Enacting Molecular Complexity: Data and Health in the Metabonomics Laboratory

Nadine Sarah Levin
Green Templeton College
Trinity Term 2013

A dissertation submitted in fulfilment of the requirements for the degree of Doctor of Philosophy in Social Anthropology.

In this dissertation, I examine how biological data practices enable researchers to interact with and enact biological life in statistical ways, and how this poses challenges to the use and integration of biological knowledge with clinical practices. Instead of considering data as a pre-existing cognitive representation of the world, I combine scholarship on the anthropology of science with scholarship from science and technology studies to consider data as a form of *material practice*. I consider, in other words, how data is intertwined with technologies, people, and values, such that data is used to make normative and naturalized claims about biology and disease.

To explore the generation, interpretation, and use of biological data, I focus on the field of "metabonomics"—the post-genomic study of metabolism—as it is carried out within the Biomolecular Medicine Laboratory (BMM) at Imperial College London. In doing so, I examine how metabonomics researchers use biochemical techniques and multivariate statistics to investigate metabolism and disease.

After providing an overview of the literature, central questions, and methodology that frame this dissertation, I examine how multivariate statistical practices are central to the historical identity and epistemic culture of metabonomics research at the BMM. From there, I demonstrate how multivariate statistics require and enable metabonomics to enact metabolism as an inherently complex entity. Consequently, I examine how researchers struggle to assign the categories of "normal" and "abnormal" to dynamic notions of metabolism and health. I then explore how the translation of metabonomics knowledge into clinical practices places value on multivariate forms and large volumes of information, eclipsing the importance of human interpretation and judgment. Finally, I examine how metabonomics research is used to develop personalized medicine, but in ways that make it difficult to address the health of individual patients.

Enacting Molecular Complexity: Data and Health in the Metabonomics Laboratory

Nadine Sarah Levin

Green Templeton College

Institute of Social & Cultural Anthropology

A dissertation submitted in fulfilment of the requirements for the degree of

Doctor of Philosophy in Social Anthropology

Trinity Term 2013

Acknowledgements

First and foremost, I want to thank my supervisors Stanley Ulijaszek (Oxford) and Simon Cohn (Cambridge) for their continuing and unwavering support for this project. Stanley – thank you for providing clarity on scientific and statistical ideas, and for keeping me honest with the scope and specificity of my claims. Simon – thank you for agreeing to supervise my project in the first place when you had no obligation. I am incredibly grateful to all of the hours spent chatting over coffee in the British Library, and—above all—for your confidence in my ideas and future aspirations as an anthropologist. Stanley and Simon, together your insights and attentiveness has made this project what it is today.

I also owe a great amount of thanks to the Biomolecular Medicine (now called Computational and Systems Biology) Laboratory at Imperial College London. Without their support, and their willingness to have an anthropologist in their midst, this project would not have been possible. In particular, I want to thank Professor Jeremy Nicholson for his generosity in allowing me to conduct fieldwork, and for providing me with financial support. I also want to thank Claire Merrifield for introducing me to the laboratory, and for working closely with me in the first few months of my fieldwork. And lastly, I want to thank not only the numerous informants who participated in this study, but also the number of good friends and co-workers—Silke Heinzmanm, Alma Villasenor, Steve Robinette, Mark McPhail, Judy Fonville, Hector Keun, Beatriz Jimenez—I came and continue to have.

Within the Institute of Social and Cultural Anthropology at Oxford, I want to first thank Elisabeth Hsu and Javier Lezaun for providing support and feedback during my Confirmation of Status examination. Your comments were crucial for helping me to re-evaluate the focus and trajectory of my thesis. I also want to thank Dr Caroline Potter for her support during my MPhil, and for encouraging me to consider extending my studies into a DPhil.

I also want to thank my peers and colleagues for their help, support, and distraction throughout my five years at Oxford: Nick, Jonah, and Nicole, who started with me on the MSc in Medical Anthropology; Tanja Schneider, who introduced me to researchers at InSIS and provided helpful feedback; and lastly Amy Hinterberger, who gave me integral feedback and support. I also want to thank Beth Morrissey, Susan Portalupi, Sarah-Ann Burger, and Charlotte Greenhalgh for helping me with my final push in editing.

I have been incredibly fortunate to have the additional feedback and support of several wonderful academics—Hannah Landecker, Kaushik Sunder Rajan, and Sabina Leonelli—outside of Oxford University. Hannah – thank you for your inspiring thoughts on

metabolism, and I hope that we have the opportunity to work together further in the near future. Kaushik – thank you for challenging me to develop my ideas on the life sciences and anthropological writing, and for providing a truly inspiring example for how anthropologists can think about and influence the life sciences. Sabina – thank you for your insights into data practices beyond the world of metabonomics, and for your support in my early (and future) career. I would also like to say a special thanks to Don Chambers, for helping me to question where best to place my abilities and to pursue a future career.

Financially, this project would not have been possible without the support of multiple institutions and funding bodies. I am especially grateful to the Rhodes Trust for enabling me to study at Oxford University in the first place. I am also grateful to Green Templeton College, the Institute of Social and Cultural Anthropology, the Biosocial Society, the PEO Sisterhood, and the Oxford University Vice Chancellor's Fund.

Finally, I would like to thank those people who are not within my field, but who have been the closest to my heart and work throughout these five years. Andrew – thank you for being an unwavering source of love and stability, and for always encouraging me to have confidence in myself. Mom and Dad, how can I say enough – thank you for helping me foster a love of science, and for supporting me in every way possible not only in Oxford but also throughout my whole upbringing.

Table of Contents

Figures

Acronyms

A&E	Accident and Emergency, otherwise known as Emergency Room
BMM	Biomolecular Medicine Laboratory
BRC	Biomedical Research Centre (UK)
CAP	Community acquired pneumonia
COMET	Consortium for Metabolic Toxicology
FDA	Food and Drug Administration (USA)
GWAS	Genome-wide association study
HGP	Human Genome Project
INTERMAP	International Collaborative Study of Macronutrients, Micronutrients and Blood Pressure
ITU	Intensive care unit
MALDI-MSI	Matrix-assisted laser desorption/ionization mass spectroscopy imaging
MAS-NMR	Magic angle spinning nuclear magnetic resonance
MELD	Model for end-stage liver disease
MHz	Megahertz
MRC	Medical Research Council (UK)
MRI	Magnetic resonance imaging
MS	Mass spectrometry
MWAS	Metabolome wide association study
NIH	National Institute of Health (USA)
NIHR	National Institute for Health Research (UK)
NHS	National Health Service (UK)
NMR	Nuclear magnetic resonance
PCA	Principal component analysis
PLS	Partial least squares
SAF	Sir Alexander Fleming
SOM	Self-organizing map
STS	Science and technologies studies

“All models are wrong, but some are useful.”

George E.P. Box, Statistician

Introduction

It is the middle of August in the summer of 2012, and the London Olympic Games are in full swing. Despite predictions of widespread chaos, the Games have gone smoothly, and with few issues for spectators or athletes. Back in Oxford, as I write my dissertation with the BBC's impressive all-day coverage of the events playing through headphones, I am googling the laboratory in which I did my fieldwork—the Biomolecular Medicine Laboratory (BMM) at Imperial College London—for background information.

I want to know what the latest research publications are for the laboratory's field of research, which is termed "metabonomics" and forms the focus of this dissertation (see Figure 1). Metabonomics examines the molecular components that make up the metabolism of organisms. It seeks to define the "metabolome"—the sum of the metabolites within an organism—and to explore how the metabolome changes in relation to disease, diet, and environment. To do so, metabonomics uses biochemical and statistical technologies to analyze the molecular composition of urine, blood, and tissue samples. As such, it seeks to understand the relationships among the various components—the genes, proteins, cells, tissues, and organ systems—that join together to make metabolism an interconnected, dynamic, and complex process. Metabonomics, therefore, is the post-genomic[1] study of metabolism: similar to other post-genomic fields like epigenetics and proteomics, it follows in the wake of the Human Genome Project (HGP) of the 1990s to examine the downstream effects of genes (Natural Environment Research Council 2006).

[1] The term "post-genomic" refers to the study of the effects and mechanisms of gene regulation and expression. Thus, post-genomic fields seek not to identify the structure or sequences of genes, but rather to study how they have higher biological meanings and functions. The best known fields of post-genomic research include epigen(etics/omics), transcriptomics, metabo(lomics/nomics), and lipidomics, but the term also encompasses broader fields such as systems biology.

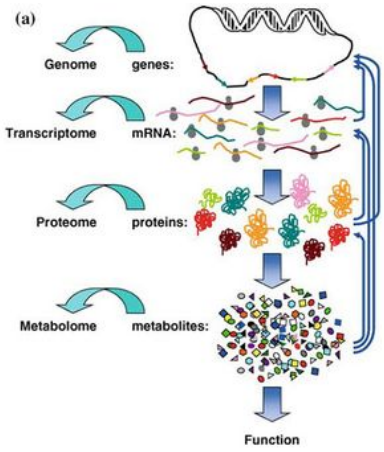


Figure 1: Metabonomics in relation to other fields of post-genomic research. Diagram of the fields of post-genomic research and their corresponding objects of investigation. The diagram is organized from top to bottom according to increasing levels of biological complexity, and shows how the field of metabonomics (pictured at the bottom) investigates the metabolome and its constituent metabolites (Goodacre 2005).

As one of the pioneering metabonomics laboratories, the BMM consists of more than 75 researchers conducting work in a diverse array of areas, which include nutrition, toxicology, surgery, cancer, epidemiology, infectious disease, and data analysis methods. The laboratory is spread out across facilities at Imperial College London's South Kensington campus, as well as several hospitals operated through the Imperial College Healthcare National Health Service (NHS) Trust. The BMM is led by Professor Jeremy Nicholson, its founder and the current Head of the 1000-person Department of Surgery and Cancer. The laboratory is an academic powerhouse: Jeremy Nicholson alone has more than 500 publications, while the BMM's research activities are funded by a variety of public and private initiatives within and beyond the United Kingdom.

Amidst my searching, up pops a link with an interesting title: "A phenomenal legacy for London 2012" (UK Department of Health 1 August 2012). In and of itself, the word "legacy" evokes notions of future expectations and promises, hinting at the fact that many biomedical projects capitalize on unrealized potentials to leverage funding and media attention (Brown et al. 2003). I click on the link, intrigued by potential connections between the global-scale sports competition and the local-scale laboratory research. Reading the article, I discover that Jeremy Nicholson has managed to secure a deal that will transform the Olympic laboratory used to test for illegal performance-enhancing drugs into a cutting edge metabonomics facility (Saini 2012).

Termed the "MRC-NIHR[2] Phenome Centre," the facility will examine people's "phenomes"—the biochemical summaries of their bodies in space and time—in order to provide information about disease risk factors and biomarkers. The MRC-NIHR Phenome Centre, I read, will house millions of pounds of laboratory equipment, and will be funded by a combination of public and private entities. It is scheduled to open sometime in 2013 at Hammersmith Hospital, one of the major Imperial College medical campuses (UK Department of Health 1 August 2012). The Centre builds on the BMM's efforts—as I discuss in Chapters 2 and 5—to extend the research activities of the field of metabonomics beyond the walls of the laboratory and to the domain of clinical work[3].

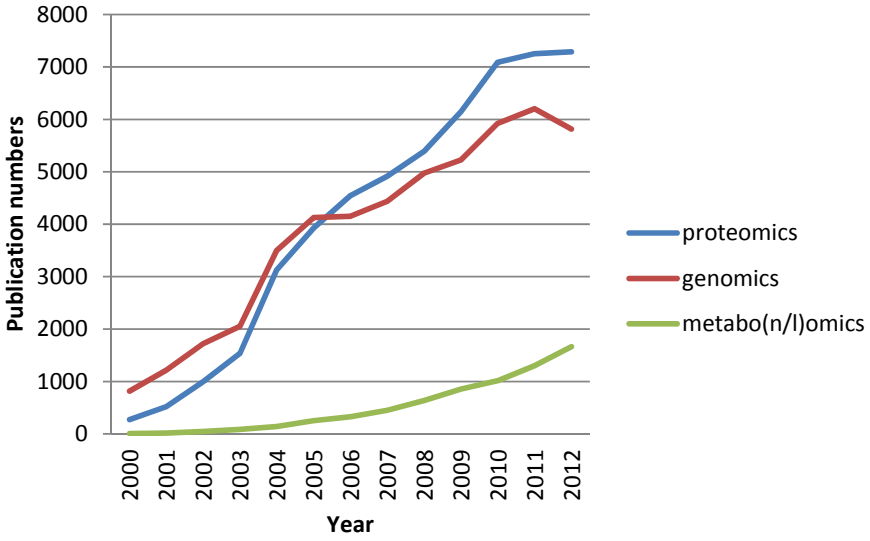


Figure 2: Metabonomics publication statistics.
Table showing the numbers of papers published in the fields of proteomics, genomics, and metabonomics/metabolomics. Data provided from the ISI Web of Knowledge, and generated according to searches in "Topic," "Articles," and "Science and Technology." Note the increasing trend for metabonomics/metabolomics publications.

[2] Abbreviation for Medical Research Council (MRC) and National Institute for Health Research (NIHR).

[3] In a published interview, Jeremy Nicholson describes the evolution of the project:

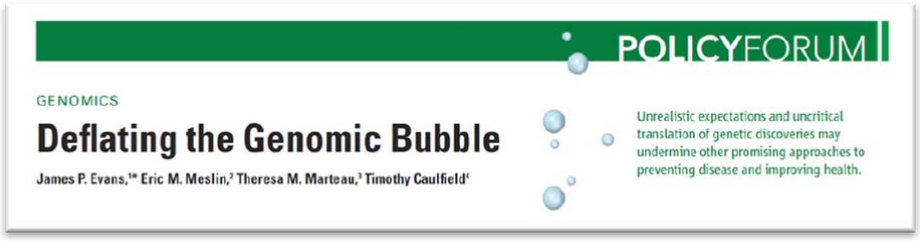
> I've been working in metabolic phenotyping and metabolic profiling for the best part of 30 years. I've been thinking about trying to build a national center for about seven or eight years to broaden and extend the research capacity and capability…to other universities in the U.K., even outside of the U.K. The infrastructure of the Olympic Games 2012 antidoping laboratory offered a window of opportunity. They had 45 mass spectrometers of various types working in parallel. They were doing up to 300 forensic assays for different drugs, metabolites, and other markers of abuse with a turnaround time of about six hours. There is no analytical laboratory in the world with that sort of capacity and throughput. (Mukhopadhyay 2013)

The "Phenome Centre" is a testament to the increasing size and strength of the field of metabonomics in UK and global academic institutions alike (see Figure 2), and follows a surge of developments at the interface between metabonomics and biomedicine. The National Institute of Health (NIH), for example, recently announced a program to invest more than $50 million USD in metabolomics research over five years between 2012 and 2017 (National Institute of Health 19 September 2012)[4]. Within the UK, numerous academic institutions carry out research in metabonomics, either as a sub-speciality within an academic department, or as a dedicated field that defines an institute or laboratory. On a broader scale, metabonomics research is carried out by a number of pharmaceutical and nutrition companies, including Nestlé[5], Unilever, AstraZeneca, Servier, and Pfizer. Together, the ties between metabonomics research and commercial and governmental developments signal—as I discuss throughout Chapter 2—that the field has evolved within and been influenced by particular historical and societal contexts.

Researchers within the BMM draw on the claim that metabonomics captures the complexity of life, in order to highlight why metabonomics provides a compelling alternative to genomics. Metabonomics, they claim, provides a real-time understanding of metabolism as the dynamic outcome of the interaction between genes, metabolic pathways, and the environment. Because of its ability to reflect gene-environment interactions, metabonomics can be used to understand and treat complex disorders of nutrition, immunity, cancer, and pathogens. As one metabonomics researcher claimed: "The genome tells us about the genetics…it gives you the alphabet, but doesn't give you the language. This is a way of linking from the genetics to the disease pathway." Thus, in criticizing genomics for only providing the blueprint of life (see Figure 3), metabonomics positions itself as the field that can move beyond the reductionism and static quality of genetics to provide a complex picture of metabolism and life (see Chapter 3).

[4] Similarly, the non-profit organization Genome Canada invested $7.5 million CAD in the development of the Human Metabolome Project in 2007, in an attempt to identify and catalogue the range of metabolites present in human beings (Genome Canada 2007).

[5] One page of the Nestle website asks "Metabonomics – What metabolic profile are you?" and claims that measuring individual metabolic phenotypes can aid in the tailoring of nutrition (Nestle 2013).


POLICYFORUM

GENOMICS

Deflating the Genomic Bubble

James P. Evans,[1]* Eric M. Meslin,[2] Theresa M. Marteau,[3] Timothy Caulfield[4]

Unrealistic expectations and uncritical translation of genetic discoveries may undermine other promising approaches to preventing disease and improving health.

Figure 3: Metabonomics - addressing the limits of genomics?
Headline article in the journal *Science* questioning the biological sciences' focus on genomics at the expense of other fields of research (Evans et al. 2011). This paper was recommended to me during my fieldwork, as an example highlighting the importance of metabonomics research amidst the popularity of genomics research.

Reading the article about the Olympic legacy, I remember back to the handful of lectures in which I watched Jeremy Nicholson talk excitedly about recent developments in metabonomics. The ability to track peoples' metabolic responses to organ transplants, the revolutionary studies in the molecular epidemiology of cardiovascular disease, the insights into the role of gut microbiota in complex conditions such as obesity or autism: all of these things were going to revolutionize biomedicine. In measuring the molecular basis of metabolism, metabonomics would allow—as I discuss in Chapter 4—biomedical diagnosis and treatment to become more objective and precise.

I first met Jeremy Nicholson at the beginning of August 2012, when I sent him an introductory email expressing my interest in conducting ethnographic research at the BMM. He wrote to me with a short and characterful email, saying:

> Sure you can come and chat- I am pretty busy until about 3rd week of August-But let's get one thing straight: "Genomics has nothing whatsoever to do with population obesity and very little to do with individual obesity- It's all down to environment, culture, lifestyle, bugs and bad habits- Discuss?"

After I crafted a response that was, apparently, to his liking, he invited me to the laboratory to give me a perfunctory tour. He was charismatic, and as expected of a high-functioning academic, extremely busy[6]. Speaking twice as fast as most human beings, he described the field of metabonomics in relation to technologies that have formed important aspects of biochemistry and toxicology research for the past several decades. Metabonomics, Jeremy

[6] See a featured interview in the magazine *Scientific American* for more information on Professor Nicholson's life and research (Wenner 17 June 2008).

Nicholson proclaimed, was: “Very practical, quick, cheap, high-throughput relative to all other things…So it has everything going for it…Game set match.”

At its core, metabonomics entails the use of biochemical technologies like nuclear magnetic resonance (NMR) and mass spectrometry (MS) to visualize metabolism as a series of biochemical compounds that can be detected in biological fluids or tissues. However, the majority of time and effort in metabonomics research is dedicated to the data practices—the computational methods, algorithms, and statistical techniques—that are used to interpret the biochemical data produced by NMR and MS experiments. Such data practices—as I discuss throughout Chapter 3—are used by metabonomics researchers to clean up, standardize, and examine the patterns contained within biochemical data. They are predominantly made up of multivariate statistics—a domain of statistics involving the observation and analysis of many variables simultaneously, often in large data sets—and are varied in form and function. Along the way, researchers use data practices to negotiate and draw biological conclusions about the meaning of the biochemical variables and data points generated in experiments.

Several months later, mid-way into my fieldwork, I sat in on a small meeting of researchers involved in an array of projects directed at the development of translational medicine. This was the notion—as I discuss in Chapter 5—that scientific research should be translated into practical clinical applications through collaborations between clinicians and laboratory researchers. At the meeting, Jeremy Nicholson unveiled grand plans for the medical future of metabonomics research. Gathering the researchers around a sheet of paper, he pointed to a complex diagram that showed patients and samples flowing through various departments and stages of hospital treatment (see Figure 4). Part of this, he explained, involved several upcoming projects to track the “patient journey” through hospitals using a combination of biochemical techniques and biobanking approaches (Kinross et al. 2011). This was essentially a vision of “personalized medicine”—as I discuss in Chapter 6—the idea that medical treatments can be tailored to an individual’s biology.

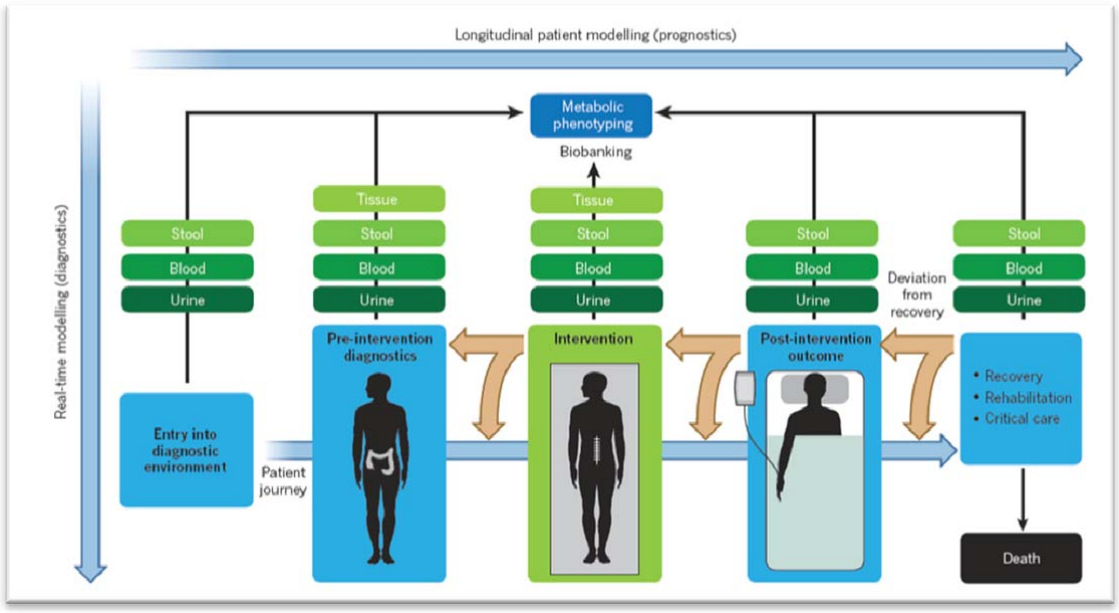


Figure 4: Translational research and personalized medicine.
Diagram showing how metabonomics researchers envision their technologies being used to diagnose and treat patients at various stages of their care in clinical environments (Nicholson et al. 2012b).

