17 multiplex families include a diabetic recombinant. No other group has found such a high recombination rate, and to avoid bias we have excluded from the present analysis the sib-pairs from these five families; in all cases they were HLA-D/DR identical.

Results

Table II shows the combined data on HLA haplotype sharing by pairs of Type 1 diabetic siblings of Caucasoid descent. There is no significant heterogeneity between duplex pairs (with no further affected siblings) and pairs from sibships with three or four diabetic children. The proportions of pairs in the three haplotype identity groups are virtually identical to those in the UK study (Fig. 2), and there is no evidence for heterogeneity between the individual reports contributing more than 10 pairs (χ^2_{14} for heterogeneity = 13.1). The overall deviation from the "expected" distribution, assuming no linkage is, of course, very highly significant ($P \ll 10^{-10}$).

Table II. HLA haplotype concordance in pairs of Type 1 diabetic siblings: combined data, 1981.

Number of diabetic sibs in family	Number of pairs			
	HLA-identical	Haplo-identical	Non-identical	Total
2	182 (55%)	128 (39%)	18 (5%)	328
3 or 4[a]	31 (65%)	13 (27%)	4 (8%)	48 [c]
2, 3 or 4:				
Observed	213 (57%)	141 (38%)	22 (6%)	376
Expected[b]	94	188	94	376 [d]

[a]One pair selected from each sibship, either at random or as first two diagnosed diabetic; [b]"Expected" = numbers expected if *no* linkage between HLA and diabetes; [c]No significant heterogeneity: $\chi^2_2 = 2.78$, $P = 0.25$; [d]Significance of difference: $\chi^2_2 = 217.5$; $P \ll 10^{-10}$.

HLA AND FAMILY STUDIES NOT INVESTIGATING LINKAGE

From the findings described above, it follows that a non-diabetic sibling of a Type 1 diabetic child is himself likely to be genetically susceptible if HLA-identical to the diabetic child, whereas haplo- and non-identical siblings are progressively less at risk. This argument does not depend on the presence or absence of the "high-risk" antigens (e.g. HLA-DR3 or -DR4) defined in population studies. Research involving HLA-genotyped families, each including a diabetic proband with one or more siblings, is already helping in the search for answers to a number of questions. Some of these are considered below, with particular reference to our own prospective family study.

The Barts–Windsor Family Study

This project was set up in 1978 to investigate immunological and environmental events before the onset of any future diabetes in first degree relatives of Type 1 diabetic probands. The families taking part live in a defined geographical area of about 1000 km^2 centred on Windsor, England. They were ascertained through local diabetic clinics using the following criteria: Type 1 diabetes in the proband; availability of at least one sibling aged under 20 years; and availability of parents, although in some larger families it was possible to deduce the HLA genotype of a missing parent. Ascertainment was independent of diabetes in the siblings, and the ascertainment rate is believed to be high. Further details are given elsewhere (Gorsuch *et al.*, 1980).

Results from 160 families are analysed here. Table III gives the numbers and age ranges of family members typed for HLA-A, B and C antigens. The ascertainment rate of siblings in these families was 95%, and none of the missing 10 sisters

Table III. The Barts–Windsor Family Study: subjects ascertained in the 160 families included.

Subjects	No.	Male	Female	Age at entry
Probands[a]	160	86	74	14 (4–29)
Siblings	288	164	134	14 (1–32)
Parents	313	154	159	42 (26–61)

[a]Age of diagnosis of diabetes = 8 (1–19) years.

and five brothers is diabetic. In 10 families (with 16 siblings) the paternal or maternal HLA haplotypes could not be distinguished unequivocally, even after HLA-DR typing.

Distribution of Siblings According to HLA Identity to Proband

In the Barts–Windsor study, 50 (18%) of the 272 siblings with unambiguous HLA genotypes are HLA-identical to their diabetic probands, 151 (56%) are haplo-identical and 71 (26%) are non-identical, including both diabetic and healthy siblings. The apparent deviation from the "expected" proportions (25%, 50% and 25% respectively) is not quite significant at the 5% level, but a deficit of HLA-identical pairs is anticipated in view of the method of ascertainment and linkage between HLA and diabetes. The actual proportions expected would depend on a number of other factors such as family size, mode of inheritance of susceptibility and penetrance.

To illustrate how this deficit in HLA-identical siblings might occur, let us consider a hypothetical example, simplifying the problem by making a number of assumptions. We assume random zygotic assortment of HLA haplotypes, and that survival to the age of ascertainment is independent of both HLA and susceptibility genes. We assume further that susceptibility genes from *both* parents are necessary for susceptibility, that penetrance is then (for the sake of argument) 100% and that the genes are tightly linked to HLA.