For the rest of the meeting, the researchers discussed which biomedical technologies and approaches would be used to gather biochemical data from various clinical projects. Discussing the state-of-the-art machines that would be used to carry out diagnoses and interventions on patients, the researchers stressed the importance of developing advanced statistical algorithms and models that could make sense of large volumes of complex metabolic information. They emphasized that collecting enough and the right kinds of data—as I discuss throughout Chapter 5—would be fundamental to the success of the project, and would involve intense collaboration with clinical practitioners throughout the network of hospitals operated by Imperial College London.

Embedded in this discussion was the recognition that much of the work in metabonomics, and post-genomic research more broadly, involves making sense of complex biological information. Metabonomics researchers use multivariate statistics—as I discuss throughout Chapter 3—to engage with metabolism as an inherently dynamic and complex biological process. By using particular technologies to engage with metabolism, they present particular visions and articulate particular ways of thinking about the components and relationships that make up biology and life. But researchers also use multivariate statistics—

as I discuss throughout Chapter 4—to negotiate the boundaries between states of health and disease through quantitative measures of the normal and abnormal. In doing so, they constantly encounter challenges with the contingency and instability of quantitative disease categories, whose binary measurements do not always match up to the variety of disease symptoms and experiences encountered outside of laboratory environments.

Ultimately, such clinical developments in metabonomics beg important questions about the increasing role of post-genomic research—and the data practices it encapsulates—in defining, measuring, and addressing issues of health and disease. They ask us to consider how the roles and forms of data practice in biomedical research are changing, and how new developments in the biomedical sciences are influencing the study of the human body. They raise questions about how metabonomics, with its embrace of the molecular and the statistical, might provide new languages for and ways of seeing the biological world.

Though metabonomics is undoubtedly producing interesting and novel approaches to biomedical research, the field's contact with clinical practice also raises several flags of caution. How optimistic should we be about developments in metabonomics that claim to augment or replace existing clinical practices? Should we interrogate more carefully biology's turn to the complexity or personalization of biomedicine, and the meanings and values such terms invoke or exclude? To what extent do quantitative statistical practices succeed in capturing the dynamic and contextual meaning of health? Furthermore, to what extent do statistical practices enable medical professionals to provide personalized or individually tailored medical care? Together, these questions signal concerns about the increasing centrality, as well as the seeming lack of limits to the power, of statistics to biomedical research and care.

Chapter 1: Anthropology in the Metabonomics Laboratory

Introduction

Social studies of science have, for the last several decades, examined what the sciences are and do. They have sought to unravel the common portrayal of science as an objective form of knowledge that is based on methodical and rigorous experimentation. Examining the practices and contexts of scientific research, studies in sociology, anthropology, history, and the interdisciplinary field of science and technology studies (STS) have revealed that science, like other forms of knowledge, constitutes a particular way of describing and knowing the world (Latour et al. 1986). Scientific practices, in other words, are crafted, contextual, and intricately tied to the material realities of the laboratories in which they occur.

To examine biomedical practice, social scientists have ventured into laboratories to examine the production of scientific facts, or have travelled to diverse locales to examine how facts come to have particular meanings and effects in society. Ultimately, the long-standing goals of such research have been "to explore how new facts and objects are progressively becoming defined and definitive, to investigate the practical, material, and organizational conditions for the establishment of such new realities" (Moser 2011:707). In this sense, emerging biomedical practices provide insight into the technologies, ideas, values, and epistemologies that characterize modern cultures of biomedicine.

Following the work of those anthropologists and sociologists who have sought to understand the nature and effects of scientific practice (Traweek 1988; Rabinow 1996; Helmreich 2000; Myers 2006), this dissertation provides an ethnographic account of the research practices and culture of the field of "metabonomics," as well as of the laboratory in

which it is being pioneered. As the post-genomic study of metabolism, metabonomics attempts to model the metabolism of organisms by applying complex statistical practices to biochemical data. As such, this dissertation explores the ways in which daily laboratory practices are used in metabonomics research within and beyond the laboratory to ask and answer questions about health and disease. It also examines the constellation of histories, technologies, ideas, and values that metabonomics practices entail.

This introductory chapter serves a number of purposes. First, I provide an overview of the field of metabonomics, situating it within developments in genetics, systems biology, bioinformatics, and informational practices. Then, I examine how both the anthropology of biomedicine and feminist studies of technoscience provide a basis for examining the entanglement and reciprocal relationship between scientific practices and cultural ideas and values. Consequently, I examine how scholarship on visualization and the ontological enactment of scientific objects provides a framework for examining data as a material practice. I then examine how scholarship on the concept of "normal" highlights the moral, social, and political role of biochemical and statistical ideas of health and disease in modern biomedicine. From there, I provide an overview of my fieldsite and methods, discussing the challenges and ethical questions raised by my ethnography. To conclude, I provide an overview of chapters that make up the analytical component of this dissertation.

Defining Metabonomics in the Post-Genomic Era

"Metabonomics" is the post-genomic study of metabolism. It is the study of the "raw materials and products of the body's biochemical reactions, molecules that are smaller than most proteins, DNA and other macromolecules" (Pearson 2007). These small biochemical compounds are called "metabolites," and the goal of metabonomics is to measure and model metabolites in bodily fluids and tissues (Nicholson et al. 1999; Nicholson et al. 2008). Metabonomics, therefore, provides a snapshot of an organism's metabolism by describing, measuring, and quantifying its "metabolome," the sum of its biochemical reactions and its "full and unique cocktail of metabolites" (Hunter 2009) (see Figure 5).

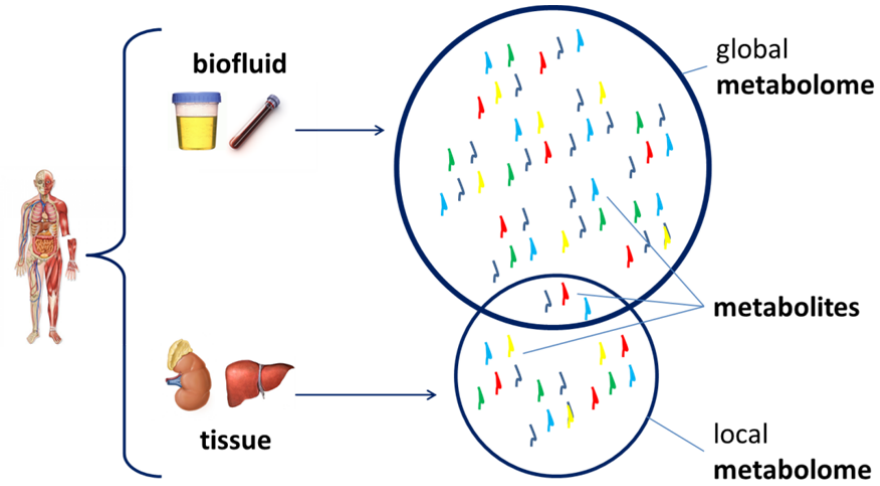


Figure 5: Metabonomics and the metabolome.
Diagram showing how metabonomics is used to examine the metabolome, the sum total of metabolites within an organism. Metabonomics can measure the metabolites contained within biofluids or tissues (left), which can be used to draw conclusions about the local metabolism (i.e. of specific organs) or about the global metabolism (i.e. of the entire organism).

An organism's metabolome, however, changes in response to environmental stimuli such as food, drugs, or diseases. While the genome, the sum total of genes, remains constant over life, the metabolome is a rich source of information about the dynamic and temporal status of an organism. Metabonomics researchers leverage this fact—that the metabolome is a dynamic and complex entity which reflects the result of gene-environment interactions—to claim that metabonomics is in many ways superior to genomics. Metabonomics, researchers assert, comes closer to highlighting the "phenotype" of organisms than genomics, and therefore has more direct applications to biomedical research and interventions. As one researcher commented during an interview: "[Metabonomics is] the most powerful [approach] from a purely theoretical point of view, because it captures gene environment interactions. So whether you're healthy, wealthy, or wise is determined by the interactions of your genes with what you do in your life, and how you're fed, and all of this other stuff."

Using this metabolic information, the end goal of metabonomics is to find ways to measure the causes and outcomes of disease processes. To do so, metabonomics seeks to determine "biomarkers" (see Chapter 3), measurable and quantifiable biological entities that can be statistically determined in relation to health and disease (Holmes et al. 2008b; Metzler 2010). Such biomarkers shed light on the "the health status of the organism as a whole" (Barton 2011), producing information that can be used to predict disease or monitor

treatments (Pearson 2007). Though the investigation of biomarkers as indicators of health and disease has been a part of biomedical practice for decades—for example in the detection of heart disease via blood cholesterol—researchers assert that metabonomics offers a more "sophisticated" and "objective" way of determining health and disease at the molecular level. This is done by using a variety of biochemical technologies, primarily nuclear magnetic resonance (NMR) spectroscopy and mass spectrometry (MS), to quantify the biochemical composition of biological fluids and tissues. These biochemical technologies are used in conjunction with multivariate statistical and computer modelling technologies, which attempt to find the patterns and associations contained within biochemical information (see Chapter 3). An explanation of these and other key technical terminologies is included at the end of this dissertation in the Appendix.

In everyday metabonomics practice, the measurement and modelling of metabolites is a challenging affair. This is due to the number and variability of metabolites among organisms, as well as to the dynamic ways in which metabolites change in time and space. As one scientific commentary observes:

> The number and nature of compounds in the human metabolome will vary depending on which body fluid is looked at and the method used for the analysis. There is also no clear division between compounds produced by the human body, those produced by our gut bacteria and fleeting products generated by food or drugs swallowed that day. (Pearson 2007)

Moreover, the human body's metabolism is not simply "human." It is made up of nutrients from food, and also from bacteria, which constitute 1 to 3% of human body mass and outnumber human cells by 10 to 1[7]. The number of metabolites in organisms, moreover, increases exponentially with biological complexity. The number of metabolites in bacteria, one of life's simplest unicellular forms, is more than 600, while the number of metabolites in plants, moving up in complexity, is roughly 20,000. The number of metabolites in human beings, which have numerous cell types, tissues, and organ systems, is massive. While there are a known 20,000 genes in the human genome, biochemists estimate that the number of small-molecular metabolites in the body is anywhere between 2,000 and 100,000 (Blow 2008).

[7] See a recent review written by members of the BMM in the journal *Science* on the metabolic interactions between humans and gut bacteria (Nicholson et al. 2012a), which is part of a Special Issue on the Gut Microbiota.

It is important to note that metabonomics is similar to—or arguably synonymous with—the related field of "metabolomics" (see Chapter 2). While metabolomics focuses on the characterization of metabolism at the cellular or organ level, metabonomics focuses on the effects of environment, disease processes, and gut bacteria on metabolism at the level of the organism. Because the divisions between these approaches to the study of metabolism are not well defined, the two fields entail different but overlapping sets of practices. For example, the development of metabolomics is attributed to work by a variety of groups with MS on model plants and organisms—including a team of researchers at the University of Manchester who coined the term "metabolome" in a 1998 paper (Oliver et al. 1998)—while the development of metabonomics is attributed to the work of Jeremy K Nicholson with NMR on bodily fluids (Nicholson et al. 1999). Despite these differences, in practice laboratories conducting metabonomics and metabolomics research use similar analytical platforms and approaches to data analysis.

Formally, metabonomics is described in relation to the developments in genomics and "big science" that have occurred over the past few decades (Glasner 2002), and which have allowed scientific research to increase in size and intensity (Gitelman 2013:2). The Human Genome Project (HGP), which started in the 1980s and culminated in the full sequencing of the human genome in the 1990s, revolutionized the scale and scope of scientific research. Scientists began to focus on the molecular inner-workings of organisms and cells, and also carried out the discovery and analysis of genetic information in increasing volumes and smaller amounts of time. But as Kaushik Sunder Rajan (2005) has commented, the technical innovation and epistemic advances of the HGP were paralleled by the increased commercialization of biological work, which saw increasing partnerships—and money making ventures—between academia and businesses. This transformed the character of biological research, such that it became increasingly intertwined with commercial-scale research machines and methods, and also with commercial outputs aimed at medical markets and settings. Metabonomics, in a sense, follows clearly from these "post-genomic" developments. As it analyses organisms at a molecular level, it is characterized by increasingly rapid, large-scale, and corporate practices. It entails collaborations with instrument manufacturers and pharmaceutical and nutritional companies (see Chapter 2), and is concerned with the development of diagnostic and therapeutic medical applications.

Ultimately, the hallmark of metabonomics research is its integration with and reliance on data practices. The metabonomics laboratory, as it conducts research on the various facets

of metabolism, is permeated by informational ideas and techniques, including statistical methods, computer programs, and training sessions. In contrast to experiments in fields like molecular biology that revolve around the physical manipulation of biological materials, experiments in the metabonomics laboratory involve generating and (re)working quantitative statistical models of biological information. This typifies the differences between "wet" and "dry"—also known as "laboratory" and "statistical"—styles of research, with "dry" research tending to occur more in post-genomic research on large quantities of data (Penders et al. 2008). Metabonomics researchers, therefore, spend the majority of their time and effort standardizing data and running algorithms (see Chapter 3). In this sense, metabonomics has much in common with "systems biology," a growing interdisciplinary field that combines the efforts of physicists, computer scientists, engineers, and mathematicians to make sense of and transform vast amounts of data into a mathematical understanding of life (O'Malley et al. 2007b; Calvert et al. 2011)[8].

Overall, metabonomics typifies the increasing centrality of data practices to modern biomedical research (Moody 2004; Thacker 2005; Swedlow et al. 2011). Such practices are generally referred to as "computational biology" and "bioinformatics," which encompass practices for "storing, searching, organizing and managing biology data…and making that data into biological knowledge" (Stevens 2011:20). Bioinformatics and computational biology are diverse in form and function, but they are fundamentally comprised of techniques for dealing with data. As Kim and Mike Fortun have observed with toxicogenomics, the work of modern biology often revolves around "information—its quantity and gaps; its circulation, standardization, and interpretation and its promise as a lever of new knowledge" (Fortun et al. 2005:49). Thus, in the landscape of post-genomic research, data practices are no longer one of several components of biological work: instead they take center stage, and form a large part of the conceptual and physical effort in laboratory settings[9] (see Figure 6).

[8] Metabonomics' relationship to genomics and systems biology is also evident in its use of the suffix "omic," which generally "highlights the study of a comprehensive collection of data" (Hotz 13 August 2012).

[9] While the use of computers and statistics in biology dates back to work on protein crystallography in the 1950s and 1960s (Hirschauer 1991; Hagen 1998), computers and statistics have become crucial to the organization and analysis of biological data in the wake of the HGP (Mackenzie 2003).

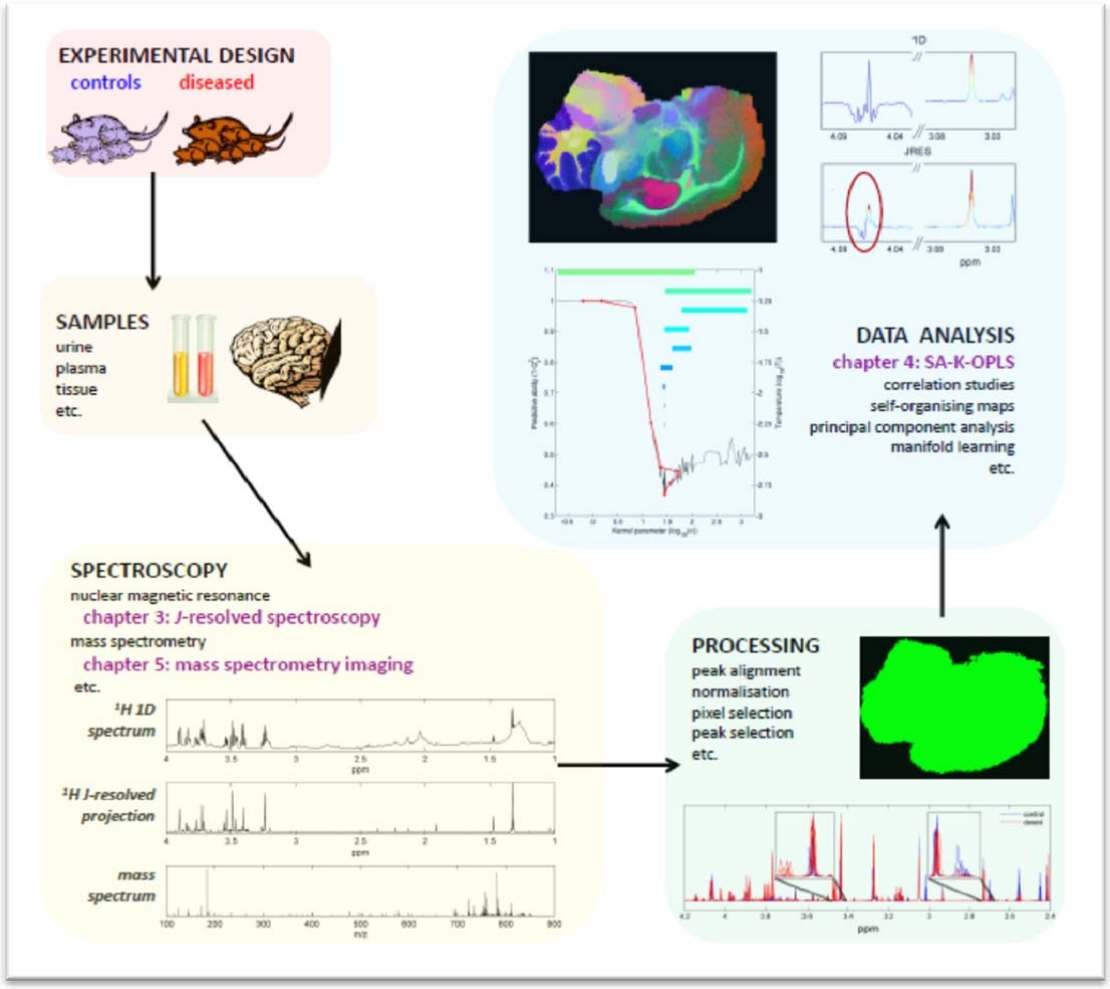


Figure 6: Data practices in metabonomics research.
Diagram taken from a doctoral thesis, showing the centrality of data practices to metabonomics research. The left side shows the steps involved in *obtaining data*, including sample acquisition ("Experimental Design"), preparation ("Samples"), and analysis with biochemical techniques ("Spectroscopy"). The right side shows the steps involved in *analysing data*, including data standardization ("Processing") and interpretation ("Data Analysis").

Data practices, however, do not simply involve the application of computing and informational techniques to "old" biological problems. Rather, they entail a shift in the practices of knowledge production, and consequently in the kinds of knowledge and values that are produced. In genomics and metabonomics alike, data practices are used not only to organize vast amounts of data, but also to interpret and make sense of biological information. In his book *Silicon Second Nature*, Stefan Helmreich writes that Artificial Life is "more than a new way of thinking about biology…[It] is a symptom and source of mutating visions of 'nature' and 'life' in an increasingly computerized world" (Helmreich 2000:11). Similarly,

Carlo Caduff has commented in his work on avian flu and biosecurity that bioinformatics processes render bodies into "informational forms, rather than corporeal ones" (Caduff 2012: 343). Thus, as scientists develop new practices for engaging with biological information, bodies and biological life acquire new materialities. New informational practices not only transform the way we interact with and experience technologies, but also entail shifts in how researchers engage with the meaning of life.

Why Metabonomics?

In engaging with metabonomics as a form of post-genomic research, this dissertation attempts to disengage a critical study of science from the hype and "newness" of genetic/genomic technologies, innovations, and discoveries. Any discussion of an emerging area of scientific research must engage with issues of novelty: scientists speak of the unique promise of their research to gain funding and publicity, and also to emphasize the fundamental role that biomedicine plays in society. Engaging with this novelty in contemporary biomedicine, many scholars have examined the impacts and rhetoric of genetic and genomic technologies, as exemplified in studies of the "geneticization" of society (Lippman 1992) and the "sociology of expectations" (Brown et al. 2003; Fortun 2008). Genomics, they write, reinforces deterministic and causal explanations of life, and also performs and normalizes the future value of emerging genetic technologies.

While such research has made important observations about the character of contemporary biomedical research, I would argue that it has suffered from several overarching problems. Firstly, social scientists have tended not to examine the material practices of biomedical research, basing their claims instead on what is written or said about science. Doing so has left the complex and varied daily workings of post-genomic science unexamined, placing primacy on public concepts of science over the materialities of laboratory practice. Social scientists have explored stabilized portrayals of scientific fields and concepts, instead of exploring the multiple and contingent ways that science is practiced. Secondly, by focusing on those biomedical research projects that have captured media and public attention, social scientists have tended to reproduce the hype and promises surrounding genetic and genomic technologies. This has led to scholarship focused on the "new," which has simultaneously obscured how emerging research is deeply connected to

historical developments and constantly evolving practices[10]. While tracking rapidly emerging fields of research is understandably difficult, these tendencies in the social sciences have reproduced dominant narratives about biomedical practice, leaving the diverse and heterogeneous practice of science unexplored.

To this end, this dissertation does not necessarily suggest that metabonomics is unique in its attempts to explore biological life. Many other forms of biological research—including but not limited to epigenetics, proteomics, and immunology—are marked by data practices, as well as the integrated work of biologists, statisticians, and computer scientists. My aim, rather, is to illustrate how metabonomics crystallizes the negotiations and challenges inherent in the emerging biomedical sciences, as they produce and grapple with large volumes and specific types of biological information. Metabonomics as a field highlights the increasing centrality of data practices, and in particular multivariate statistical forms, to the organization and analysis of biological work. Because metabonomics is an emergent and relatively under-explored field, its daily work is explicitly oriented—in ways that more established fields like genomics are perhaps not—towards developing and working through problems of biological information.

The importance of metabonomics as an object of anthropological attention, however, is not limited to the insight it provides into data practices. Metabonomics—like the fields of epigenetics and proteomics mentioned above—represents a shift away from the supposed reductionist and gene-centric approaches of the 20th century HGP (Kay 2000; Keller 2002), and towards the network and systems approaches of the 21st century (Glasner 2002; Landecker 2011; Niewöhner 2011). Metabonomics, in its embrace of gene-environment relations, highlights the holistic and ever-changing nature of life. Emphasizing the ways in which metabolism is made up of relationships, flows, and permeable boundaries, metabonomics produces an overarching narrative of biological complexity and dynamism.

The extent to which this narrative is reflective of biological and laboratory practice, however, remains unclear. Such an emphasis on the complex and dynamic nature of life is paralleled by increasingly informational and statistical ways of investigating life. As such, metabonomics as a field entails investigations into metabolic systems, which, in reality, entail

[10] Exceptions include, among others, Hans-Jorg Rheinberger's account of molecular biology (1997), Lily Kay's account of genetics (2000), and Hannah Landecker's account of cell culture (2007).

the reduction of biological processes and diseases to numerical values and statistical patterns. As rhetorics and narratives meet during everyday practices, metabonomics highlights how certain types of information are not only produced and valued—but also amenable to quantification—over others. Consequently, a study of metabonomics affords a critical examination of the claims surrounding the relationality and complexity of the post-genomic sciences.

Metabonomics Research as Culture

As a starting point, this dissertation explores the ways in which metabonomics research—as an emerging form of post-genomic knowledge, as a constellation of techno-scientific practices, and as a "new" realm of biomedical investigation into the body—can be understood as a form of cultural practice. Drawing on literature in medical anthropology, feminist studies of technoscience, and science and technology studies (STS), I argue that my study of the metabonomics laboratory is deeply embedded within recent efforts in anthropology to engage with the ways in which the biomedical sciences both shape and are shaped by society and culture.

Examinations of the social and cultural dimensions of science and technology have a long history within anthropology. Beginning with its examination of Western biomedicine in the 1980s and 1990s, scholars explored the distinction between disease and illness (Kleinman 1982; Singer 1990), critiqued the hegemony of biomedical understandings of the body (Lock et al. 1993; Lock et al. 1998), and provided important critiques of Western medical practice (Good 1994; Farmer 2001). Building on those roots, anthropologists have contributed to studies of scientific practice by highlighting the importance of performativity and phenomenology of the human body (Scheper-Hughes et al. 1987; Csordas 1994). In doing so, they have shaped an interest in the embodied rather than representational aspects of medical practice (Hadolt et al. 2012:181). This has focused anthropological attention on practices—on what people do and experience in everyday life—instead of exclusively attending to cultural signs, symbols, and meanings (Lévi-Strauss 1962; Douglas 2002).

Turning their attention to the practices of science and technology, and building off of a rich tradition of questioning the concept of "culture" (Franklin 1995), anthropologists have also explored how science and culture are interwoven. Biomedicine, already established as contingent and social, is not carried out "in a world scissored off from the culture in which it exists" (Helmreich 2000:19). Characterized by Emily Martin's (1995) figure of the "string

figure," scholars have shown how science is part of the fabric of culture, in which biomedical facts penetrate into and interact with diverse components of society. Chipping away at the primacy of scientific knowledge, anthropologists have highlighted how emerging medical technologies have local meanings, as well as diverse impacts on practitioners, patients, and advocacy groups (Lock et al. 2000; Collier et al. 2005; Franklin et al. 2006; Fassin 2007; Gibbon 2007; Fullwiley 2011). For example, Rayna Rapp's *Testing Women, Testing the Fetus* (1999) examines the social impact of amniocentesis, a medical technology used for the prenatal diagnosis of genetic conditions. As she details the history and practice of amniocentesis and genetic counselling—showcasing how practitioners deal with sensitive and complex issues like genetic risk—she untangles the factors that contribute to and create the tensions between scientific and lay conceptions of medical technologies, heredity, and risk.

In a similar vein, STS scholars have examined the ways in which the laboratory itself can be viewed as a distinct culture (Traweek 1988). For example, Karin Knorr Cetina's *Epistemic Cultures* (1999) documents the ways in which molecular biology and high-energy physics laboratories have distinct power relations, social networks, technologies, and ways of communicating. In showing that laboratories are not monolithic, and that distinct configurations of technologies, people, and histories give rise to unique practices and ideologies, her study is an important critique of the idea that science is homogenous and devoid of social and cultural dimensions.

There is no question that anthropology has produced a rich body of literature exploring the ways in which "diagnostic and therapeutic tools and practices profoundly affect bodies, identities, and infrastructures in medicine" (Hadolt et al. 2012:182). But, as Bernhard Hadolt et al. note, anthropology has also been slow to embrace the material and technical aspects of biomedicine, and in ways that attend to them as central theoretical matters rather than topical interests. In a similar way, it has been argued that anthropology—in its examinations of the embodied experience of illness and the social character of biomedical knowledge—has tended to "black box" scientific practice[11]. By creating artificial distinctions

[11] As Ian Hacking defines the term "black box":

> These include not only off-the-shelf apparatus but also all sorts of systems for operating on symbols, for example, statistical technologies for assessing probable error. [They] include standardized pieces of apparatus bought from an instrument company, borrowed from a lab next door…the laboratory worker seldom has much idea how the box works…yet it encodes

between the biological and the social, it is argued, anthropology has not attended to the material laboratory practices that are conceptualized as being upstream of society (Franklin 1995:169; Lock et al. 2000; Gibbon 2007).

Responding to this critique of anthropology, feminist studies of technoscience provide a fruitful line of investigation into the "performative and material aspects of biomedical techniques" (Hadolt et al. 2012:183). Emerging from studies of kinship (Strathern 1992a; Strathern 1992b) and biological conceptions of gender (Haraway 1991b; Haraway 1995), such studies have examined both the practice and also the "history and foundations of the natural" within scientific research (Franklin 1995:170). A seminal contribution to this line of inquiry is Marilyn Strathern's examination in *After Nature: English Kinship in the Late Twentieth Century* (1992b) of the ways in which kinship invokes and collapses categories of the "natural"—what is biological or genetic—and the "cultural"—what is deemed by society to be a relative or family relationship. Strathern argues that new reproductive technologies have changed the "biological" and "natural" facts of reproduction, revealing that the natural is as much socially constructed as the cultural. Thus, neither the natural nor the cultural are pre-existing categories: cultural assumptions play key roles in creating the categories of natural and cultural, and in structuring the distinctions between them.

Building on Strathern's imperative to break down the ontological distinction between such categories as nature and culture, Donna Haraway's work also provides a seminal contribution to feminist studies of technoscience, by exploring how modern biomedical technologies inherently collapse the boundaries between existing cultural categories to generate new hybrid categories or "border crossings" (Haraway 1997:60). In *A Cyborg Manifesto* (1991a), Haraway employs the concept of the "cyborg" to show how science and culture are intricately bound through modern society, leading to hybrid entities such as mind/body, human/animal, and organism/machine. Haraway's point is that modern society, which is filled with scientific ideologies and technologies, reconstructs the boundaries between nature and culture or human and non-human, and creates new ways of intervening into the natural world and into bodies. Ultimately, her agenda is political: by questioning the construction of nature and other categories like race, class, and gender, her work is a call to

---

> in material form a great deal of pre-established knowledge which is implicit in the outcome of an experiment. (Hacking 1992:42)

reconceptualize feminism by revealing and challenging the assumptions of dominant discourses and narratives.

Feminist studies of technoscience, as exemplified here by the works of Marilyn Strathern and Donna Haraway, have "redefined the possibility of studying science as culture" by questioning the ontological distinction between the concepts of the natural and the cultural (Franklin 1995:173). They have opened up new lines of anthropological inquiry into the cultural basis of the natural objects of scientific research, providing anthropologists with the conceptual framework to explore how the biological is actively made social rather than relegated to the status of natural. Consequently, scholarship on "new" biomedical technologies—organ donation in brain-dead patients (Lock et al. 2000), cell culture (Landecker 2007), amniocentesis (Rapp 1999), among others—has expanded the scope of anthropological research by challenging us to examine the entanglement between the natural or scientific and the cultural.

Further exploring the interactions between culture and science, scholars have focused on the ways in which science and culture reciprocally interact with and influence each other in daily scientific practice. Scholars such as Ian Hacking (2007) and Sheila Jasanoff (2004) have explored how cultural ideas and values are taken up and embodied in laboratory practices, and how they are consequently remade through those same practices to flow in and out of scientific arenas and society. In *Kinds of People: Moving Targets* (2007), Hacking develops the concept of "looping effects" to examine the reciprocal effects of scientific classification on human beings. Providing examples such as autism and obesity, Hacking argues that people actively take up the labels of classification, such that classifications change the people who are classified and lead to new classifications. Here, Hacking's work emphasizes how categories and labels are enforced by experts and institutions, which consequently legitimize practices like counting, norms, clinical medicine, and genetics (Hacking 2007). Although Hacking's primary focus is on the ways in which classifications affect peoples' identities and subjectivities, his argument highlights the reciprocal relationship between technoscientific practices and society.

Also exploring the reciprocal relationship between scientific practice and culture, other scholars have examined the entanglement between technical practices and ways of seeing the world. Approaching modern biomedical sciences as situated practices, Peter Chow-White and Miguel García-Sancho (2011) have developed the concept of "bidirectional

shaping" to examine how the introduction of computers and informational technologies into scientific settings affects not only the functioning and practice of computers, but also their development within institutions. Similarly, Natasha Myers has examined how within the world of proteomics, molecular animations not only draw on pre-existing knowledge of proteins, but also "refigure these ways of seeing and the very theories they enact" (Myers 2006:16). Together, these concepts emphasize the ways in which scientific and cultural practices mutually inform one another, such that the technologies used to conduct laboratory research shape and are shaped by societal ideas and values.

Taken together, such anthropological, feminist, and STS scholarship on biomedicine forms a fundamental starting point for this dissertation. This literature establishes how science must be examined as a cultural form of practice, and also how biological or "natural" objects of research have inherently cultural meanings and consequences. Thus, this dissertation explores how metabonomics research and the ideas it produces are the result of particular practices and techniques, and are also imbued with cultural ideas, values, and histories. With this line of inquiry—which critically examines the claim that biomedicine is an objective or epistemologically superior form of knowledge—this dissertation explores how metabonomics research results from the active work of scientists carrying out work within particular cultural and historical settings. It also explores how metabonomics research interacts with society to produce situated understandings of biology and the body.

Data as a Material Practice

In the era of post-genomic research and bioinformatics, much STS scholarship has focused on the visual graphs, charts, models, maps, and representations that form the end products of scientific practice (Carusi 2012). These visual representations are diverse in form, and are increasingly generated with computerized interfaces and statistical practices. Importantly, these visual representations of data analysis are integral to "making data meaningful," in that they summarize relationships and communicate meanings within data, and also enable particular modes of understanding and seeing data (Dumit 2003). This is not to say that the pathways from scientific experiments to the communication and visualization of results are linear: the rendering of biological substances into information, and consequently of experimental information into visual data, involves active work, interpretation, and standardization.

To this end, STS and anthropological studies of visualization have examined the use and production of visual images to gain insight into the varied modes and ideologies of scientific practice. Scholars have explored how scientific objects become "visible" through embodied practices of production (Alac 2008), and how seemingly pre-given visual objects are brought into being—and to have meaning—through historically and culturally specific practices and ideologies (Dumit 2003; Daston et al. 2007; Myers 2008). Importantly, studies of visualization have shown how the visual does not merely represent an objective reality, but rather intervenes into nature to create such a reality (Lynch et al. 1990). In this sense, visual representations cannot be understood in isolation from the contexts in which they are produced and used. Studies of visualization have also shown how the production of visual objects transforms data into something objective and universal, removing uncertainty from scientific practices and observations. As such, visualizations naturalize particular ways of seeing and interacting with the world, by privileging some points of view and silencing others (Haraway 1997).

Questions about the objective nature of visualizations harken back to many of the first investigations into scientific research within the fields of sociology and STS, which questioned the "objective" nature of scientific knowledge, and challenged the status quo of Western or technological approaches to knowledge. These studies argued that scientific knowledge only appears to be different, superior, and unquestionable, and is instead socially constructed. According to this social constructionist school of thought, scientific knowledge operates through the careful process of laboratory craftwork, and is fundamentally shaped by social, economic, and historical configurations (Latour et al. 1986). This process of craftwork involves diverse processes and practices, including the protocols that standardize research (Timmermans et al. 1997), and the social and technical configurations that legitimate science through experiments (Shapin et al. 1985). Ultimately, such studies of laboratory science were fundamental for establishing that data and facts are not pre-given and do not precede their discovery in scientific settings.

While questioning the objective nature of scientific knowledge, scholars have also sought to question the very nature of "objectivity" itself. Objectivity implies rigor and neutrality to which scientists strive in the everyday production of knowledge. It implies, in other words, a set of "epistemic virtues" that provide norms and values for how particular forms of knowledge should be produced. But as Lorraine Daston and Peter Galison (2007) argue, objectivity is embedded within different histories and contexts. Detailing three epochs

of objectivity—truth-to-nature, mechanical, and trained judgment—they show how notions of the "objective" are associated with different meanings and scientific practices. Notions of objectivity therefore obscure the contingency of knowledge, and the ways in which it is produced relative to particular constellations of technologies and values. This raises important questions about how and why scientific forms of knowledge are rendered as more "objective" than other more embodied forms of knowledge.

Despite the large body of social scientific work that has critically examined practices of visualization, there has been comparatively little investigation into the practices of information, data, and statistics that give rise to such visualizations. Within the social sciences, data—and consequently the practices that give rise to it—have been black-boxed. They have been rendered as pre-given, objective, and stable entities that function as a backdrop to, rather than a fundamental aspect, of modern biological work[12]. Despite this apparent gap in the literature, the sociological (Alonso et al. 1987; Desrosières et al. 2002) and historical (Hacking 1990) aspects of data practices have been examined in various ways: Joan Fujimura and Rayma Rajagopolan (2011) have examined how statistical practices in genomics draw upon and enact particular conceptions of race and ethnicity, and Peter Keating and Alberto Cambrosio (2012) have examined how statistics are adapted to the field of microarray analysis. Similarly, Adrian Mackenzie (2003) has examined how bioinformatics articulates biopolitics and property relations, and Sheila Jasanoff (2002) has examined how statistics have been used in court settings to override subjective and individual claims.

Overall, these studies exhibit attempts to grasp the complex effects of data practices and knowledge. They do not, however, engage with the ways that statistics are practiced and negotiated in everyday laboratory settings in *material ways*. Although social scientists have begun to acknowledge the dense relationships and challenges that surround data practices (Stevens 2011), they have, for the most part, relegated statistics to the cognitive domains of reasoning and mental activity (Greiffenhagen et al. 2011). Data practices are no different from other forms of scientific practice: they appear as objective and natural, but they are caught up in material and social networks. As Nathan Ensmenger comments in his analysis of artificial intelligence algorithms, this is due largely to the "inherently amorphous nature of

[12] One statistics-savvy researcher, who was formerly a member of the metabonomics laboratory, commented about the prevalence of black boxes in metabonomics practices. He said during an interview: "I guess one extreme instance you see is the black box – so you have data, you train your black box, then the black box can predict something that is of importance to you."

software…[which]…is generally invisible, ethereal, and ephemeral. In many cases, it exists only as a unique – and temporary – arrangement of digital bits buried deeply within a tiny microprocessor" (Ensmenger 2012:8). Thus, studies exploring the negotiation of contingent and unstable data practices have remained critically absent from the anthropological and science studies literature. To this end, this dissertation aims to foreground and explore the production and use of data practices—and their relationship to other laboratory practices—in contemporary biomedical research.

This dissertation also attempts to further the theorization of statistics as material objects and practices, by following an ontological approach that takes scientific practice rather than knowledge as its point of departure. Here, practice refers to the study of what things *do* rather than *represent*. Though practice is linked with and shaped by culture, it focuses on the material world and the people who inhabit it, and on the ways in which they reflect important ideas values and power dynamics. This focus on practice is important for shifting attention away from the intellectual products, published accounts, or perceptions of data practices, and towards the active work that goes into making and performing the ideas, technologies, and relationships that surround data within the laboratory.

Such a focus on scientific practice highlights a shift from "representationalism" to "performativity." It moves away from the notion that there is an inherent distinction between representations and that which they represent, or conversely, that practices are independent of the practices of representing. Taking inspiration from work on materiality, performance, and feminism, such an approach moves away from asking how scientific descriptions relate to reality, or as Karen Barad (2003) explains, from asking how representations mirror important dualisms such as nature or culture. Rather, studies of scientific practice examine how science is done, and what people and objects come to mean (Abram et al. 2011:7). Studies of practice are, in essence, studies of ontology rather than epistemology: they are studies of the lived world of cultures and technologies, rather than the cultural perspectives and beliefs that such worlds entail.

By focusing on performativity and practice, this dissertation also follows Donna Haraway's (1991b) call to consider the "material-semiotic practice of technoscience." Material-semiotic refers to the ways in which semiotics—the textual and symbolic dimensions of the world—are intertwined with and embodied in material practices. For Haraway, a material-semiotic approach reveals the ways in which social relationships and

meanings are imbued into objects, such that there is no separation between "meaning systems [and] the material worlds and bodies they structure, and are structured by" (Moreira et al. 2011:334). Considering the material-semiotic practices of science, Haraway insists that examinations of material practice should not discount cultural metaphors, symbols and beliefs, but rather should see them as intertwined with the performative dimensions of science[13]. With this approach, I define data practices—which I also refer to as data techniques (Hadolt et al. 2012)—not just as material realities, but also as constellations of people, technologies, objects, ideas, and values.

Following an ontological and material-semiotic approach to practice (Haraway 1988; Pickering 1992; Butler 1993; Mol 2002; Barad 2003), this dissertation engages with everyday work in the metabonomics laboratory to examine how data and statistics are enacted in practice: how they are brought into being and temporarily realized and stabilized, reflecting norms and assumptions about the world (Woolgar et al. 2013). Despite the seeming intangibility of data practices, they exist as material and performative objects within social, economic, and political networks. Data practices are more than cognitive or stable representations: they are enacted through diverse material and discursive practices, and are imbued with qualities, capacities, and values (Coopmans 2011:156). As with other material-semiotic entities, the meaning in data is only construed when ideas are turned into actions, such that data "encompas[s] not only computers, codes, algorithms, and ideas, but also people, practices, and networks of interaction" (Ensmenger 2012:8). Data practices, therefore, play a pivotal role in the enactment of biology and metabolism in the metabonomics laboratory, as they are carried out by a diverse range of people and practices.

Ultimately, taking an approach that combines these insights from studies of laboratory science and visualizations, this dissertation explores how metabonomics researchers learn to see metabolism and other concepts as "technologically mediated vision[s]" (Coopmans 2011:158) through particular data practices. In this way, this dissertation avoids black-

[13] Similarly emphasizing the material-semiotic nature of science, scholars have developed Actor Network Theory (ANT) to theorize the ways in which human and non-human "actors" are in networks of relationships and values (Latour 1987; Law 2009). While ANT is a valuable and widespread approach—particularly in its attention to non-humans and agency—its utility from an anthropological perspective is limited in several ways. Firstly, anthropologists argue that ANT problematically reduces humans and non-humans to equal "actors" within the network, obscuring the qualitative differences between relationships, and minimizing the agency of humans (Strathern 1996). Secondly, anthropologists argue that ANT struggles to differentiate between the various levels at which networks operate, as illustrated by examinations of how laboratories operate within broader networks of clinics and research agendas (Oudshoorn 1990).

boxing data practices, which form a key and unavoidable component of metabonomics, as things that are pre-given or immaterial. Instead, it views data practices alongside other laboratory practices—such as the biochemical analysis of metabolic samples or the writing of research reports—and as technologies that "embody and enable specific values, agendas, and possibilities" (Ensmenger 2012:26). With such an approach, this dissertation examines how data practices are both representations and performances of ideas and values, whose "objective" nature and normalization must be questioned.

Theorizing "Normal" and "Health" in Post-Genomic Research

Given the primacy of data practices in post-genomic research, one of the central aims of this dissertation is to unravel the processes through which statistical visions of metabolism, the body, and health come to occupy a central place in the everyday work of metabonomics. This dissertation examines, therefore, how metabonomics researchers struggle to reduce the dynamic and qualitative notions of health and disease to quantitative and statistical measurements. Consequently, the normal and abnormal are produced as the dominant narratives and objective forms of knowledge for engaging with health and disease in the metabonomics laboratory. As the normal and abnormal are used within biomedical and social settings to articulate concerns with health and disease in fundamentally statistical ways, they emerge as key areas of anthropological inquiry. Normal values are established in blood tests, urine samples, and other biological parameters in order to judge the health of patients, while behaviors and preferences are judged or measured against the normal, as a means for asserting what is ideal or moral. As philosopher Ian Hacking writes: "People, behavior, states of affairs, diplomatic relations, molecules: all these may be normal or abnormal" (Hacking 1990:160).

Though the concepts of normal and abnormal are central to the broad field of biomedicine, this has not always been the case. The word "normal" came into use in a non-mathematical settings in the 19th century (Ernst 2006:2), when, concurrently with the development of probability and statistics, it was used to describe the social and biological characteristics of populations. Recognizing the importance of normal and abnormal to biomedicine, key thinkers in medical anthropology have documented the challenges and issues surrounding concepts of normal and abnormal—and in particular their application to notions of health and disease—since the mid-19th century (Lock et al. 2010a). Ultimately,

this body of research has established not only that the seemingly banal categories of normal and abnormal are ubiquitous, but also that their production and use are morally, culturally, and politically charged.