Consider a population of 400 families with each father and each mother carrying susceptibility gene(s) on one HLA chromosome only, and with three children in each family. Of the total 1200 children, 300 (25%) would be expected to be susceptible, having inherited genes from both parents. The expected distribution of these 300 children among the 400 families can be calculated by simple probability. Thus, the probability that any child is *not* susceptible is 0.75 (i.e. 1–0.25), and the probability that *none* of the three children in a given family is susceptible is 0.75^3. Hence the expected number of families with at least one susceptible child is: $(1 - 0.75^3) \times 4000 = 231$.

Now, as penetrance is here taken as 100% all these families could be ascertained in a family study through a diabetic child. There would then be 231 diabetic probands with 462 siblings. The expected number of *susceptible* siblings in that study would then be: 300 231 = 69. Therefore, with tight HLA-linkage, the proportions of siblings expected to be HLA-identical to their probands would be: $69 \div 462 = 0.15$ or 15%.

Of course, this is only an illustration, and the answer would have been modified had our assumptions been different. Nevertheless, there seems no need to invoke "negative selection" before birth against zygotes carrying susceptibility genes on both chromosomes – a possibility suggested as an explanation for the "disturbance" of HLA haplotype distribution among non-diabetic children in multiplex families (Bengsch *et al.*, 1978). Perhaps a more rigorous treatment, taking into account the variables, given assumed values in our example, might help in the formulation and testing of mathematical models of inheritance.

Risk in HLA-genotyped Siblings of Diabetic Children

Observed Prevalence

At the time of writing, 14 (5%) of the 288 siblings and nine (3%) of the parents in the Barts-Windsor study are diabetic. The mean age of the siblings is 16 years. Table IV shows the prevalence in each group of siblings, defined according to HLA identity to their probands.

Table IV. The Barts-Windsor Family Study: prevalence of Type 1 diabetes in siblings grouped according to HLA haplotype identity to probands.

	Number of haplotypes shared with proband					All siblings
	2	2/1[a]	1	1/0	0	
Total No.	50	8	151	8	71	288
No. diabetic	7	0	6	0	1	14
Prevalence	14%		4%		1%	5%

[a]Either two or one haplotype in common with proband. These siblings are from families where the two haplotypes of one of the parents are indistinguishable.

These prevalence figures are, of course, based on rather small numbers of diabetics. How reliable are they, and by how much is the risk of diabetes in a child increased when that child has a diabetic sibling? We have approached these questions in two ways: by making indirect estimates of the prevalence figures from independent data to see whether they agree and by comparing the prevalence in the general population.

Indirect Calculation of Prevalence

Here we use figures from the UK multiplex family study (Gamble, 1980; Walker and Cudworth, 1980); approximately 6% of siblings of diabetic children are themselves diabetic at the age of 16 years, and 56% of diabetic sib pairs are HLA-identical. If we assume that 20% of all siblings of diabetics are HLA-identical to them (see above), then the calculated cumulative frequency of diabetes in such HLA-identical siblings by age 16 would be about:

$$\frac{0.06 \times 0.56}{0.20} = 0.16 \text{ or } 16\%$$

For haplo-identical siblings the indirect method gives about 5%, and for non-identical 1.6%. These estimated prevalence figures are not strictly comparable to those actually observed (Table IV) since the ages of the siblings in the Barts-Windsor Family Study cover a wide range, although the mean age is indeed 16. Nevertheless, there is a marked similarity. Furthermore, the indirect method gives virtually identical estimates in a third population (Platz *et al.*, 1981).

Comparison with General Population Prevalence

There are no accurate figures for prevalence of Type 1 diabetes at different ages in the UK. The Paediatric Register of the British Diabetic Association suffers from patchy notification, and includes only those cases developing before age 16. In areas of highest ascertainment, the calculated frequency at age 16 is 0.22% (Gamble, 1980). The Danish incidence study, smaller but with high ascertainment, gives an equivalent figure of 0.26% (calculated from Christau *et al.*, 1977). For our comparison, however, we used a true prevalence study of insulin-dependent diabetes in 45 500 schoolchildren aged 5–15 years, in Northampton, England (Beardmore and Reid, 1966). We have applied an age correction factor (based on the distribution of age of onset of diabetes in Liverpool to predict that about 88 (0.19%) of these children would be diabetic when their mean age reached 16.