Studies of the normal and abnormal were officially inaugurated into the oeuvre of medical anthropology with the work of philosopher Georges Canguilhem, who served as a mentor and inspiration for Foucault and his work on "biopolitics" and power in biomedicine (Foucault 1990; Foucault 2003). In particular, Canguilhem's work examines how normal and abnormal are not objective concepts, but rather are fundamentally imbued with medical values and judgments. In *On the Normal and the Pathological*, Canguilhem (1989) emphasizes how statistical and physiological concepts of normal and abnormal emerged in the 19th and 20th centuries as the predominant way for judging health and disease. Discussing the rise of physiology in the19th century, Canguilhem shows how biomedical concepts of normal were conceptualized as differing from the abnormal in a quantitative, measurable way. As such, health or disease came to be defined according to statistical "deviations" from physiological definitions of the normal. With the rise of physiology, normal thus became a proxy for health, as well as the primary object of investigation in biomedicine.

In examining the concepts of normal and abnormal in the 20th century, Canguilhem explores how modern biomedicine adopted a statistical rather than a normative conception of health. For Canguilhem, health was a function of the "biological normativity" inherent in organisms: it was the vital capacity of organisms to respond dynamically to the environment (Canguilhem et al. 2001). Health was the ability to establish new norms—new versions of the normal—in response to the environment, while disease was a lack of normativity, an inability of organisms to respond to the environment. In this line of reasoning, health and disease were qualitatively different and could not be measured in terms of quantitative variation from the normal. Normal could not be substituted for health, because it represented a normative condition rather than the capacity for normativity[14].

Drawing on Canguilhem's work, the concept of "the normal" in biomedicine emerges as a fundamentally moral and evaluative statement, and therefore represents the imposition of

[14]This line of reasoning also states that the normal does not imply the absence of disease, while the abnormal is not the same as disease and instead represents another type of norm. To consider this in less abstract terms, a competitive athlete can have an "abnormal" body weight but be in a state of health, while a bulimic can have a "normal" body weight but be in a state of disease.

biomedical practices, ideals, and values onto the vitality of body. In emphasizing that the biomedical normal is the selection of one norm from many possible norms, Canguilhem's work critiques the conflation of the normal with the healthy: it shows how statistical concepts of health arise through biomedical research practices and values, rather than through the general conditions of life. Investigations of normal as a way to understand health are "an artefact of the decontextualized clinical and laboratory methods used in biomedical research" (Mol 1998:275). Ultimately, Canguilhem's point is that the normal is never merely statistical or objective: instead, it is embedded within particular contexts, technologies, and represents an organism at a particular moment, place, and time. To equate normal with health is to make an evaluative judgement about what types of normal—and consequently what types of practices to investigate normal—are ideal or preferable.

While Canguilhem's work successfully interrogates the moral and evaluative qualities of the normal, philosopher Ian Hacking's work establishes how the statistical normal came to be a culturally-salient descriptor and category. Tracing the historical emergence of probability throughout Europe, Hacking examines in *The Taming of Chance* (1990) how normal became a guiding principle for describing, understanding, and dealing with societal and biological issues in the 18th and 19th centuries. Normal came to have a powerful statistical meaning in conjunction with the emergence of statistical technologies for data collection about populations. These included civic statistics in France in the 1820s and mortality statistics in Britain in the early 1800s, both of which enabled the control and surveillance of various populations, and, in particular, of societal "deviants."

In tracing the history of famous 19th century thinkers like Quetelet, Broussais, and Comte, Hacking demonstrates how the statistical normal became the predominant concept to which society strived, as it acquired both a descriptive and prescriptive meaning. Hacking details, for example, how the "normal state" as a key biological descriptor emerged through the physiological work of Broussais (Hacking 1990:166), whose research was also a subject of concern for Canguilhem. As the normal state became a guiding principle for physiology and biology, it was taken up and elaborated upon by Comte, one of the founding fathers of sociology, during a bought of mental illness. Comte, in applying the concept of the normal to his experience of mental illness, solidified its meaning not only as a statistically-measured population average, but also as a qualitative ideal to which individuals should strive. Ultimately, Hacking demonstrates how the concept of normal acquired a social meaning, and statistical concepts of normal became linked to notions of societal and biomedical progress

(Hacking 1990:168). A statistical normal thus emerged not only as a way to *describe* biological functioning, but also as a way to *determine* what was ideal with regards to disease diagnosis and treatment[15].

While the works of Canguilhem and Hacking are important for understanding the moral qualities and social ubiquity of normal in biomedical practice, the work of Michel Foucault is fundamental for exploring how the normal is implicated in processes of normalization and power. In his work on techniques of power—which include *Discipline and Punish* (1977) and *The History of Sexuality* (1990)—Foucault, explores how the rise of institutions and technologies for data collection promoted the exercise of power at the level of populations and individuals. As this occurred, institutions established and enforced particular concepts of normal—that is, normal relative to certain populations, practices, technologies, and values—as well as practices that contributed to the realization of such concepts. Consequently, biomedicine gained the power to control individuals and populations by linking statistical notions of the normal to notions of health and vitality.

As this occurred, writes Foucault, there was a shift from "anatomo-politics"—activities involved in the objectification and disciplining of bodies as machines—to "bio-politics"—activities involved in the regulation and management of bodies imbued with biological productivity (Foucault 1990). As the biological became a legitimate and natural matter for government and political action, forms of statistical knowledge like demography and epidemiology were co-developed with methods for visualizing and acting upon populations. Consequently, it was not only that populations were controlled through practices of "governmentality," but also that individuals were prompted to monitor and control themselves through "technologies of the self" (Foucault et al. 1988; Rose 2007; Lock et al. 2010b).

By exploring the ways in which statistical concepts of the normal are linked to the health of the body and populations, and through medical and social practices and

[15] Hacking details, for example, how Belgian statistician Quetelet examined the physical characteristics of Scottish soldiers in order to craft a "normal distribution" of height and chest circumference. What was important about Quetelet's work, according to Hacking, was that he transformed a formerly abstract notion of the average into a measurable, population-based quantity. Only by developing the technologies and by selecting a population to measure specific characteristics was Quetelet able to understand human anatomy in terms of a "normal distribution." Normal was not a phenomenon that occurred in nature, but rather was created by—and subsequently enabled—statistical techniques. Thus, Hacking's discussion of Quetelet demonstrates how the concept of "normal," as it came to be understood as a population-based average, was imbued with statistical meaning (Hacking 1990).

infrastructures, Foucault's work establishes that the normal is an explicitly political object. The dominant concepts of normal, health, and disease—seen more broadly as forms of knowledge and truth—are established through power relations within a variety of institutional structures, and at a variety of levels and scales (Rajan 2006). For example, as Foucault explores in *The Birth of the Clinic* (2003), the clinical gaze emerged as a way to propagate particular visions of the normal and the pathological within tissues, bodies, and populations. As biomedical institutions became sites of biological and social control, power and knowledge were vested in the processes of "life itself" (Rose 2007), as illnesses—and consequently the people who experience them—became defined and reduced to the binary of the normal and abnormal.

Ultimately, this dissertation draws on the combined work of Canguilhem, Hacking, and Foucault to explore how health and disease are negotiated—largely as the normal and abnormal—in the metabonomics laboratory. Taken together, these works establish how the seemingly objective categories of normal/abnormal and health/disease are imbued with moral, sociocultural, and political qualities, and are linked to the practices and power-structures of metabonomics research. This dissertation thus examines how statistical notions of normal and abnormal are enacted in relation to the post-genomic practices of the metabonomics laboratory, as well as in relation to the translational and personalized medicine (see Chapters 5 and 6) research agendas in which they are embedded. In particular, this dissertation explores the ways that metabonomics practices—and the wider context in which they operate—influence the production, use, and meaning of concepts of health and disease. In addition, it explores how metabonomics' concepts of normal and health are entangled with and influenced by the notions of complexity and statistical thinking that run through the practices of the laboratory.

Methods: Researching Metabonomics in the Biomolecular Medicine Laboratory (BMM)

This dissertation is based on one year of ethnographic research between October 2010 and September 2011 in the Biomolecular Medicine Laboratory (termed throughout this dissertation as "the BMM") at Imperial College London. The BMM comprises more than 75 students, career academics, and staff (Imperial College London 2011b), and is headed by Professor Elaine Holmes, the PhD student of the former head and founder of the BMM Professor Jeremy Nicholson. It occupies a place within the Division of Surgery and Cancer,

a “super-department” of more than 1000 researchers and staff, which has Jeremy Nicholson as its Head. To understand the “laboratory life” of metabonomics research at the BMM, I carried out laboratory research, conducted interviews, read the key papers and theses from the laboratory, attended training sessions and seminars, co-wrote a review article with a prominent laboratory figure, and coordinated a weekly discussion group on the application of metabonomics research to clinical issues. Consequently, this dissertation is as much an ethnography of the BMM as it is an ethnography of the field of metabonomics. The two are intertwined historically and conceptually (see Chapter 2), such that many of my observations about the BMM shed light onto the practices of metabonomics—and the practices of the post-genomic sciences—more broadly.

The research in this dissertation builds out of an effort to engage both critically and productively with emerging research in the biomedical sciences. Before coming to Oxford, I immersed myself in and defined my intellectual passions in relation to science. In 2008 I graduated from the University of Chicago with an undergraduate degree in biological sciences, which included many years practicing the scientific research about which I am currently writing. In total, I spent four years working in a laboratory conducting immunology research, an experience which encouraged me to approach ideas with quantitative logics and tools, and which also gave me a level of comfort and familiarity with laboratory practices and techniques. After beginning a research degree in Anthropology and training to be a social scientist, I developed different ways of approaching and engaging with the biological world. Ultimately the juxtaposition of these two disparate fields has made me increasingly self-reflexive about my research and writing, and shaped my choice of topics and methods for my doctoral research.

Despite my familiarity with biological research, conducting ethnographic research on the processes through which scientists create, make sense of, and use biological data was a challenging affair. Throughout my experiences conducting immunological research, I carried out the “wet” practices of biology, and had minimal engagement with complex data practices. As I discuss in Chapter 3, my engagement with statistics was limited to its use as a confirmatory tool: biological numbers spoke for themselves, and statistical tests were used to validate and add certainty to results. In contrast, throughout my fieldwork, the physical aspects of experimentation were largely seen as a nuisance, and the majority of time and effort was devoted to the experimentation and analysis of data via computers. To this end, my fieldwork required me to undertake the challenge of learning to speak the language and

carry out the practices of statistical data analysis. Methodologically, this proved challenging because it required participant observation of scientists working at computers, where much of the work is silent, embodied, and cognitive. But conceptually, this proved even more challenging because it required me to engage with data in a multivariate—and fundamentally different—way of thinking about biology and life.

Acknowledging that broad-scale changes are occurring in the biomedical sciences, and that emerging research is using "increasingly integrated, sophisticated and empirically productive approaches" (Parkin et al. 2007:215), my work views biomedical research as a source of productive critique and thought stimulation for anthropological issues and perspectives. My approach to metabonomics attempts to avoid the delegitimation of scientific work, as is sometimes seen in the "social constructionist" approach to laboratory studies (Latour et al. 1986). Instead, it seeks to critically explore the practices and meanings of metabonomics research by, on the one hand, questioning the production and use of facts, but on the other hand, trying to engage with and learn from modes of scientific research that might be changing the way we think about the world.

I came to the BMM somewhat accidentally, after approaching its head—Professor Jeremy Nicholson—about the possibility of exploring the laboratory's collaborative efforts to develop "personalized nutrition" with the food company Nestlé. After meeting a doctoral student in the laboratory at a personalized nutrition conference, I developed the idea to ethnographically track the research efforts of the "Nestlé-ICL Research Alliance," a collaborative research project on personalized nutrition (Imperial College London 2012c), at both Imperial College London and the Nestlé Research Centre in Lausanne, Switzerland. While Jeremy Nicholson welcomed me into the BMM—lightly dismissing issues of intellectual property and privacy with the statement "This is Britain, we are gentlemen!"—Nestlé was not as interested in allowing an anthropologist among their ranks. One month into my fieldwork, Nestlé invited me to their facilities in Lausanne on an all-expenses-paid trip. Following a whirlwind afternoon of meetings and facility tours, Nestlé politely declined my offer to collaborate with an email that read: "Although your proposal is scientifically interesting, expanding this area with a substantial program and extra-resources is not possible from our side due to other scientific focus and project prioritization." Consequently, I decided to focus the entirety of my ethnography on the research activities of the BMM, and on their efforts to negotiate the practices and ideas contained within the laboratory's speciality: the field of metabonomics.

The BMM is technically a laboratory, in the sense that it carries out biomedical research in a controlled set of conditions. The term "laboratory," however, implies a unity and coherence that is absent in the daily practice of the BMM, which encompasses multiple research locations, experimental practices, disciplines, and interests. Throughout my fieldwork I interacted with researchers between the ages of 20 and 70 years, who came from both academic and commercial backgrounds. I spoke to people trained in the fields of genetics, nutrition, statistics, computer science, computational chemistry, structural chemistry, veterinary medicine, bowel surgery, vascular surgery, and epidemiology. Although I spoke to people about a wide range of technologies and practices, this ethnography focuses on those that surround NMR as a biochemical technique and multivariate statistics as a range of data analysis methods. While doing so omits those practices and histories that surround MS, a technology which is becoming increasingly integral to metabonomics experiments within and beyond the BMM, it allows me to focus on those experimental practices I directly observed during my fieldwork.

Within the heterogeneous environment of the BMM, my ethnography took me to four floors within the Sir Alexander Fleming (SAF) building on Imperial's South Kensington Campus, and also to a shared experimental facility in an adjacent building. Because of the BMM's ties to clinicians and clinical research projects, my fieldwork also took me to the St. Mary's and Charing Cross Hospitals, both of which are part of the Imperial College research network. Furthermore, because of the BMM's ties to commercial entities such as Bruker Biospin and Waters Corporation, and also to companies such as Nestlé, AstraZeneca, Unilever, and Servier, my research took me to locales such as the Nutrigenomics Organization annual meeting in Glasgow, Scotland, as well as the Nestlé Research Centre in Lausanne, Switzerland. The BMM's ties to laboraties across the world—in the US, Germany, Switzerland, China, the Netherlands, Sweden, Spain, to name a few—also took me to a dense network of published research papers, which span a vast range of topics and disciplines in molecular biology, systems biology, medicine, computer science, and statistics.

The BMM, because of its loose, complicated, and spread-out structure, is not an easily-defined or "bounded" entity. As I have indicated, the practices that make up metabonomics research and the daily activities of the BMM are spread across multiple people, places, and realms of knowledge. This is not to say that my ethnography of metabonomics at the BMM is "multi-sited," but rather that it required me to trace biomedical phenomena across multiple arenas and sites of practice (Marcus 1995). My fieldwork at the

BMM therefore took on a unique character for two reasons. The first was due to the institutional, interconnected, and multiple nature of my fieldsite, which encompassed the activities of a wide variety of scientists, subjects, and scientific practices. The second was due to the fact that my fieldwork involved a constant movement back and forth between my home in Oxford and my fieldsite in London, as I commuted several times per week on the train. This created a "a constant shifting of positions between situations, people, identities and perspectives" (Amit 2000:11). This challenged my ability to find my place or identity within my fieldsite, but also provided a space and time for daily insight and reflexivity. Each day, as I travelled to and from my fieldsite, I gained the physical and mental separation from my ethnography to allow me to reflect on the meaning of my daily experiences.

Carrying out research at a high-powered institution that flourished because of its ability to network and make global ties placed particular demands and requirements on my activities, and affected the shape and character of the narrative that this dissertation provides. Work in the institutional setting of the BMM was fast paced and, at times, stressful. Researchers worked to tight deadlines for funding and publications, and often spent long hours in the laboratory trying to finish experiments or write up results for thesis chapters or journal articles. Work in the BMM, therefore, involved what Laura Nader (1972) has referred to as "studying up." In working with elite professional groups, I was involved in an inversion of the usual power relations, in which the people who anthropologists study are in positions of lesser power (Franklin 2002:354). As anthropologist Diana Forsythe has noted, "the relocation of fieldwork and fieldworkers to powerful institutions in this society has major implications for the conditions under which field research takes places and the kinds of relationships that develop between anthropologists and their informants" (Forsythe 1999:6).

Within the BMM, my relationships with researchers—and consequently my access to ethnographic data within a setting of complex power dynamics and hierarchical structures—varied. I gained relatively easy access to the work of graduate students, who were happy to let me help with their experiments or tag along to social outings to nearby pubs or movie theatres. With higher-level academics it was easy to conduct semi-formal interviews, but gaining access to daily research practices—which included closed group meetings or data analysis brainstorming sessions—was more difficult. This was not due to a lack of trust, but rather to the time and work pressures placed on higher-level academics. In an elite research institution such as Imperial, interacting with an anthropology doctoral student was, understandably, not always a priority.

Given the challenges associated with studying up, I had to come up with innovative strategies to gain acceptance and trust, and to become integrated into the world of metabonomics research. For a start, my background as a scientist facilitated my immersion into the laboratory, as those people I worked with recognized my skill in, comfort with, and passion for conducting scientific research. Metabonomics researchers frequently asked me, for example, if I had considered pursuing a degree in metabonomics after finishing my DPhil in Anthropology. While at times I was tempted to go "epistemically native" (M'charek 2005:175) and to prioritize laboratory work over ethnographic observations, my nativeness also enabled me to become truly immersed in the laboratory work, and gave me deeper access into the lifeworlds and perspectives of the researchers with which I worked (Löwy 1996:22). It enabled me to move beyond the surface-level meanings of science, and to understand the less apparent complexities in research practices and concepts.

Faced by these experiences and challenges, I made the decision early on in my fieldwork not to maintain the supposed "separation" between my life and the lifeworld of my informants. To become more immersed in the laboratory, I not only observed laboratory practices and participated in metabonomics experiments, but also brought my world of anthropology into the laboratory. I openly discussed the contrast between social science and metabonomics research, prompting discussions of methodologies, and giving my key informants texts like *Laboratory Life* (Latour et al. 1986). Overall, by exposing anthropology as a discipline—and therefore my own methods and questions—to the researchers at the BMM, I was able to learn how my informants viewed their own research, and whether they viewed my research as interesting and valid[16].

Ultimately, this dissertation is an inherently partial view of metabonomics and of the BMM (Helmreich 2000:26). The pictures I present of metabonomics research practices are always relative to a particular context, and to a particular constellation of people, objects, histories, and ideologies. As such, the objects of study in the dissertation have not been "discovered," but have rather been "laboriously constructed, prised apart from all the other

[16] Upon sharing part of my dissertation with one of my informants, I received the following reply via email:

> Read your chapter and enjoyed it. Probably not what you intended, but I think it is a great illustration of challenges in thinking statistically- you cover everything from the problem of sampling to confounding variables to the reductivity of models (be those 1s and 0s in regression or referring to a range of signs and symptoms as one "disease"). A great discussion on the challenges of practicing quantitative research in the real world.

possibilities for contextualization," and "inescapably shaped by the conceptual, professional, financial and relational opportunities and resources available to [me] the ethnographer" (Amit 2000:6). The methods that I have used to carry out this ethnography have intervened into the world of metabonomics research in the BMM—sometimes subtly, sometimes explicitly—to enact my objects of study (M'charek 2005:172). In this way, my portrayal of scientific practice focuses as much on metabonomics as a field as on the BMM as a specific laboratory.

Consequently, the people who are represented in this ethnography are also enacted and partially represented. Though it has been my intention to respectfully portray the informants with whom I worked, in writing about the metabonomics laboratory and its practices, this has meant engaging with and questioning issues of representation. Most notably, this dissertation is based around a series of vignettes of metabonomics research, which present in-depth examinations of practices and values. These vignettes are not singular instances, but rather have been deliberatively chosen because they are representative of metabonomics research more broadly. In addition, I have taken the liberty to interpret what has been said to me in interviews or informal dialogues, and I have made suggestions about unexpected or surprising connections that are not always obvious to researchers themselves. By attempting to uncover such "partial truths" (Code 2000), this dissertation includes some information that is critical of metabonomics research, or about which metabonomics researchers would likely disagree. However, I hope that my work does not come across as an attack on metabonomics research, but rather as an exploration of the many challenges faced by the field and its researchers.

Amidst the challenges of representing the characters and practices of ethnographic research, I have necessarily encountered issues of ethical transparency and confidentiality. The majority of my ethnographic and interview data was collected with verbal consent[17]: with the understanding that people were speaking to me in confidence, and with the stipulation that names would be anonymized. However, in the writing of my dissertation, it became apparent that it would not always be possible to anonymize the identities of the biomedical researchers whose work is discussed throughout this dissertation. This is true in particular in my discussion of the evolving history of the field throughout Chapter 2, which

[17] My initial plan was to carry out interviews with written consent, as is common to much ethnographic research. But when I presented metabonomics researchers with sheets of paper that they were required to sign before an interview, they reacted very negatively and became worried about the interview. Subsequently, when I made the decision to switch to verbal consent, researchers were much more relaxed and willing to collaborate.

features descriptions of the roles played and research conducted by specific individuals. Omitting the names and citations of such individuals would render the history of metabonomics impersonal, and would remove the agency and constant negotiation of contingency that characterizes the ongoing project of metabonomics research. Citations in-and-of-themselves are a form of ethnographic data, without which my account of metabonomics would lack historical and epistemological specificity.

Consequently, it is my view that the specific and local characteristics of research—which associate scientific facts with particular individuals, technologies, and ideologies—are integral to my account of metabonomics. Given these challenges, I made the decision—following Stefan Helmreich's approach in his study of artificial life researchers (Helmreich 2000:26)—to give all researchers whose work appears in this dissertation the chance to choose whether their names and works would be anonymized. For the reasons discussed previously, I have largely identified those individuals who have established and enduring roles in metabonomics research, and have given pseudonyms to those individuals who have more temporary roles[18]. While doing so runs the risk of perpetuating academic hierarchies and power structures, I would like to argue, instead, that it identifies those individuals who are central to the emerging narrative of metabonomics, while maintaining anonymity for those individuals who provide a portrait into the everyday practices of the BMM. Thus, for those researchers whose names appear publically, I have done my best to share my writing with them, and to check that they are happy with their portrayals or with the accuracy of their words. Within the text, pseudonyms appear as first names only, while true names appear as first and last names.

Overview of the Dissertation

In this dissertation, I explore the daily practices of the metabonomics laboratory, focusing on how biochemical and statistical practices engage with and shape the dynamic concept of metabolism in efforts to define health and disease. In **Chapter 2**, I discuss the context and history of research within the Biomolecular Medicine Laboratory (BMM) and within the field of metabonomics. I discuss how the academic setting of Imperial College London, particular biochemical and statistical research practices, and groups of people

[18] More often than not, researchers at all levels within the BMM's hierarchy were happy to associate their names with metabonomics research and narratives. This is not altogether unsurprising, as it reflects the increasing trend for scientific research to be a matter of personal ownership, attribution, and pride.

constitute metabonomics research at the BMM as an "epistemic culture." From there, I provide a historical overview of the formation of metabonomics as a field, detailing its relationship to NMR and multivariate statistical techniques. Subsequently, I provide an overview of the historical foundations of the BMM, discussing how its research has been characterized by two periods, one focusing on drug toxicity and the other focusing on translational and personalized medicine.

Building on my exploration of the historical and technological roots of biochemical and statistical practices, in **Chapter 3** I explore how multivariate statistics have become conceptually central to metabonomics' visions of the complexity of metabolism. I provide a brief history of metabolism as it is conceptualized and rendered into a contemporary form by metabonomics researchers, and consequently describe how metabonomics uses biomarker panels—made of combinations of biochemicals that are "discovered" through the application of multivariate statistics to biochemical information—to conceptualize health and disease. From there, I discuss how multivariate statistics are naturalized as the obvious and correct way to engage with metabolic data, and consequently how multivariate statistical methods are integral to the enactment of molecular complexity. Subsequently, I suggest that biological meanings are entangled with technological practices, such that metabonomics researchers both draw upon and produce particular and limited notions of molecular complexity in their everyday work.

Extending the challenges inherent in applying complex and statistical thinking to notions of metabolism, in **Chapter 4** I explore how researchers grapple with the dynamic states of health and disease with the unstable concepts of "normal" and "abnormal." To begin, I discuss how the normal and abnormal are enacted in the metabonomics laboratory relative to particular technologies, practices, ideas and values. I examine how statistical concepts of the normal emerge as "objective" ways of thinking about biology, despite the fact that metabonomics practices render normal and abnormal as inherently contingent entities. From there, I explore how metabonomics researchers engage with the instability of the normal and abnormal, by examining the strategies researchers use to grapple with the multiple and contingent nature of disease objects. Subsequently, I suggest that metabonomics, in enacting statistical concepts of the normal, struggles to engage with the dynamic, vital, and contextual meaning of health.

Building upon researchers' invocations that objectivity with disease diagnosis and treatment can be achieved with molecular and statistical technologies, in **Chapter 5** I discuss the challenges that arise when metabonomics researchers attempt to "translate" laboratory research into clinical applications. I highlight the centrality of informational practices to metabonomics, and show these practices, in contrast to histopathology, enact biological entities as statistical patterns. I then discuss how metabonomics researchers place value on the collection and analysis of data—and in particular multivariate statistical forms—as the key to disease diagnosis and treatment. From there, I discuss the challenges that arise in the interpretation of such data, arguing that researchers struggle to translate laboratory findings into biological processes and outcomes. Subsequently, I suggest that researchers' valuations and technological visions of data struggle to overcome the fundamental role that human judgment and interpretation play in clinical practice.

Following such claims that multivariate statistics can help to improve and standardize medical care, in **Chapter 6** I examine how metabonomics research is mobilized to investigate the emerging concept of "personalized medicine." I show how personalized medicine, despite a rhetorical focus on the individual, draws upon and co-produces information about both individuals and populations. To do so, I explore the BMM's efforts to develop the concepts of "pharmacometabonomics" and "molecular epidemiology" in newly established state-of-the-art facilities. From there, I focus on research on acute-on-chronic liver failure to highlight how increasingly statistical and "personalized" methods for diagnosing disease conflict with individually-tailored forms of clinical practice. Subsequently, I suggest that the statistical basis of the BMM's efforts to develop personalized medicine does not enhance the treatment of individual patients, but instead focuses on the probabilistic treatment of populations.

Chapter 2: Metabonomics in Historical Context

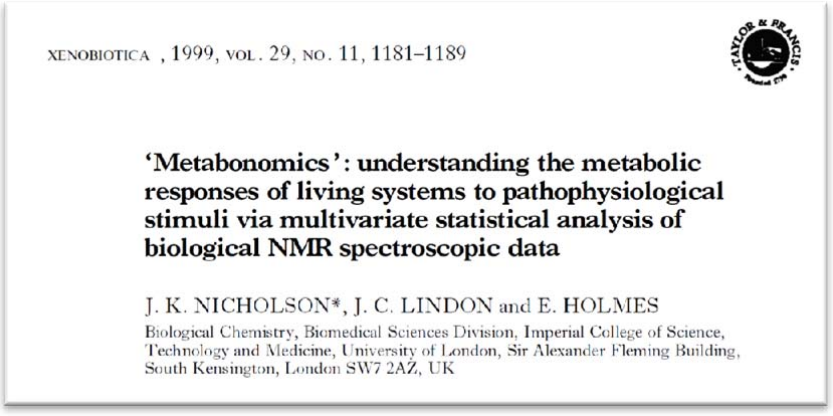
Introduction

I am interviewing Professor John Lindon, one of the founding members of the BMM, in his third-floor office. A kindly looking gentleman in his late sixties, John Lindon is now semi-retired after a successful research career in both industry and academia. At the BMM, he spends most of his time mentoring graduate students, and is widely regarded as the laboratory's expert in nuclear magnetic resonance (NMR) spectroscopy, a technology that is used to detect the biochemical composition of samples. Today, I am speaking with John Lindon because he has devoted a significant amount of time in the later stages of his career to chronicling the history of the field of metabonomics at the BMM. He has published this history on the BMM's website (Lindon 2010), but I want to speak to him in person to hear how this published account came to exist.

During our interview, John Lindon describes how the term and field of "metabonomics" came about. Though it entered into official scientific narratives in 1999, the term was first coined, says John Lindon, in 1996 at a brainstorming meeting at Birkbeck College that included Jeremy Nicholson, another founding member of the BMM. The term was derived in response to the growth of biomedical fields like genomics in the 1990s, as researchers sought to place a name—and official stamp—on their field of work. It was an attempt to describe the science that focused on the "metabolome," the sum total of the biochemicals that make up the metabolism of an organism, rather than the more well-known "genome," the sum total of genes. John Lindon says:

> We've been doing what we call "metabonomics" since, now, say, 1988. We never had a name for it, we didn't think there needed to be a name for it. But in about, well I'd say it was the mid-1990s—1995—it became obvious that people started talking about genomics. And we saw that there needed to be a name for this.

Though a literature search of the term "metabonomics" yields no results before 1999, the building blocks of metabonomics, as John Lindon stresses, had already existed as a series of practices, technologies, collaborations, and ideologies for decades. In particular, metabonomics drew on pattern recognition technologies, early forms of multivariate statistics that emerged in the field of econometrics in the 1950s, and were subsequently used in drug discovery research in the 1980s and 1990s. In contemporary times, these statistical technologies encapsulate the data practices that metabonomics researchers use to interpret information and make sense of metabolism, and form a central focus of this dissertation. In addition, metabonomics drew on NMR spectroscopy, a technology originally used in physics to detect the physical composition and structure of compounds, and which is now used in metabonomics to determine the biochemical composition of liquid and solid samples.



XENOBIOTICA , 1999, VOL. 29, NO. 11, 1181–1189

'Metabonomics': understanding the metabolic responses of living systems to pathophysiological stimuli via multivariate statistical analysis of biological NMR spectroscopic data

J. K. NICHOLSON*, J. C. LINDON and E. HOLMES

Biological Chemistry, Biomedical Sciences Division, Imperial College of Science, Technology and Medicine, University of London, Sir Alexander Fleming Building, South Kensington, London SW7 2AZ, UK

Figure 7: 1999 *Xenobiotica* paper that provides the original definition of metabonomics.

Though John Lindon's historicized account of metabonomics emphasizes that the field has been practiced longer than its name would suggest, it is also a way of performing the field as something strong and enduring. Metabonomics is concerned with establishing its place and role within the history of biomedical research, as it competes with other fields of post-genomic research to secure recognition and funding. John Lindon continues on:

> A lot of people think metabonomics started in 2001, because, you know, the papers started appearing in 2001. They don't realize that there was a huge history before that, [on] a much smaller scale, but the thinking was the same. The technology might have been more primitive, and the numbers might have been lower, but the thinking essentially hasn't changed.

After informally defining the term "metabonomics" at their 1996 meeting, the researchers published an official definition in the journal *Xenobiotica* in 1999, in an article titled "'Metabonomics': understanding the metabolic responses of living systems to pathophysiological stimuli via multivariate statistical analysis of biological NMR spectroscopic data" (Nicholson et al. 1999) (see Figure 7). The article—which as of the writing of this thesis has been cited 1925 times —outlined the working definition of the emerging field of metabonomics as: "The quantitative measurement of the dynamic multiparametric metabolic response of living systems to pathophysiological stimuli or genetic modification" (Nicholson et al. 1999:1181).

Describing the origins of the term "metabonomics," John Lindon says:

> We were having a discussion one afternoon, and we ended up with the term "metabonomics" to describe what we were doing. And we had, sort of, various ways of getting to metabonomics. There were all sorts of names thrashed about… But we thought of "economics," and it had to be "omics." And so metabonomics, we thought. And we didn't come up with "metabolomics" like "metabolite," just "metabonomics." And I think afterwards we rationalized it like genomics and economics, with the "n." And so we coined the term "metabonomics," and we started using it in about '96. But we never published it.

Metabonomics' use of the suffix "omics," as John Lindon comments, is a testament to the field's entanglement with the rise of genomics and other post-genomic fields of study. Though the word "genome" was first used in 1920 by German botanist Hans Winkler, the suffix "omic" entered into popular use after the growth of the Human Genome Project (HGP) in the 1990s, when the concept of the "genome" gained a popular definition as the sum total of the genes in the body (Baker 27 February 2013). Consequently, metabonomics encapsulates the increased scale and molecular capabilities of post-genomic research, and is one of many attempts to broaden the study of the "gene" to include other parts and systems of the body. Though the suffix "omics" signals the fact that metabonomics follows in the wake and shadow of the HGP, I argue throughout this chapter that the field of metabonomics entails a distinct and specific history that has evolved separately from genomics. Metabonomics is tied to the technologies and practices of—but is also engaged in a constant project to distinguish itself from—genomics and other fields of post-genomic research.

John Lindon's discussion of the origins of the term "metabonomics" also signals the field's rivalry with "metabolomics," a similarly named field of post-genomic research on

metabolism (see Chapter 1). Metabolomics, he explains, was developed in parallel to metabonomics by a separate group of researchers—in particular Oliver Fiehn, who now leads a metabolomics laboratory at the University of California Davis—who were working on the application of the technology mass spectrometry (MS) to model organisms and plant systems (Lindon et al. 2007:ix). Though both metabonomics and metabolomics involve convergent methods and approaches (Lindon et al. 2007:3), in our interview, John Lindon emphasizes their differences. Metabolomics is the study of endogenous metabolism, and seeks to determine the metabolites that make up cells or tissues. Metabonomics, in contrast, is the study of exogenous metabolism, and seeks to understand how metabolism changes in response to perturbations such as environment or disease, rather than to identify metabolites (Small Things Considered: The Microbe Blog 2009).

As John Lindon emphasizes, metabolomics provides *an analytical rather than a biological definition*: it involves the quantitative characterization of metabolism, rather than an understanding of the way metabolism changes in relation to health and disease (Lindon et al. 2003). By emphasizing their ties to applied biomedical research, metabonomics researchers place value on their field's application to biological outcomes and clinical issues of health and disease. John Lindon's words signal, in other words, the ways in which metabonomics researchers derive scientific prestige from associating their research—with its roots in chemistry and statistics—with biological questions and applications. As Park Doing (2009) argues in his account of the changing research agenda of a synchrotron laboratory, scientific prestige and the biological focus of research are intertwined.

When I ask John Lindon about the difference between metabonomics and metabolomics, he emphasizes that the rivalry between the two fields is mostly semantic. Metabonomics researchers write that the distinction is more philosophical than technical, and in agreement John Lindon says: "We're cool about it, what will happen will happen." In practice and in the literature, both terms are used interchangeably. John Lindon describes:

> I think we could take the view that the word metabolome is a valid word, the metabolic complement of a single cell. Now there are thousands of metabolomes that you study when you study metabonomics, does that make sense? So that's sort of our take, I might be the Imperial party line. So we use both terms, and if you don't want to offend anyone you call it "metabolic profiling."

Although researchers in the BMM are not outwardly concerned with the distinction between the two fields, their continued use of the word metabonomics—as well as their reference to its unique history—signals the "boundary work" (Gieryn 1983; Löwy 1990) that

metabonomics researchers do to assert the prestige and originality of their field. Thus, as researchers attempt to increase the research capacity and reputation of metabonomics, they retrospectively construct a narrative about the origins, context, and history of their field. Doing so conceals how metabonomics, rather than being a naturalized object, is the result of active efforts to stabilize a constantly changing field. In the end, such boundary work signals the constant negotiations of metabonomics researchers about what their field is and will come to be.

Overall, this chapter is based on a combination of my first-hand observations as an anthropological participant-observer, as well as second-hand accounts of the history and practices of the laboratory. I learned a great deal about BMM—its practices and setting—through my ethnographic fieldwork, and through interviews with founding members John Lindon and Jeremy Nicholson. Beyond the information gathered in interviews, many of the details, dates, and names that are presented in this chapter are based on my informal interactions with researchers who worked within or were affiliated with the BMM. This information, which constitutes this "oral history" of metabonomics research at the BMM, has been cross-checked against seminal publications in the field, which I cite throughout the text. These publications provide reference to specific dates and people, and show how metabonomics researchers portray their field to the public eye.

Before beginning my overview of the history and research of the BMM, I want to clarify that my exploration of oral histories and narratives is not a search for the "truth" about metabonomics as a field. Instead, it is an examination of how people talk about the past, in order to better understand the current configuration of metabonomics research. By analyzing the dominant narratives that characterize the evolving history of metabonomics research at the BMM, this chapter explores the ways in which *history is actively being made, and boundaries are actively being drawn around a discipline before people are fully aware of what it contains.* Thus, following in the footsteps of Karin Knorr-Cetina's (1999) ethnography of high energy physics and molecular biology, I explore how metabonomics research has come to encompass a particular "epistemic culture," which is made up of practices of knowledge production that are interwoven with various histories, locations, and 20th century developments. I also explore, as Park Doing (2009) has done in his ethnography of the synchrotron laboratory, how changes in the nature of metabonomics research reveal the ways in which disciplines and facts become established.

Consequently, this chapter illuminates the shifting and evolving nature of metabonomics over time, by showing how it has emerged as a particular configuration of biochemical technologies, statistical practices, and research ideologies. To this end, the first section of this chapter provides an overview of the research, technologies, and people that make up the BMM at Imperial College London. After discussing the present context and setting for my fieldwork, the second section details metabonomics' historical ties to the technologies of multivariate statistics and NMR. The third section details how after the founding of the BMM at Imperial College London in 1998, the research agenda and practices of the BMM were shaped by its efforts to collaborate with private companies, and also by Jeremy Nicholson's promotion to the head of the Department of Surgery and Cancer in 2009. In conclusion, this chapter examines the rhetorical, political-economic, and cultural factors involved in the establishment of metabonomics as a biochemical and statistical field.

Navigating the Biomolecular Medicine Laboratory (BMM)

Many months before meeting John Lindon, I enter the fourth floor of the Sir Alexander Fleming (SAF) building on a rainy autumn day. As I step out of the elevator, I enter into a giant, multi-floor atrium. The building, which is only a decade old, has been cleverly designed to let light in through the ceiling of windows, and the effect is welcome on the grey October day. To my right are a group of graduate students sitting around a long table, chatting and laughing during their lunch break. To my left, and throughout the large atrium area, are rows of computer workstations, housing primarily graduate students, but also a few recently graduated post-doctoral researchers. These rows of desks form a circular ring around a large, open central space, which spans four stories and is flanked by small laboratory spaces and meeting rooms along its outer edges. The physical space of the BMM is a testament to the social hierarchies that run through the laboratory, and also to the role that certain technologies and practices play in metabonomics experiments.

The BMM is located within Imperial College London's main campus in South Kensington (see Figure 8), which is flanked by other campuses in Chelsea, Hammersmith, and Paddington. These campuses encompass Imperial's four main academic units: the Faculty of Natural Sciences, the Faculty of Medicine, the Faculty of Engineering, and the Business School. As a public research university, Imperial has more than 13,000 students and 3,000 faculty and staff, who specialize in research in the fields of science, engineering, medicine, and business.

Figure 8: Entrance to the SAF building in South Kensington, London.

Imperials' reputation in the international academic community is that of science and engineering prowess. It has 14 Nobel Prize winners in its history, and consistently falls within the top ten of the World University Rankings, which are published each year by the Times Higher Education (Times Higher Education 2012). Amidst other well-known London universities with a science focus—such as University College London and King's College London—Imperial promotes itself as being on the cutting edge of scientific research, and as having strong ties to industry, commerce, and healthcare (Imperial College London 2013). Consequently, Imperial conducts research and training at multiple hospitals throughout Greater London, which include St. Mary's Hospital, Charing Cross Hospital, Northwick Park Hospital, St. Mark's Hospital, and Hammersmith Hospital. All of these are managed by the Imperial College Healthcare National Healthcare Service (NHS) trust, the second largest NHS trust in England and one of the first academic health science centres in the United Kingdom[19].

Such ties to healthcare and industry are intricately linked to Imperial's research and funding climate, which entails large investments from the private and charitable sectors. Imperial has one of the smallest endowments—funding for operational or capital

[19] NHS trusts are public sector corporations—headed by a board of directors—that provide services on behalf of the NHS. The establishment of trusts occurred under the re-organization of the NHS at the turn of the 20th century: internal markets were introduced such that local health authorities were given management over their own budgets and activities.

investments—of the major global universities. In the year 2010/2011, for example, Imperial's total income was £705 million, of which £299 million was from research grants and contracts (Imperial College London 2011a)[20]. At the time of my doctoral research, this translated into a unique research climate, in which Imperial was known for encouraging a large number of public-private partnerships in order to gain access to government and industry funding. In the BMM, this not only raised money for research projects and scholarships, but also provided opportunities for patents and translational applications of the laboratory's research. Consequently, the high number of public-private partnerships gave way to Imperial's reputation—at least amongst those researchers with whom I worked—for being money-conscious, and for being concerned with publication outputs and grant numbers.

Figure 9: SAF building, home of the BMM.
The fourth-floor atrium of the SAF building in which the BMM is located, showing the multiple floors of desks and computers for graduate students. Note how the space is occupied by computers rather than the test tubes and pipettes associated with imaginaries of biomedical laboratories.

I walk around the atrium (see Figure 9), taking in the sights and sounds. It is filled with the constant tapping of keyboards, rustling of papers, and the low-level chatter of graduate students. Those students affiliated with the BMM are located primarily on the fourth floor of the atrium, while researchers from other labs are scattered across the other three floors. Many of the BMM's graduate students will have arrived at the BMM at around

[20] In contrast, the University of Oxford had a total income of £880 million, of which £367 million was from research grants and contracts. But while Oxford's endowment was £3,900 million, Imperial's endowment was only £75.6 million (Coughlan 26 September 2011; Wikipedia 2013).

10am, but will stay until 8pm or later. As they work, they stare at the large screens of desktop computers, many wearing headphones, others sipping coffee or eating lunch at their desks. As is typical of graduate students in most scientific fields, they spend long and irregular hours at the BMM, eating meals and taking breaks with fellow laboratory members.

Figure 10: Entrance to the sixth floor NMR laboratory of the BMM.
The entrance to the sixth floor laboratory of the BMM. Note warning signs for strong magnets from NMR spectrometers.

Leaving the atrium, I wind my way up to the sixth floor of the SAF building. When I arrive at three large elevators, I come face-to-face with the entrance to the sixth floor laboratory space of the BMM, the area where many of the NMR and MS machines are housed[21]. The door at the entrance has several laminated signs tacked onto it, all in bold capital letters, that read "WARNING: NUCLEAR MAGNETIC RESONANCE SUITE," "NO UNAUTHORIZED ENTRY!!!" and "NO MOVING OR METAL OBJECTS IN THIS LAB!" (see Figure 10). Slightly intimidated, I use my card to swipe through the door, and enter into the BMM's main NMR facility.

With access to this space, I have gained official approval to work as a "visiting researcher," in the form of a letter signed by Jeremy Nicholson and a picture ID. I have also

[21] In addition the analytical facility on the sixth floor of the SAF Building, the BMM also shared access to the "Cross-Faculty Centre for NMR facility" in an adjacent Chemistry building, which housed a high-resolution 800 MHz NMR spectrometer. This building also housed the Waters Laboratory for Mass Spectrometry, which contained a variety of MS machines manufactured by the Waters Corporation. As of the writing of this dissertation, an NMR facility was also being installed in St. Mary's Hospital in the "Clinical Phenome Centre."

completed a "safety induction," a one hour tour of the laboratory's facilities, which gives a nod to its various dangers. I have been sternly told to maintain all of the laboratory equipment in good order, and to help keep the BMM's six NMR spectrometers up and running. I am told not to travel in the freight elevators when carrying large dewars of liquid nitrogen—which is required to keep the superconducting magnets inside of the spectrometers running—as a precaution in the case of a leak. I am also told to wear safety goggles when filling the NMR spectrometers with liquid nitrogen, in case the metal tubes freeze and explode. In addition, I am told not bring any electronics inside of the yellow and black striped line that signals each NMR spectrometer's magnetic field. This, researchers in the BMM emphasize, will wipe any magnetic information off of credit cards or informational storage devices.