Table V gives relative risks (RR) for the occurrence of diabetes in siblings from the Barts-Windsor Family Study, grouped according to number of HLA haplotypes shared with the probands. Relative risk is more often encountered as a measure of association between two genetically determined factors in population studies, but as used here it can be thought of as the ratio between the risk of diabetes in the siblings and that in children in the general population. It is used here instead of a simple prevalence ratio to show the 95% confidence limits, since RR and its standard error can be calculated from prevalence data and corrected for bias and discontinuity (Haldane, 1955; Woolf, 1955). The RR for *all* siblings, irrespective of HLA identity to probands, is 27 – virtually identical to the prevalence ratio calculated by Gamble (1980). In the HLA non-identical group, only one sibling is diabetic, and the binomial test shows that this frequency is not significantly greater than that in the general population ($P = 0.12$). The most striking results in Table V are the high RR in the HLA-identical group, and the difference between this value and that in the haplo-identical group.

Table V. Estimated increased risk of developing Type 1 diabetes according to HLA-linked susceptibility based on a comparison between the Barts-Windsor Family Study and the Northampton survey of Beardmore and Reid (1966), see text.

	HLA-identical siblings	Haplo-identical siblings	HLA-non-identical siblings	All siblings
Relative risk	88	23	11	27
95% confidence limits	40–197	10–52	2–56	15–48

It might be argued that the prevalence of childhood diabetes in the UK might have increased since the Northampton study of Beardmore and Reid (1966), in which case our figures for RR would be overestimates. There is no good evidence for any such increase, and annual notifications to the Paediatric Register of the British Diabetic Association were more or less constant between 1973 and 1980 (Gamble, Personal Communication). However, we have for comparison repeated the calculations using the general population prevalence figures estimated from the Danish study (see above), to give RR values of 45 (95% confidence limits 20-97), 12 (5-27) and not significantly increased in the HLA-identical, haplo-identical and non-identical siblings respectively.

It seems safe to say that the risk of diabetes in a child is probably about 60- to 90-fold greater when that child is HLA-identical to a diabetic sibling than when he has no diabetic sibling.

Further Questions Approached Through Family Studies

A recurring theme at this congress concerns the nature of genetic susceptibility to diabetes. How do the genes act? Can the gene products be identified? In Type 1 diabetes, the "family study" approach is likely to prove a powerful research tool in attempting to answer these questions.

By investigating HLA-genotyped families of affected probands, comparisons may be made between "potentially susceptible" and "non-susceptible" healthy subjects, judged on the basis of HLA haplotype identity to probands. For example, it appears that the increased prevalence of organ-specific auto-antibodies in siblings in such studies is independent of HLA identity to the probands, which does not favour the suggestion that HLA-linked susceptibility to diabetes involves generalized enhancement of humoral organ-specific autoimmunity (Gorsuch *et al.*, 1980).

Prospective follow-up of such families adds a further dimension. In this context, Dr Bottazzo (this volume, p. 79) discusses islet-cell antibody as a marker to identify future diabetics early in the prediabetic phase. A different abnormality has been suggested as a possible marker in this respect by Barbosa *et al.* (1980[b]). Using these or similar techniques it becomes feasible to investigate possible pathogenetic factors (such as autoimmunity and viral infections) (Cudworth and Gorsuch, 1981; Gorsuch and Cudworth, 1981) at the time they are active, without having to rely on retrospective investigations after the onset of clinical diabetes.

CONCLUSIONS

As we have seen, research involving HLA-geneotyped families contributes at several levels to our understanding of the aetiology and pathogenesis of Type 1 diabetes. Diabetic sib-pair studies confirm linkage between HLA and the disease, and the quantitative data obtained provide fodder for arguments on the various possible modes of inheritance of susceptibility. The assessment of risk in HLA-genotyped siblings of affected children, although as yet rather tentative, has important implications for genetic counselling. Further studies of this type, taking into account HLA-DR specificities and other "risk factors" such as islet-cell antibodies, would be feasible only with larger series of families.

Perhaps the most exciting prospects appear where the disciplines of genetics, immunology, virology and biochemistry are co-ordinated in major, long-term prospective studies of HLA-genotyped families. In our opinion, it is this approach which offers the greatest hope of working out the exact details of aetiology and pathogenesis, with the eventual possibility of prevention of the disease.