The sixth floor space that unfolds before me has a large, open floor plan, which is a combination of desks and machines (see Figure 11). It is inundated by bright fluorescent lighting, and is peppered with color-coded signs for health and safety hazards, as well as research posters spanning the laboratory's work over the past decade. Recently published papers[22]—all of which are co-authored with the head of the laboratory Jeremy Nicholson—and cut-outs from magazines of media attention have also been plastered to the walls, enlivening the laboratory's sterile white colour scheme. The space hums with the constant ticking and white noise of machinery, which is interspersed with the voices of research fellows, post-doctoral researchers, and staff. This setting is more familiar to me than the fourth floor atrium: it has the feeling of being a "wet" laboratory, in which scientists manipulate physical samples and conduct experiments at machines.

As I continue walking through the laboratory, passing several administrative offices and an impressively large colour copier, I arrive at the flagship area of the laboratory. This space, which is the size of a large cafeteria, is peppered with six-foot tall NMR spectrometers, all of which are flanked by computers, tubing and wires, metal boxes full of machinery, and dewars of liquid nitrogen. Each machine is encircled by a 10 foot diameter yellow and black striped line, which signifies the powerful magnetic fields in which the NMR

[22] Many of the BMM's papers were published in the *Journal of Proteome Research* (JPR), which is published by the American Chemical Society. Founded in 2002, JPR publishes papers on protein analysis and function, and also on topics relating to the post-genomic sciences. Researchers referred to JPR as the laboratory's "in-house" journal, in the sense that it was the go-to journal—which had a high acceptance rate among metabonomics researchers—for publishing the results of studies.

spectrometers operate. Interspersed throughout the space, there are also several MS machines, which appear as box-like apparatuses that are connected to a variety of bottles filled with clear fluid. To my right, there is also a small "prep lab" where biological samples are prepared at benches or in sterile hoods to prevent contamination.

Figure 11: Inside of the sixth floor NMR laboratory of the BMM.
The sixth floor laboratory, showing the various machines operating within the open floor plan. To the left is a computer console used to control NMR spectrometers, while to the right is a large, round NMR spectrometer, whose magnetic field is designated by the black and yellow markings.

Within the NMR laboratory, researchers walk around in plain clothes, largely without lab coats, goggles, or other protective wear. Because their work does not typically involve hazardous, biologically active, or sensitive materials, they do not need to take the same physical precautions as other biomedical researchers. Instead, researchers take steps to avoid the large magnetic fields generated by the NMR spectrometers, carefully checking that their pockets are free of cell phones, ipods, keys, or other metal objects before they conduct experiments. Moving around the laboratory, they shuffle back and forth between machines holding trays of long, thin test-tubes filled with diluted urine or serum. Some researchers are perched on ladders at the side of NMR spectrometers, loading samples into robotic autosamplers, while others are staring intently at computer screens, adjusting the parameters of NMR spectrometers as they set up experiments.

The People and Culture of the BMM

Retracing my steps back towards the entrance to the sixth floor laboratory, on my right I see a large space full of cubicles that is home to academic staff and post-doctoral researchers. They sit at their desks, staring intently at screens or pouring over printed manuscripts and research papers. Across the hall is the office of Professor Jeremy Nicholson—the unofficial "boss" of the laboratory—whose door is plastered with science cartoons and notable figures from recently published articles (see Figure 12). A large sign, in which is written "DO NOT ENTER WITHOUT ASKING" in bold capital letters, covers the window to his office. The sign, though it seems unfriendly, is instead a reflection of the busy life that Jeremy Nicholson leads. After being appointed the head of the Department of Surgery and Cancer in 2009, Jeremy Nicholson spends the majority of his time out of office, attending meetings or travelling to international conferences. Though he still supervises graduate students in the BMM, his time is precious, and it is not unusual to see him entering and leaving his office followed by a train of people trying to ask a quick question or schedule a future meeting.

Figure 12: Article from *Scientific American* taped to the wall outside of Jeremy Nicholson's office, featuring his research in the BMM.

Adjacent to Jeremy Nicholson's office is the office of Professor Elaine Holmes, the acting Head of the BMM. The former student of Jeremy Nicholson, Elaine Holmes took over

Jeremy Nicholson's role as head of the laboratory[23]. As the person responsible for the work of more than seventy people in the BMM, Elaine Holmes spends the majority of her time applying for grants, and reading and proofing research papers. When she is in the office, she is hurriedly running from one meeting to the next, often pushing her recently born child in a stroller down the hallway. Given the growth of the BMM, researchers often lamented that Jeremy Nicholson and Elaine Holmes struggled to keep up with the pace of work: they were too overworked to review the laboratory's research, and would often take weeks if not months to provide feedback on drafts of thesis chapters or publications.

The busy schedule of Jeremy Nicholson and Elaine Holmes is an indication not only of the research productivity of the BMM, but also of its pyramid-like administrative and power structure. The placement of researchers within this academic hierarchy is governed by academic merit and experience, but also by skills in networking and self-promotion. Working under Jeremy Nicholson and Elaine Holmes are upper-level academic staff, primarily Professors and Lecturers, who occupy a separate set of cubicles and offices on the third floor of the SAF building[24]. These researchers shape the direction and funding of the BMM's research: they are responsible for supervising students, teaching courses, and chairing project meetings, and also for procuring funding to support their own research projects. They have specific and distinct academic specialties—which include cancer, systems biology, surgery, parasitology, environmental toxicology, aging, bioinformatics, biochemistry, and pharmacology—and often work within larger interdisciplinary and inter-institutional projects. The general activities of the BMM are also supported by several staff administrators, who carry out the organizational aspects of the laboratory, such as booking conferences, making travel arrangements, organizing grant payments and reimbursements, and submitting orders for laboratory supplies.

---

[23] As with many biomedical laboratories, the BMM is headed by a principal investigator (PI), a senior researcher and leader who sets the research agenda and expectations of the laboratory. PIs are responsible for securing funding and grants, for recruiting students, and for publishing and promoting the results of the laboratory's research.

[24] The BMM also hosts several visiting academics—which include senior researchers from affiliated Universities and commercial entities—as well as visiting researchers—which consist of former members of the laboratory and clinicians conducting short-term research projects. These visiting academics and researchers conduct collaborative research with the BMM, providing samples, analyzing data, or consulting on large research projects.

Working under the academic staff are research fellows and post-doctoral researchers, a group of people who occupy the intermediary level between senior academics and graduate students. Working from desks and cubicles nestled in a corner of the sixth floor laboratory, some research fellows carry out independent research on multi-year grants, while others work on six-month and one-year extensions of their PhD research. Working at the lowest hierarchical level within the BMM are the graduate students and visiting researchers, who form the youngest and largest group of researchers. With their numbers hovering between twenty and thirty, graduate students work primarily from the fourth floor atrium and conduct experiments in the sixth floor laboratory. It was common for less than half of the graduate students to be physically working in the SAF building at any given time, as many preferred to analyze their data or write up their results from home. The majority of graduate students work on three-year funded studentships, which come from Government Research Councils, charitable bodies, overseas agencies, and pharmaceutical and food science companies (Imperial College London 2012a).

The BMM contains a diverse group of researchers, who come from a variety of intellectual and geographical backgrounds. Researchers encompass a range of ages—from early twenties to early seventies—as well as an array of academic backgrounds—including chemometrics (see Chapter 2), veterinary medicine, nutrition, biochemistry, and microbiology. Researchers in the BMM also comprise a large range of nationalities, including people who had lived or completed their previous education in Spain, the Netherlands, Sweden, France, Germany, Greece and Cyprus, China, Russia, Mexico and the United States. Most of these researchers speak English in professional settings, but talk colloquially with friends in native tongues. Consequently, the international character of the BMM's research reveals the ways in which the laboratory participates in a global community of academic research, as its boundaries extend beyond the confines of London to encompass informal research networks throughout the world.

Overall, the physical space and people within the BMM reflect the particular "culture" of metabonomics research that is practiced at Imperial College London. Knowledge and expertise are partitioned in physical spaces, such that the spatial and social hierarchies that I have described represent distinct groupings of researchers at different stages in their academic careers. Graduate students are placed in the crowded, open space of the fourth floor atrium, which is located away from the expensive machinery and equipment of the sixth floor NMR laboratory. Higher-level post-doctoral researchers are given larger

desks, faster computers, and larger monitors. Lecturers and professors are bestowed with private offices, and with them, the autonomy—as occurs in most academic settings—to carry out their own research projects.

Thus, the BMM, like other laboratories at premier academic institutions, is intertwined with "economies of privilege, education, and travel that connect elite cultures in Europe and the United States" (Helmreich 2000:34). Its location within Imperial College London gives rise to certain socioeconomic groupings of people, and affords metabonomics researchers access to a wide range of funding, research facilities, and collaborations. Within this political economy of science, the laboratory's research activities are fundamentally shaped by its collaborations with public and private corporations, which provide funding and resources—physically and in expertise—without which the laboratory could not operate. I explore this in more detail later in this chapter, as I discuss the BMM's ties to private pharmaceutical and nutrition companies, as well as to public health initiatives funded by the UK government.

Despite these cultural aspects of the BMM, however, I want to emphasize that the laboratory is neither monolithic nor a clearly bounded entity. Though metabonomics research occurs physically at the BMM, it also stretches across multiple laboratories and social spaces. Key research activities are carried out at the affiliated St. Mary's and Hammersmith Hospitals, and in collaboration with scientific groups at other research institutions within and beyond the UK. In addition, researchers use a variety of technologies and approaches to their work, which as I discuss throughout Chapter 3, entail a combination of institutionalized and creative practices.

The Foundations of Metabonomics

In this section, I provide an overview of the historical and technological foundations of the field of metabonomics, detailing how biochemical and statistical technologies became central to the study of metabolism. I explore how John Lindon's work on pattern recognition and multivariate statistics combined with Jeremy Nicholson's work on NMR to create the modern practices that we call "metabonomics." Though the history I provide here reveals the creative and hard work of John Lindon and Jeremy Nicholson, it also *omits* the work of other individuals. It presents, as I mentioned in the Introduction to this chapter, the historical narrative of the BMM and the field of metabonomics as it has been constructed by researchers. While it could be argued that focusing on the role of key players privileges the

voices of those in power and maintains academic hierarchies, doing so also reveals the active work that goes into stabilizing and crafting the history of the field. Thus, the aim of this section is not to give a voice to the less powerful or to write the field's history, but rather to examine the predominant historical narratives that researchers present in an attempt to establish and legitimate their field.

Pattern Recognition and Multivariate Statistics

The history of metabonomics, according to researchers in the BMM, began when John Lindon met Jeremy Nicholson in the mid-1980s in London. John Lindon had been working at the now closed Wellcome Research Laboratories[25], carrying out what was known as "rational drug design" for pharmaceutical development[26]. With this approach, researchers did not rely on trial-and-error testing of chemical substances, but rather worked backwards to design tailor-made biochemical compounds to work with known biological pathways or "drug targets" (Venture Navigator 2007). Researchers used biochemical techniques, including NMR, to synthesize compounds with desired physical and biological properties (Gwynne et al. 2007). But even more importantly, they relied on advanced computer modelling tools to make complex theoretical calculations about the physical properties of compounds, such as the size of molecules, charges on atoms, accessible area to water molecules, and dipole moments.

[25] The Wellcome Research Laboratories were founded by Sir Henry Wellcome, an American born pharmacist and philanthropist who lived between 1853 and 1936, and who made his fortune from selling medicine in tablet rather than power or liquid forms. Wellcome established the Physiological and Chemical Laboratories in the late 19th century, which were consolidated into the Wellcome Research Laboratories in 1946. Throughout the early 20th century, the Wellcome Laboratories played a key role in research on the diphtheria anti-toxin and histamine, and also supported wartime efforts to manufacture drugs of German origin. In 1924, Wellcome consolidated his enterprises and wealth into the Wellcome Foundational Limited. In 1932, shortly before his death, he signed the entire share capital of Wellcome Foundational Limited to the Wellcome Trust, stipulating that all profits of the Wellcome Foundation would be used by the Trust to advance medical research and the understanding of its history. As a result, the Wellcome Foundation operated as a profitable pharmaceutical company through the early 1990s, pioneering "rational drug design" in drug discovery research. The Wellcome Trust remained the sole shareholder of the Wellcome Foundation until the 1980s, until the Wellcome Foundation was taken over by Glaxo plc. in 1995. This resulted in the new company Glaxo Wellcome plc., which merged with SmithKlineBeecham in 2000 to create GlaxoSmithKline (GSK) (The Wellcome Trust 2013).

[26] "Rational drug design" occurred as pharmaceutical researchers began to accumulate a greater understanding of the structural and molecular properties of drug-target interactions, and was broadly classified as a type of synthetic organic chemistry (Hopkins et al. 2007). It gave way to technologies of genetic engineering and monoclonal antibodies, which allowed existing biochemicals to be produced at a greater scale and lesser cost, and consequently to industrialized technologies of high-throughput screening (HTS) and genomics (Nightingale 2000).

In this era of rational drug design, researchers at institutions like the Wellcome Research Laboratories attempted to calculate hundreds of properties per compound, because they believed it was the best way to understand the biological activities and toxicities of potential drugs (Xu et al. 2002; Gasteiger 2006). Doing so required state-of-the-art computing facilities, which in the early 1980s meant something truly different to what it does in the 21st century. As John Lindon recalled, researchers at the Wellcome Laboratory had begun to use Digital Equipment Corporation (DEC) "minicomputers," which replaced larger and centralized International Business Machines (IBM) mainframes. Such minicomputers were the first personalized and user-friendly computers available to scientists, and were replaced in the early 1990s by personal "microcomputers" (PCs) manufactured by modern brands like Sun Microsystems, Apple, and IBM (Malerba et al. 1999). Consequently, by adopting such computing technologies, researchers introduced techniques that had been previously used for the management of data in corporate and management settings into their biochemical work (Chow-White et al. 2011:138).

It was at this time that scientific researchers first began to conduct research at their own workstations and to control their own personal computers, using a mix of commercially available software, and writing their own code in programming languages such as "R" (Esbensen et al. 1990:391). As computer technologies became increasingly central to the generation and analysis of biochemical data, scientists became versed in the "art" or "craft" of data analysis. Without access to stabilized, black-boxed computer software[27], researchers designed customized statistical programs and used personal styles of data analysis. The growth of personal computers, as Brian Pfaffenberger (1988) has written, gave scientists control over their own research and facilitated creativity. It gave them access to new possibilities for using and interpreting biochemical and statistical data. Likewise, as Peter Chow-White and Miguel Garcia-Sancho (2011) have noted, scientists began to adapt their approaches to biological research to the capabilities of computer technologies.

The practices required to carry out rational drug design involved a combination of data analysis techniques, which consisted of pattern recognition tools and algorithms focused on the classification of molecules and their biological properties (Farrant et al. 1992). Such

[27] In this field, one of the first black-boxed pieces of commercial software was "ARTHUR," which was developed in the 1970s with the computer programming language FORTRAN (Varmuza et al. 2009). It ran on mainframe rather than personal computers, and comprised basic procedures for multivariate analysis, including PCA.

pattern recognition tools involved early forms of multivariate statistics, notably principal components analysis (PCA) and related techniques (Fonville et al. 2010). Though these statistical techniques were applied to the fields of chemistry and biology in the 1970s, they were used throughout the mid-20th century for a diverse range of applications, which included "handwritten and printed alphanumeric character recognition, weather prediction, medical diagnosis, speech analysis" (Kowalski et al. 1972).

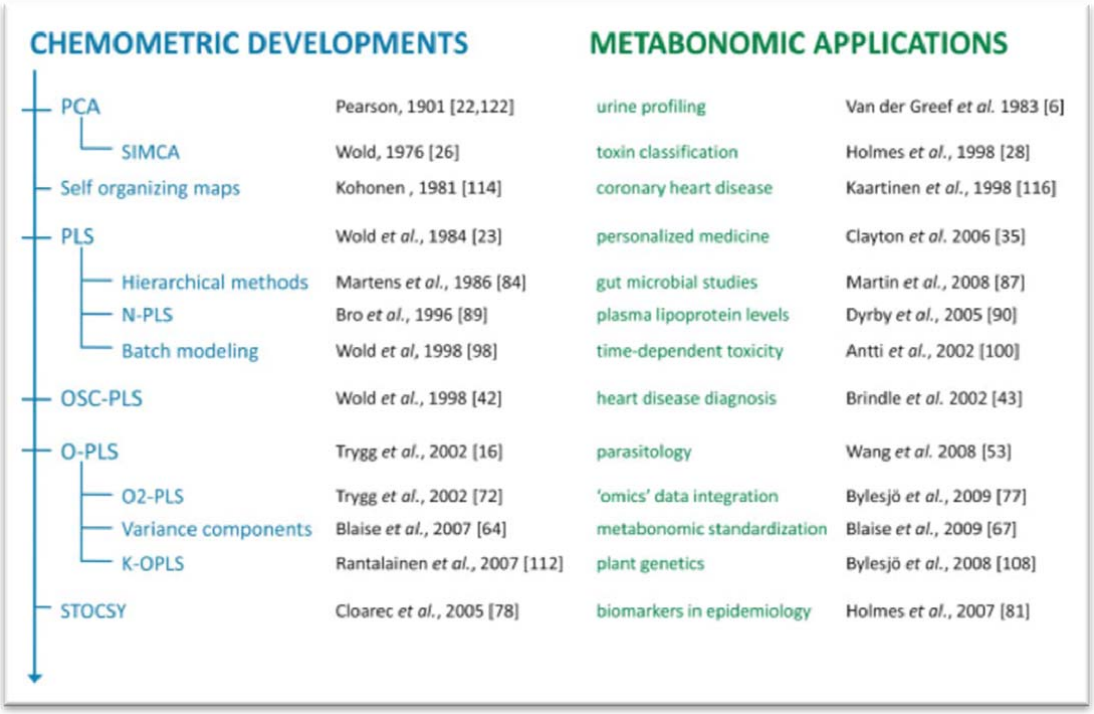


Figure 13: Chart showing the historical evolution of pattern recognition and multivariate statistical technologies from the perspective of metabonomics researchers.
Note how the development of PCA is attributed to Karl Pearson in 1901, while the development of PLS is attributed to Svante Wold in 1984 (Fonville et al. 2010).

As pattern recognition tools and multivariate statistics were applied and adapted to the analysis of complex chemical data in the 1970s, they became the foundation for the field of "chemometrics" (Geladi 1988), which in modern times occupies a central place in metabonomics research. Chemometrics, which is defined as the application of mathematical and statistical techniques to biochemical data, enabled researchers to cope with the demands of increasingly complex datasets. As Esbensen and Geladi (1990:390) write:

> While univariate mathematical and statistical methods were available for chemical data in the early years, many problems could be solved without the use of them…With multivariate data the intuitive capabilities of the human mind were no longer capable of observing what was going on, and scientists were forced to do something extra.

The multivariate statistical techniques from which chemometrics derived, however, had a history that long predated rational drug design. PCA was invented in 1901 by Karl Pearson, and was based on the work of Francis Galton in the late 19th century (Hacking 1990; Umetrics 2012), whose work on anthropometry is central to the history of anthropology. As one of the building blocks for multivariate statistical analysis, PCA was taken up and elaborated upon in the mid-20th century by Herman Wold, an econometrician and statistician, who developed partial least squares (PLS) regression through his work on econometric models (Geladi 1988)[28]. Proponents of chemometrics write that techniques like PCA and PLS became core components of the field through the work of Svante Wold, the son of Herman Wold and the first documented person to use the term "chemometrics" (Varmuza et al. 2009) (see Figure 13). Though the evolution and practice of chemometrics is diverse and beyond the scope of this dissertation, it should also be noted that Svante Wold went on to found Umetrics in 1987—the company that produces the software SIMCA-P+[29], one of the basic toolkits for statistical analysis in the BMM—as an attempt to bring multivariate data analysis to the chemistry issues faced by the pharmaceutical industry.

Overall, the use of pattern recognition tools in rational drug design signals the deep history of multivariate statistical methods within metabonomics as a field. Through John Lindon's work on rational drug design, chemometrics became central to the field of metabonomics, and with it multivariate statistical techniques were established as key practices for data analysis. In metabonomics, multivariate statistical techniques have continued to evolve in parallel to the increasing computational capabilities of computers, and also in parallel to the changing requirements for the analysis and interpretation of complex biological data. The history of pattern recognition and chemometrics, therefore, showcases

[28] Many other developments in statistics—some of which have surprising connections to the field of chemistry—were happening prior to Herman Wold's work on PLS. In the early 1900s, William Gosset, a chemist working for the brewer Guinness, developed the Student's t-test as a way to monitor the quality of stout. Gosset was hired because of Guinness' policy of recruiting top graduates from Oxford and Cambridge, but was not allowed to publish his findings under his true name, and therefore used the pseudonym "Student." Several years later, Ronald Fischer, a geneticist and statistician working on Mendelian inheritance—built on Gosset's work to develop seminal work on design of experiments and variance.

[29] Similar to John Lindon's comments on the evolution and importance of personal computers, Svante Wold remarked in a published interview by Geladi and Esbensen (1990) that the arrival of microcomputers (early PCs) enabled SIMCA to develop into a commercial program.

metabonomics' continual emphasis on both the capabilities of and necessity of using multivariate statistics, which are seen as advanced tools for approaching complex data, and for understanding metabolism and biology.

Nuclear Magnetic Resonance (NMR) Spectroscopy

While John Lindon was using pattern recognition tools and early forms of multivariate statistics at the Wellcome Laboratories, Jeremy Nicholson was working with NMR spectroscopy at Birkbeck College. NMR is a technique that exploits the magnetic properties of atomic nuclei to determine the chemical and physical properties of molecules. The term "nuclear magnetic resonance" refers to the fact that the nuclei of atoms absorb and re-emit energy at various frequencies in response to magnetic fields of various strengths. By detecting the unique magnetic signatures of atoms and compounds, NMR allows researchers to identify the chemical structure—and also confirm the identity—of unknown biochemicals.

NMR itself has a rich history that dates back to experimentation on radars and radio frequency power during World War II. The discovery of NMR spectroscopy in the early 1950s—the application of the technology to the study of liquids and solids—is credited to two groups of physicists working independently of one another, the first led by Felix Block at Stanford University, and the second led by Edward Purcell et al at Harvard[30]. NMR was first applied to the study of biological samples in the 1950s, as researchers used NMR to look at the effects of hydration on DNA, one year after Watson and Crick announced the discovery of DNA (Evans 1995). From the 1960s onwards, NMR was routinely used to determine the structure of molecules in the field of organic chemistry. Following rapid advances in superconducting magnetics and pulse-fourier transform methods in the 1970s (Grant et al. 1997), NMR was capable by the 1980s of identifying a wide range of small molecules in biological samples[31]. In this way, metabonomics—with its devotion to the study of the biological molecules found in biological fluids—is tied to the capabilities and history of NMR spectroscopy.

---

[30] Their work built on physicist Wolfgang Pauli's theoretical work in the 1920s, which established that atomic nuclei have spins and therefore generate magnetic fields. From this, physicist Isador Rabi empirically measured the spin and magnetic moment of atomic nuclei, paving the way for applied experiments in NMR. The Block and Purcell groups shared the Nobel Prize in Physics in 1952 for their work on NMR (Filler 2009).

[31] NMR was also used to develop clinical imaging—in the form of magnetic resonance imaging (MRI)—in the 1970s (Joyce 2008).

After completing a PhD in chemical oceanography on metal biochemistry in sea birds, Jeremy Nicholson began to conduct inorganic chemistry research on the mechanisms by which metal ions enter into cells and bind to molecules to cause cancer. In this research, he used NMR to investigate the mechanisms through which cadmium metal enters into and affects signalling in blood cells. Jeremy Nicholson realized, however, that this would require investigating how cells behaved in their natural blood plasma environment, rather than in an isolated chemical buffer (Wenner 17 June 2008). He recounted:

> And so when we did that, I did that, there was a whole load of extra signals from the plasma. Yeah [laughs] there's going to be stuff in the plasma as well as inside the cell! And so we ran the plasma and it gave these amazing spectra, and from that day onwards I was kind of hooked into the idea that the profile is obviously telling you something about the biochemistry of the plasma, the plasma tells you about the person or the individual. And that's where metabonomics really came from, the idea that you could profile body fluids or tissues and extract pattern information that relates to health or toxicity.

Conducting these experiments, Jeremy Nicholson began to sediment the use of NMR in investigations of the biochemical composition—and therefore metabolism—of liquids and solids. At the same time, John Lindon had developed extensive experience with NMR through his work in academia, prior to his involvement with the pharmaceutical industry and the Wellcome Research Laboratories. For example, John Lindon had completed his PhD in chemistry and NMR spectroscopy in the 1969 at Birmingham University, and subsequently completed a post-doctoral position at Columbia University, where he used NMR to conduct research on the liquid crystals that drive digital displays. In the 1970s he returned to the United Kingdom to a job at Southampton University, but soon thereafter left academia to work at the Wellcome Research Laboratories.

The field of metabonomics began to truly form when Jeremy Nicholson and John Lindon met by chance in 1987 at an NMR meeting in London. John Lindon described hearing a talk by a "young whippersnapper named Jeremy Nicholson," who was conducting research on NMR spectroscopy in body fluids. During his presentation, Jeremy Nicholson described having difficulty interpreting the biological patterns in his data from NMR experiments on blood plasma. At that moment, John Lindon had, as he described, "the realisation that…we could marry…research in NMR and chemometrics into a really worthwhile application, and one that would benefit pharmaceutical R&D, where I was employed at the time." Thus, John Lindon realized that the pattern recognition methods he had been using in the Wellcome Research Laboratories could be combined with Jeremy Nicholson's work on metabolism (Lindon et al. 2008). Following the NMR meeting, John

Lindon approached Jeremy Nicholson and initiated a collaboration "to investigate the use of such multivariate statistics to classify samples according to their biological status" (Lindon 2010). Though others whose names are not mentioned here were no doubt involved in the creation, and the historical narrative, of metabonomics, John Lindon and Jeremy Nicholson's meeting and collaboration marks the beginning of the field of metabonomics as it is told by researchers within the BMM.

Figure 14: Metabonomics researchers before they came to Imperial College London.
A photo of John Lindon (left), a researcher named Maria Anthony (middle), and Jeremy Nicholson (right) at the Wellcome Laboratories in the early 1990s, taken from a section of the BMM's website titled "The Early History of Metabonomics" (Lindon 2010).

John Lindon and Jeremy Nicholson began an official collaboration at Birkbeck College in 1995, when John Lindon left the Wellcome Research Laboratories to join Jeremy Nicholson's research team at Birkbeck (see Figure 14). They were joined by one of Jeremy Nicholson's early PhD students, Elaine Holmes, who later came to play a key role in the establishment of the BMM at Imperial College London. By the 1990s, the first papers combining NMR spectroscopy and pattern recognition methods began to appear (Gartland et al. 1990). Thus, at Birkbeck, John Lindon and Jeremy Nicholson began to establish the validity and success of combining multivariate statistics with NMR spectroscopy to study health and disease. The field of metabonomics as we know it began to officially form.

Given the intersections between John Lindon's work on rational drug design and Jeremy Nicholson's work on NMR spectroscopy, the story of metabonomics' foundation as a field is, in reality, a story of two experts coming together to combine techniques for the analysis and collection of data. This division between the two realms of practice is, of course, not so clear cut, as signalled by John Lindon's extensive research experience with NMR prior to his collaboration with Jeremy Nicholson. Overall, the work of John Lindon and Jeremy Nicholson established multivariate statistics and NMR as the conceptual and experimental practices that are central to the efforts and negotiations of the field of metabonomics in contemporary times. These technologies became knowable objects within metabonomics through developments across a surprisingly diverse array of disciplines and fields. On the one hand, multivariate statistics were practiced in fields such as econometrics, psychology, computer science, and analytical chemistry, while on the other hand NMR was used for research in fields such as nuclear physics, structural biology, toxicology, and crystallography. Thus, it was not the presence of multivariate statistics and NMR that defined metabonomics as a field, but rather their specific use by specific people and in specific contexts to study metabolism.

Metabonomics Comes to Imperial College London

In 1998, John Lindon and Jeremy Nicholson moved from Birkbeck College to Imperial College London, accompanied by Elaine Holmes. There, they established the Biomolecular Medicine Laboratory (BMM) on the sixth floor of the brand-new Sir Alexander Fleming (SAF) building in South Kensington, London. As John Lindon described, Jeremy Nicholson was "courted" by Imperial College London, and was invited to become a Professor in the new Section of Biological Chemistry. Jeremy Nicholson described the move to Imperial as a "step up in Universities," saying that "it gave a …completely new environment to work in, with a very big medical school…and I was basically continuing the work, but with more close clinical connections." Soon thereafter, the newly established group recruited several other researchers to join them[32], and set out to purchase biochemistry equipment such as NMR spectrometers.

[32] These included Timothy Ebbels (now a Senior Lecturer in the BMM), Jules Griffin (now at Cambridge University), Richard Barton (now a Research Fellow in the BMM), and Hector Keun (now a Lecturer in the BMM)

During this time, Jeremy Nicholson and John Lindon developed a close collaboration with the company Bruker Biospin, one of the leading global manufacturers of NMR spectrometers[33]. John Lindon and Jeremy Nicholson's collaboration with Bruker Biospin had begun much earlier, starting with John Lindon's exclusive use of Bruker Biospin NMR spectrometers at the Wellcome Laboratories in the mid-1970s. When John Lindon joined Jeremy Nicholson at Birkbeck College in the 1990s, he formed an official collaboration with Bruker Biospin, agreeing that the company would provide substantial funding for experiments—as well as Elaine Holmes' post-doctoral position—as long as Jeremy Nicholson and John Lindon conducted experiments exclusively on Bruker NMR spectrometers. In this way, throughout the 1990s the BMM also began to develop close research ties to Manfred Spraul, the acting Direct of the NMR Application Group at Bruker Biospin. Through these ties, the BMM began to develop expertise on magic angle spinning nuclear magnetic resonance (MAS-NMR), a particular type of NMR that analyses the composition of solid tissues (see Chapter 4).

Over the next few years, the space occupied by the group gradually expanded throughout the SAF building, and came to house increasingly comprehensive technologies. Richard Barton, one of the first people to be hired by Jeremy's new Imperial research group, described the BMM's early days: "The lab wasn't really here… I didn't think anything about metabonomics. I was just like, oh there's some NMR, and there's some maths, and this will be fine." During this time, metabonomics researchers in the BMM carried out key work in toxicology, drug metabolism, and solid-state NMR, and began to diversify the range of projects and publications in which metabonomics research was involved.

An Era of Drug Toxicity Research

Over the next few years, the laboratory grew in size and prestige, and went through, as Jeremy Nicholson described, "various evolutionary processes." Throughout the mid-2000s, the group continued John Lindon and Jeremy Nicholson's exploration of drug mechanisms and toxicology. In 2003, the Consortium for Metabonomic Toxicology (COMET) was established as the laboratory's first large-scale collaboration between academia and industry.

[33] Bruker hold more than 70% of the world market (Agrocos 2012). Although Varian, a competing NMR manufacturer, was the first company to make NMR spectrometers commercially available after their development in the 1950s, within 20 years Bruker had the leading edge in the markets. This was because, as John Lindon described, they were able to keep up-to-date with modern developments in NMR research.

This collaboration with commercial entities shaped the laboratories research practices and agenda throughout the 2000s, ensuring access to funding and research technologies, but also ensuring that the BMM focused on topics relating to drug metabolism.

It is important to note that the COMET project—and its overarching focus on drug toxicity—was not the only research activity carried out by the BMM during this time. Researchers continued to focus on a variety of topics, including the gut microbiome, statistical and chemometric methods, parasitic diseases, molecular epidemiology, and nutrition and obesity. My aim in focusing on the COMET project is to show how this period in the BMM's history was pivotal for its future activities. It enabled the laboratory to establish key partnerships with commercial entities, and provided a launching pad from which the BMM was able to increase the scope, scale, and prestige of its work. Thus, during the COMET project, the laboratory's research activities were largely shaped by the context of its collaborations and funding opportunities, supporting the notion that science and culture are deeply intertwined.

To establish the COMET project, Jeremy Nicholson drew on the close ties to leading European pharmaceutical and nutrition companies, which he and John Lindon had begun to establish during their time at Birkbeck College and the Wellcome Research Laboratories. In totality, the three-year COMET project involved the collaborative work and funding of six different pharmaceutical companies[34] (Lindon et al. 2005). The COMET project entailed the application of metabonomics techniques to issues in drug toxicology, with the objective of investigating whether particular drug compounds were toxic to the liver or kidney. The drug companies that participated funded five post-doctoral research positions, along with one Bruker Biospin NMR spectrometer, in exchange for the laboratory's expertise in using metabonomics to investigate drug toxicity. In the end, COMET investigated more than 147 toxins and treatments—supplied by the drug companies—and built databases and statistical

[34] According to John Lindon, the COMET project typified the shifting nature of the pharmaceutical industry throughout the earlier 21st century. "Pharmaceutical genealogies is sort of an interesting subject, the way drug companies appear and disappear," he said. Over the course of its three year duration, the COMET project grew from a collaboration between three pharmaceutical companies to a collaboration between six: Bristol-Myers-Squibb, Eli Lilly, Hoffman-La Roche, NovoNordisk, Pfizer, and The Pharmacia Corporation (now within Pfizer). By the end of the project, however, only four companies—Hoffman-La Roche, NovoNordisk, Bristol-Myers-Squibb, and Eli Lilly—remained due to mergers and closures.

models to predict disease outcome[35]. The project typified many of the BMM's collaborations with commercial entities, in which the BMM received resources in exchange for its knowledge of NMR and data analysis methods.

Metabonomics researchers depict the COMET project as one of the most successful times in the BMM's early history. The project produced several high-impact and highly-cited papers, and through a spin-off collaboration with the company Pfizer produced the research and papers that led to the concept of "pharmacometabonomics" (Clayton et al. 2006) (see Chapter 6). Researchers in the BMM attribute the lucrative and successful COMET collaboration to Jeremy Nicholson's success with recruiting industry partners to participate in the research of—and also donate funds and supplies to—the laboratory. As one researcher commented:

> He got the pharmaceutical companies together and banged their heads together, and said 'There's something in this for all of you, but I've got have all of you online. And by the way we want money.' So, it's very hard to get these guys to agree and to talk to one another when they're all competitors…so he did a pretty amazing job on the COMET project.

According to researchers in the BMM, Jeremy Nicholson had "worked miracles" to get competing pharmaceutical companies to fund research activities, and to agree on data sharing and access. Jeremy Nicholson had convinced the pharmaceutical companies that the COMET project would be highly beneficial to drug development, by claiming that it would allow them to learn about the mechanism of action, and the potential toxicity, of their drugs. The BMM would provide the best analytical capabilities, and consequently the best results. As one researcher said: "Well, he's pulled good funding. He's a very, very good fundraiser…the reason we've got such good facilities is because he basically raised all of the money for them…everything you see [upstairs] was put together by Jeremy." In this sense, the COMET project entailed a particular form of "biocapital" (Rajan 2006; Helmreich 2008), a term used to describe how biological substances and promises are transformed into objects for profit making. By making the future potential of metabonomics explicit, Jeremy Nicholson was able to promote metabonomics as a commercial investment.

The public-private research collaboration that characterized the COMET project was also a key feature of Jeremy's Nicholson's leadership, and consequently his way of securing

[35] The COMET Project was continued as COMET-2 in subsequent years, but this project was less well known and involved fewer pharmaceutical companies, such as Bristol Meyers Squibb, Servier, Pfizer, and Sanofi-Aventis.

funding and research opportunities for the BMM. As institutions throughout the United Kingdom struggled to gain funding for physical and biological research—due in part to incremental governmental budget cuts from UK-wide funding bodies such as the Engineering and Physical Sciences Research Council (EPSRC) (Jha 7 October 2010; Jha et al. 15 August 2011)—Jeremy Nicholson developed his own strategy. He attempted to supplement government funding with public-private partnerships, which he used to build the research capabilities and infrastructure of the BMM. As one researcher described, Jeremy Nicholson went about "charming the industry," developing collaborations not only through the COMET project, but also with companies such as AstraZeneca, Nestlé, and Pfizer. Jeremy Nicholson's ties to industry, therefore, were a key feature of the success of the field of metabonomics. They provided the BMM not only with monetary and technological resources, but also with opportunities to publish a large number of papers in high-impact journals such as *Nature*, *Science,* and *Proceedings of the National Academy of Sciences*.

The historical developments of the COMET project, and of the BMM more broadly, beg the important question of how central a figure Jeremy Nicholson has been to the development of metabonomics as a field. There is no question that his research success and personality—and in particular his skills in networking and promoting a centralized research vision—have been central to the development of the BMM. As John Law (2009) emphasizes in his discussion of 19th century inventor Thomas Edison, such networking skills enable researchers to establish high-profile collaborations, and to gain access to extensive funding and resources. Jeremy Nicholson's established place at the center of the activities of the BMM and metabonomics hints at a strong sense of personal ownership and pride. What I want to call attention to here, however, is the absence of alternative narratives—ones which mention other names and historical events—to the one presented in formalized accounts of the BMM's history.

Although the BMM's collaborations with private pharmaceutical and nutrition companies were fundamental to the laboratory's success, they also provided distinct challenges for metabonomics researchers, particularly at lower levels of the laboratory's hierarchy. Conducting collaborative research with commercial entities generated a more competitive and high-pressure research atmosphere, as researchers within an academic environment were held accountable to the deadlines and monetary interests of commercial companies. In particular, researchers commented on the intellectual property (IP) issues that arose from collaborations with the nutritional company Nestlé, and how it constrained their

ability to publish experimental results. Due to the company's interest in patenting research that was of potential commercial use, several researchers struggled to publish results that were based on samples they had received from Nestlé. One graduate student was forced to wait many months for the company's approval, and was consequently rejected for publication because her findings had been published several months prior by a competing research laboratory. Thus, while the BMM's collaborations with commercial entities were beneficial for some researchers in positions of power, they were detrimental to others who occupied the lower ranks of the laboratory.

An Era of Translational Medicine Research

Following the success of the COMET and other contemporary research projects, Jeremy Nicholson was appointed Head of the Department of Surgery and Cancer in 2009, as the Faculty of Medicine underwent a large-scale reorganization[36]. As the person in charge of the Surgery and Cancer "super department," Jeremy Nicholson became responsible for more than 1000 researchers in laboratory and clinical specialties, whose activities were spread over six different research and hospital campuses. With this promotion, his activities focused more on management than active research, though he continued to drive the research vision of the BMM.

As a result of the power and prestige that his new position afforded, Jeremy Nicholson began to spread metabonomics research throughout the varied research activities of the Department of Surgery and Cancer. By liaising with the Sections and laboratories in the Department, he increased awareness of the capabilities and potential applications of metabonomics. He encouraged clinical researchers to conduct collaborative research projects in the BMM, and conducted media interviews about the evolution and future of the field (Bhattacharya 14 December 2009; Wenner 17 June 2008). But in doing so, he risked alienating researchers whose work entailed different biological approaches and practices, and who could no longer gain access to funding without involving metabonomics. Consequently, Jeremy Nicholson used metabonomics research as a political tool to unite a diverse set of disciplines and to promote his own field of research.

---

[36] During this change, the Department of Surgery and Cancer—formerly known as "SORA" for Surgery, Oncology, Reproductive biology, and Anaesthetics—was subdivided into the Division of Cancer and the Division of Surgery. Consequently, the Division of Cancer came to include the Sections of Oncology and Reproductive Biology, while the Division of Surgery came to include the sections of Biosurgery & Surgical Technology, Biomolecular Medicine, and Anaesthetics, Intensive Care & Pain Medicine.

Thus, Jeremy Nicholson began to make metabonomics central to almost everything that occurred in the Department of Surgery and Cancer. As John Lindon commented:

> He took the view that he would go around every outpost of his empire – and it's a huge department, as you've probably realized, what is it, 1000 people, I don't know? – and go around and find out what all the major people were doing. And he would take Elaine or me or both, so we would go to St. Mary's [Hospital] for a morning and meet the surgical group, and [they would] tell us what they were doing. Or go to the Charing Cross [Hospital] and hear the bone surgeons tell us what they were doing, and the hip replacement people and the varicose veins people… Or go to the Hammersmith [Hospital] and see the cancer people. So, Jeremy got a good vision straight away of what his empire was all about, and that enabled him to [position] metabonomics as a sort of driver for all sorts of things.

To this end, the BMM established several high-profile clinical research projects (see Chapters 5 and 6) that encompassed the broad agenda of "translational medicine" and "personalized medicine." Such projects focused on the application of laboratory research to clinical issues, but also advanced the notion that metabonomics could aid in the development of diagnostics and therapies tailored to individual patients within hospital settings (Kinross et al. 2011). As Jeremy Nicholson himself described, if he had not become Head of the Department of Surgery and Cancer, "none of it might have come [about]…only one or two projects might have come through, because of the nature of the hierarchical structure of Imperial."

The BMM's turn to clinical projects built off of broader efforts to promote translational and personalized medicine throughout the UK, particularly through large-scale public funding opportunities. The BMM's activities were supported through a variety of public and private entities, which included more than £7 million from the Imperial College National Institute for Health Biomedical Research Council (NIHR-BRC)[37], and more than £20 million from the analytical instrument companies Waters Corporation and Bruker Biospin. In the broader context of the UK, such funding opportunities occurred concurrently with other translational and personalized medicine initiatives. In biomedical research centres throughout the country, large amounts of funding for translational and personalized medicine were being made available through Cancer Research UK (Scowcroft 2011), the Francis Crick institute, the Medical Research Council (MRC), and the Wellcome Trust (EuroBioForum 2013). Funds to develop personalized medicine were also made available by the Stratified

[37] NIHR-BRCs are established to drive progress on innovation and translational research in biomedicine into NHS practice. They are based within outstanding NHS and University partnerships, and provide funding for a variety of translational projects. Imperial College is one of several research institutions that houses BRCs (Imperial College London 2012d).

Medicine Innovation Platform, which was launched in 2011 by the UK Technology Strategy Board and included a five year partnership—as well as £11 million of first round funding—to develop national-level research activities (Mansell 14 October 2010)[38].

To support the increased research activities of these translational and personalized medicine projects, the BMM established two new analytical facilities: the "Clinical Phenome Centre" at St. Mary's Hospital and the "MRC-NIHR Phenome Centre" at Hammersmith Hospital (see Chapter 6). The first Centre was developed through a combination of public and private funding, which was provided by the NIHR-BRC, the US National Institute of Health (NIH), the Gates Foundation, and renewed collaborations with the analytical instrument manufacturers Bruker Biospin and Waters Corporation (Wong 18 May 2011). This resulted in the installation between 2010 and 2013 of a series of customized MS and NMR spectrometers, which were designed to be used in clinical settings, and were equipped with specialized capacities for automation (see Chapter 4).

The second Centre was developed following the 2012 London Olympics, as the BMM—in a joint venture with King's College London—commissioned the use of the Olympic doping testing facilities. Though the Centre was originally slated to open at GlaxoSmithKline's (GSK) New Frontier's Science Park in Harlow, the original site for the Olympic testing laboratory, it was relocated for "operational reasons" to Hammersmith Hospital. The second Centre was funded by the NIHR-BRC and the MRC, and was equipped upon its opening in 2013 with more than 20 NMR and MS machines. Together, Jeremy Nicholson emphasized that these two Centres would become a locus for clinically-directed metabonomics experiments, which would analyze patient samples and help to build a database for future clinical trials (see Chapter 6).

[38] Within the United States, the National Institute of Health (NIH) and the Food & Drug Administration (FDA) produced a road map to personalized medicine in 2010, which promised to commit money and research efforts to the development of molecular drugs and tests (Hamburg et al. 2010), and involved cooperation with pharmaceutical companies and regulatory bodies. Within Europe more broadly, the European Commission committed more than €900 million between 2007 and 2011 to research enabling the development of personalized medicine approaches through the Health Theme of the Seventh EU Framework Programme for Research and Technological Development (FP7) (Draghia-Akli 2012:151; European Commission 2013b). It is also expected that personalized medicine will figure significantly into the funding allocation for Horizon 2020, the European Union's premier funding program for research and innovation (European Commission 2013a).