ACKNOWLEDGEMENTS

We would like to thank Dr J. Lister and the many other physicians for their help and support in allowing us to study their patients, and Dr A. Bloom for the use of the British Diabetic Association Paediatric Register. Dr Angela Walker was responsible for visiting the BDA multiplex families.

Invaluable technical assistance was provided by Bridget Watson, Varina Drummond and Vivienne Algar in relation to the HLA studies. We are grateful to Dr B. M. Dean and Dr G. F. Bottazzo for the auto-antibody estimations and to Miss P. Matharu for secretarial assistance.

This work has been supported by the Medical Research Council, the British Diabetic Association, the Joint Research Board of St Bartholomew's Hospital and the Juvenile Diabetes Foundation.

REFERENCES

Barbosa, J., King, R., Noreen, H. and Yunis, E. (1977). *J. clin. Invest.* **60**, 989–998.

Barbosa, J., Chern, M. M., Anderson, V. E., Noreen, H., Johnson, S., Reinsmoen, N., McCarty, R., King, R. and Greenberg, L. (1980a). *J. clin. Invest.* **65**, 592–601.

Barbosa, J., Cohen, R. A., Chavers, B., Michael, A. F., Steffes, M., Hoogwerf, B., Szalapsi, E. and Mauer, M. (1980b). *Lancet* **ii**, 330–333.

Beardmore, M. and Reid, J. J. A. (1966). *Br. med. J.* **ii**, 1383–1384.

Bengsch, N., Köbberling, J., Eckert, G. and Willms, B. (1978). *Diabetologia* **15**, 447–451.

Christau, B., Kromann, H., Andrews, O. O., Christy, M., Buschard, K., Arnung, K., Kristensen, I. H., Peitersen, B., Steinrud, J. and Nerup, J. (1977). *Diabetologia* **13**, 281–284.

Cudworth, A. G. and Gorsuch, A. N. (1981). *In* "Diabetes Mellitus: Theory and Practice" (M. Ellenberg and H. Rifkin, Eds.), 3rd edition.

Cudworth, A. G. and Woodrow, J. C. (1975). *Br. med. J.* **iii**, 133–135.

Deschamps, I., Lestradet, H., Bonaiti, C., Schmid, M., Busson, M., Benajam, A., Marcelli-Barge, A. and Hors, J. (1980). *Diabetologia* **19**, 189–193.

Gamble, D. R. (1980). *Diabetologia* **19**, 341–344.

Gorsuch, A. N. and Cudworth, A. G. (1982). *In* "Paediatric Immunology" (J. F. Soothill, H. Greenwood, A. R. Hayward and B. B. S. Wood, Eds.), Blackwell, Oxford (In press).

Gorsuch, A. N., Dean, B. M., Bottazzo, G. F., Lister, J. and Cudworth, A. G. (1980). *Br. med. J.* **i**, 145–147.

Green, J. R. and Woodrow, J. C. (1977). *Tissue Antigens* **9**, 31–35.

Haldane, J. B. S. (1955). *Ann. hum. Genet.* **20**, 309–311.

Hsu, S. H., Hsu, T. H., Chase, G. A. and Bias, W. B. (1977). *Transplantation Proceedings* **9**, 1855–1858.

Ludwig, H., Schernthaner, G. and Mayr, W. R. (1977). *Diabète et Metabolisme* **3**, 43–48.

Platz, P., Jakobsen, B. K., Morling, N., Ryder, L. P., Svejgaard, A., Thomsen, M., Christy, M., Kromann, H., Benn, J., Nerup, J., Green, A. and Hauge, M. (1981). *Diabetologia* **21**, 108–115.

Spielman, R. S., Baker, L. and Zmijewski, C. M. (1980). *Ann. hum. Genet.* **44**, 135–150.

Suciu-Foca, N., Rubinstein, P., Nicholson, J., Susinno, E., Weiner, J., Godfrey, M., Hardy, M., Rayfield, E. and Reemtsma, K. (1979). *Transplantation Proceedings* **9**, 1309–1313.

Svejgaard, A., Platz., P. and Ryder, L. P. (1980). *In* "Histocompatibility Testing 1980" (P. I. Terasaki, Ed.). UCLA Tissue Typing Laboratory, Los Angeles.

Thomson, G. (1980). *Ann. hum. Genet.* **43**, 383–398.

Walker, A. and Cudworth, A. G. (1980). *Diabetes* **29**, 1036–1039.

Winearls, J. and Bodmer, J. G. (1980). Cited as personal communication by Thomson (1980).

Woolf, B. (1955). *Ann. hum. Genet.* **19**, 251–253.