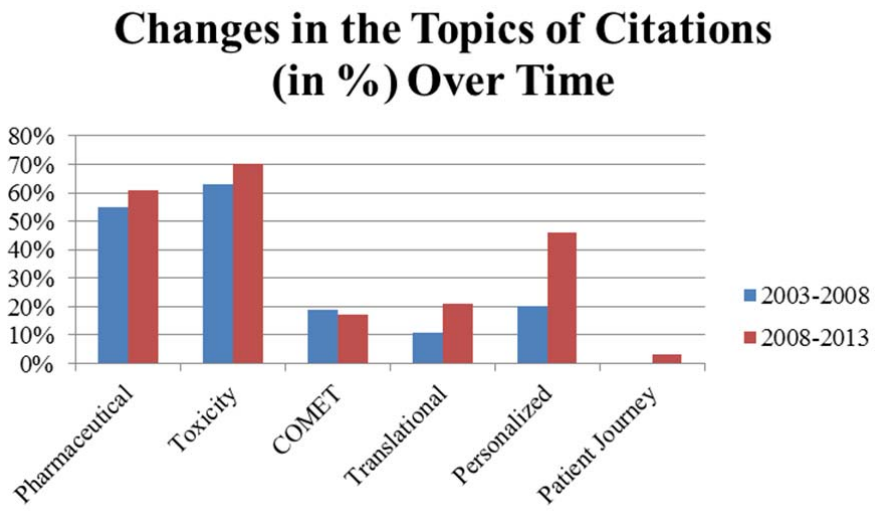


Figure 15: Changes in topics of papers published by Jeremy Nicholson between 2003-2008 and 2009-2013.
Data derived from a Google Scholar search of papers published by "JK Nicholson," and with the search criteria (words) indicated in the left-most column. Papers on pharmaceutical/toxicity dominate publications between 2003-2008, while a combination of papers on pharmaceutical/toxicity and translation/personalized medicine make up the publications between 2009-2013, with a significant increase in the number of papers on translation/personalized medicine occurring during this latter time period. Note that a total of 243 papers were published in 2003-2008, while a total of 150 papers were published in 2008-2013. Note also that multiple search criteria can appear in the same paper.

Concurrent with attempts to grow its translational research capacity, the BMM sought to develop its analytical capacities and expertise with the technology MS. Similar to NMR, MS generates information about the molecular composition of biological samples, but with the use of ionization rather than magnetism. To acquire a number of MS machines and MS-focused staff, the BMM built on its collaboration with the Waters Corporation, which began in the early 2000s and expanded in 2006 with the establishment of the "Waters Laboratory of Molecular Spectroscopy" (Waters Corporation 26 September 2006). This was an attempt to be seen, as one researcher commented, "as not just an NMR metabolic profiling group." Researchers commented, for example, that nearly 90% of the BMM's research was focused on NMR, and that the lack of MS research was due to the BMM's relatively late investment in 2006 in MS expertise and technologies. Thus, in expanding its MS capacities, the BMM sought to increase its competitiveness with metabolomics groups in the United States, which tended to use MS more than NMR because of its higher sensitivity (see Chapter 3).

Such research activities signal strategies to demarcate, promote, and gain funding for metabonomics as a field. While the BMM's early research focused on drug metabolism and toxicity, upon his appointment as head of the Department of Surgery and Cancer Jeremy Nicholson shifted the BMM's research focus in line with topics—translational and personalized medicine—that were being promoted widely by public and private biomedical institutions (see Figure 15). This was an effort, in part, to justify metabonomics as one of many fields of post-genomic research that could be used to develop translational medicine. It was also an effort to ensure that metabonomics received funding and support following the perceived failures of genomics and the Human Genome Project (HGP). As Ian Wilson, a researcher from AstraZeneca who had a long-standing collaboration with the BMM, commented:

> There are three or four omics…and unfortunately in the beginning there was genomics, and that was never-endingly oversold…And like all things that were going to solve every problem, it didn't, but it cost a lot of money. And then along came the proteomicists and said, forgot genomics, we can solve all your problems. And by the time it came to metabonomics, most companies were fairly exhausted at having been promised the world and actually being delivered very little. And now that's because these technologies take a while to mature. And now the genome [and] the transcriptome are becoming really quite important in personalized medicine, and the proteome will begin to get there as well. The metabolome is a bit further down the list, and that's because most companies don't have the resources to have something of everything.

Overall, this section shows how the BMM used the interface between metabonomics and personalized medicine as a social, economic and political tool, and as a way to gain access to funding and facilities from Imperial College London, private corporations, and the UK government. This was in part enabled by the "culture" of research at Imperial College London, which encouraged the development of public-private collaborations and networking for research. It also resulted from the increased scope and number of personnel involved in metabonomics research, as Jeremy Nicholson emphasizing the use of metabonomics in a wide range of surgical and oncological disciplines. Thus, by drawing on personalized medicine as an area of scientific interest and funding, the BMM demarcated and promoted its realm of research activity, and also unified the research activities of the Department of Surgery and Cancer around its own field of expertise.

Discussion

In conclusion, metabonomics research in the BMM has a long history that combines and builds upon research in multiple fields including, but not limited to, drug discovery, toxicology, chemistry, and chemometrics. In telling the history of the BMM, metabonomics

researchers retrospectively construct their field as the intuitive combination of research on pattern recognition—an early form of multivariate statistics—and biochemistry—the application of NMR to the study of biological fluids and tissues. They perform their field as something long and enduring, as they emphasize that metabonomics was established before its official definition in 1999, and that it dates back to the invention of the widespread technologies NMR and PCA in the mid-20th century.

Tracing the history and context of metabonomics research in the BMM shows how metabonomics as a field does not exist as a naturalized object, but rather is undergoing a constant project of self-realization and stabilization. By virtue of the fact that metabonomics contains the suffix "omic," it is positioned in relation to—and in the wake of—the genomic research of the HGP. Metabonomics, as one of several post-genomic fields of research that include epigenetics and proteomics, is engaged in a constant effort to define the future potential of the field[39]. Moreover, as a laboratory that is competing with other "metabolomics" laboratories, the BMM is engaged in a constant effort to assert the originality and legitimacy of metabonomics. At stake in the establishment of either metabonomics or metabolomics as the primary term to describe the post-genomic study of metabolism is the legitimation of a particular set of practices and individuals in the history of biology. Thus, as metabonomics researchers in the BMM attempt to expand the status and reach of their work, they must carry out rhetorical "boundary work" to cordon the boundaries and to clarify the epistemic status of their laboratory and field.

Somewhat surprisingly, the historical narrative of metabonomics has much in common with that of proteomics, the post-genomic study of proteins. As Glyn Moody describes in *The Digital Code of Life* (2004), the term proteomics was coined in 1994, though the field is said to have developed concurrently with genomics throughout the 1980s. According to Moody, proteomics, much like metabonomics, emphasized that it was able to account for variable and dynamic changes in the proteome, the sum total of proteins. Much of the early work of the field was consequently directed at standardizing technological platforms and informational practices. In addition, researchers directed their efforts at the application of proteomics technologies to clinical applications, attempting to improve the

[39] As a recent commentary in the *Wall Street Journal* observes, the proliferation of the "ome" suffix following the discovery and exploration of the genome is a product of scientific hype. "Ome," which is now found in more than 400 terms, has become a biological buzzword: it indicates the explosion of scientific sub-fields that "sound futuristic…[and] computational." (Hotz 13 August 2012).

diagnosis and treatment of disease conditions. Like metabonomics researchers, researchers in proteomics were concerned about their field's ability to engage with the range of human variability and complexity, and to relate proteomic information to understandings of biological function[40]. Such claims and challenges—which I discuss further in Chapters 4 and 5—echo those faced by metabonomics researchers as they negotiate the historical roots and future applications of their field.

Amidst attempts to establish and stabilize the field, metabonomics research at the BMM emerges as a unique epistemic culture of histories, locations, technologies, networks, and practices of knowledge production. Metabonomics research is highly intertwined with the academic environment of Imperial College London, which encourages not only academic rigor and competition, but also the establishment of public-private partnerships. It is also intricately linked to the economic climate of the 21st century United Kingdom, which has encouraged the BMM to establish partnerships with both analytical instrument manufacturers and nutritional and pharmaceutical companies, in order to gain access to resources that would not otherwise have been available with public funds. In turn, such commercial alliances have shaped the over-arching research agenda and climate of the BMM, leading to an era of research expertise and publications on the metabolism of drug toxicity, and also intensifying the BMM's ties to commercial concerns.

As the broad research focus of the United Kingdom has shifted during the early 21st century to focus on translational and personalized medicine, however, so too has the research focus of the BMM. Launching a series of projects on the clinical applications of metabonomics technologies (see Chapter 5), the BMM has begun to establish itself as a hub for biomedical research on individual and population phenotyping (see Chapter 6). But importantly, the BMM's endeavors to carry out research on translational and personalized medicine have been facilitated by the appointment of Jeremy Nicholson as Head of the Department of Cancer and Surgery. Overseeing the research and clinical practice of more than 1000 people, Jeremy Nicholson has not only promoted metabonomics, but also used it as a political tool to unite the diverse activities of the Department. Consequently, such strategies to gain funding and recognition highlight the ways in which the BMM's research is

[40] Moody writes about one well-publicized experiment, in which proteomics researchers were able to distinguish between healthy and diseased patients with ovarian cancer with a pattern of five proteins. But as Moody emphasizes, the researchers claimed that "it was not necessary to determine exactly which [proteins] these were, since it was only the pattern that mattered" (Moody 2004:284).

intertwined with the political economy of the United Kingdom, and ultimately the ways in which science and society are co-produced (Jasanoff 2004).

Throughout this history, the technologies of NMR and multivariate statistics have played a central role in the development and epistemic culture of metabonomics research at the BMM. To this end, it could be argued that metabonomics research at the BMM is an "instrumental community" (Mody 2011), a network or community of people who are defined by their use of particular technologies. Within this view, communities are ratified by their involvement with particular technologies—in this case NMR and multivariate statistics—such that particular technologies foster consensus and strengthen the community (Mody 2011:6). Consequently, this chapter foregrounds the ways in which multivariate statistics, as I discuss in the following chapter, form key resources for generating, handling, and interpreting the biochemical data generated in NMR experiments. Ultimately, this chapter demonstrates how NMR and multivariate statistics figure heavily not only into the material practices, but also into the historical identity and epistemic culture of metabonomics as a field.

Chapter 3: Multivariate Statistics and Biological Complexity

Introduction

I am interviewing Jacob, a statistics-savvy researcher, in the atrium of the SAF building at Imperial College London. I have only been doing fieldwork for a few months, and he has kindly agreed to speak to me about his work developing methods for the analysis of metabonomics data. In our interview, Jacob describes his work on the development of statistical tools to analyze, visualize, and model metabolism. These tools allow him to make sense of the myriad of information generated in metabonomics experiments. The highly-customizable computer programs that Jacob uses to write codes and algorithms for the analysis of this data, he says, "let me get [my] hands dirty." They allow him to see and interact with his metabolic data in novel and situated ways, such that urine and blood samples exist as highly complex sets of computerized information, which after statistical manipulation, can be transformed into biochemical molecules and metabolic pathways.

Jacob came to the Biomolecular Medicine Laboratory in 2003 after finishing his PhD in analytical chemistry in France. Over his four years as a post-doc, he worked broadly in the field of chemometrics, the application of statistical ideas and methodologies to chemical data. As he developed chemometric data analysis techniques, Jacob used an array of multivariate statistical approaches: techniques for processing, handling, and analyzing data that considered many variables simultaneously, rather than one or two in isolation (see Chapter 2). These included well-known, institutionalized algorithms such as principal components analysis (PCA) and partial least squares (PLS) regression, as well as algorithms that were crafted and "hand-made" with computer programming tools. Though Jacob left Imperial in 2006 to do contract work for various companies—including those involved in chemistry and finance—he continued to collaborate and act as a consultant on various projects in the BMM.

The goal of metabonomics research, explains Jacob, is to determine "biomarkers," measurable and quantifiable biological entities that can be statistically determined in relation to health and disease (Holmes et al. 2008b; Metzler 2010; Beger et al. 2012:3). In the era of post-genomic research, biomarkers are molecular entities whose presence is correlated with—but not the causative agent of—conditions of health and disease (Frank et al. 2003). Biomarkers can technically correspond to any physiological or anatomical measurement, although the most common biomarkers are proteins, genes, or other small molecules. In metabonomics, biomarkers are predominantly made of the biochemicals or metabolites deduced from metabonomics experiments on urine and blood.

Biomarkers in metabonomics, as well as other post-genomic fields of research such as genomics and proteomics, are molecularly measured and determined. To investigate the biomarkers involved in metabolism, metabonomics researchers use technologies like nuclear magnetic resonance (NMR) and mass spectrometry (MS) to measure the metabolites that underpin metabolic processes. This is similar to how genomic researchers use gene sequencing and microarray technologies to determine the molecular components DNA and RNA. In the end, these metabonomics technologies transform biological samples into biochemical information, rendering urine and blood samples into chemical signals (Lindon et al. 2008; Barton 2011). This is the "wet" part of the experimentation, which although it has been the core of laboratory work in "traditional" fields such as cellular biology and physiology, takes up relatively little time and effort in metabonomics experiments.

The practices that Jacob uses to transform biochemical data into biomarkers highlight the centrality of data practices and quantification to the modern life sciences. Jacob, along with other researchers in the metabonomics laboratory, is participating in a shift from science that is practiced at the bench to science that is practiced at the computer. Metabonomics is one of many fields of "data-centric science," in which laboratories function as centres for the storage and analysis of biological data (Moody 2004; Swedlow et al. 2011:463). As scholars in both the natural and social sciences have argued, developments in computing and biomedical research—in particular genetics and genomics—co-evolved between the 1970s and 1990s, causing biology to converge with the informational sciences (Mackenzie 2003; Thacker 2005; Chow-White et al. 2011; November 2012).

As Jacob describes the statistical practices he uses to understand and analyze metabonomics data, he describes metabolism as something that is inherently and utterly

complex. The many pathways, biochemicals, and bytes of data that make up metabolism, says Jacob, evade linear interpretations or explanations. Metabolism is a complicated internal process that is the result of multiple pathways, organ systems and timescales, and also multiple genes, proteins, and metabolites.

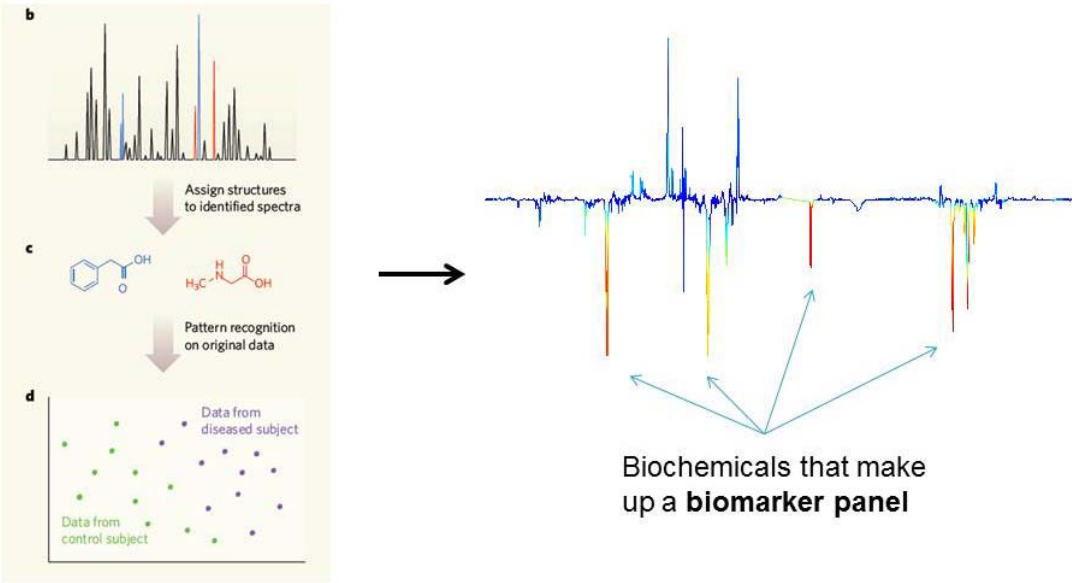


Figure 16: From biochemicals to biomarkers.
A diagram showing how biochemical information from metabonomics experiments is transformed into biomarker panels with the help of multivariate statistics. The figure at left—taken from Nicholson and Lindon (2008)—shows how NMR data (top left) is used to assess the disease status of experimental subjects through multivariate statistics (bottom left). The colourful figure (right) shows how this data is used to determine those biochemicals—which together make up a "biomarker panel"—that are associated with disease states.

Within this worldview of metabolic complexity, biomarkers exist not as single entities, but rather as "biomarker panels": as groups of multiple biochemical compounds that in combination make up a biomarker (see Figure 16). Though "biomarker panels" is my own terminology, the term condenses several prevailing concepts in metabonomics literature. These are "metabolic profile," "metabolic signature," "metabolic phenotype," and "metabotype" (Gavaghan, Holmes et al. 2000; Ebbels, Holmes et al. 2004), all of which build upon the idea that combinations of molecular entities can be used to identify samples, disease processes, or people. Jacob explains:

> For biomarkers, for me, you can understand why I don't consider a single molecule as a biomarker, for me it doesn't make sense anymore. Not just with the data, but also with a biological process. For example…I don't think a biological process is related to one enzyme or one metabolite …And if you extend that to the whole pathway, then you realize that characterizing one single biomarker is ridiculous.

Jacob's statement crystallizes the overall focus of this chapter. It highlights how Jacob's approach to data analysis informs his understanding of the biological processes that underpin metabonomics experiments. By engaging with multiple sets of biochemical information, Jacob views biomarkers in a similar way to statistical data: as complex patterns of information that cannot be seen in isolation. As Jacob describes:

> All of the [metabolites] which are in a network, they are connected, they cannot be on their own taken as a biomarker. It's the combination that's a biomarker. One up one down, one up one down. No! It's the whole thing which has to be defined as a biomarker, and not only one thing. Unfortunately people tend to think about that in terms of "so, what is the metabolite?" And I'm like, no, don't look at that, just look globally! And it's very hard, and that's my everyday struggle.

For Jacob, metabolism is an interconnected network that reflects the combined effects of genes and environment. It occurs at the level of individual cells, but also at the level of the organism, and in a way that connects many levels and layers of biological functioning. Consequently, metabolic biomarkers are multiple and dynamic entities: they exist as the combined effects of multiple metabolites, which reflect the holistic entity that is metabolism. Thus, in their search for biomarkers, metabonomics researchers enact a particular vision of metabolism that is intertwined with the biochemical and statistical tools they use in everyday research. Data practices, therefore, are intertwined with ways of seeing the biological world.

In this chapter, I examine how ways of thinking about metabolism are intertwined and co-produced with methods for investigating biology. I examine how biomarker panels are not only produced through biochemical and multivariate statistical practices, but also draw on these same practices to generate their meanings. Thus, as I explore how researchers use statistical methods to make sense of complex biochemical data, I also explore how researchers use such methods to enact metabolism as an inherently complex entity. To begin, I provide a historical overview of the ways in which metabolism and biomarkers have emerged as central concerns for the field of metabonomics. I then examine the various practices and technologies that make up multivariate statistics in the metabonomics laboratory. From there, I explore how multivariate statistics and biomarker panels emerge as the natural and obvious practices and outcomes, respectively, of metabonomics research. I then explore how complex, statistical views of metabolism are made real through the daily

actions and negotiations of materials, technologies, and people. Ultimately, I discuss how the value metabonomics researchers attribute to multivariate statistical practices reveals the ways in which researchers grapple with the inherent complexity and interconnected nature of metabolism.

Historical Perspectives on Metabolism and Biomarkers

Historical developments in the fields of biochemistry, molecular biology, and pathology emphasize the ways in which complex and biochemical views of metabolism have emerged as the central area of concern in the field of metabonomics. To this end, metabonomics researchers envisage, and perform, molecular forms of metabolism as a historically important area of study. Much like the narrative surrounding the foundation of the field and BMM (see Chapter 2), their narrative retrospectively constructs metabolism as an internal and complex biochemical process, and renders "modern" forms of metabolism—which as a concept did not exist before the 19th century—as "old."

The concept of metabolism has multiple meanings, and is tightly woven with the history of research in the fields of biology and biochemistry. As the sum of the body's chemical processes—whether they are involved in growth, reproduction, healing, eating—metabolism came into scientific focus in the 19th century. As Hannah Landecker (2011) writes in her historical exploration of epigenetics and metabolism, researchers first explored metabolism as a form of energy balance and chemical combustion, and subsequently as a series of biochemical pathways and enzyme-catalyzed chemical reactions. While Landecker asserts that the concept of metabolism evolved in relation to studies of nutrition and food, metabonomics' concept of metabolism is concerned more broadly with the determination of health and disease at the molecular level.

Metabonomics researchers trace the history of studies in metabolism to the first "experiments" conducted hundreds of years ago on the relationship between biological tissues, fluids, excrements, and disease. In a commentary in the journal *Nature*, researchers assert that "the idea that changes in tissues and biological fluids are indicative of disease goes back to at least as far as ancient Greece" (Nicholson et al. 2008:1054). Doing so, they draw on and construct a history in which the Greeks recognized that bodily fluids—classified as

the four humors—could be used to diagnosis disease[41] (Ackerknecht 1982), and from which Galen formalized his theory of humoralism in 131 AD. This stated that bodily essences could be used to characterize diseases, which were was caused by an imbalance among the four humors of black bile, yellow bile, phlegm, and blood.

Continuing a logic in which the physical characteristics of bodily fluids could be correlated to conditions of health and disease, metabonomics researchers assert that Galen's theory of the humors was incorporated into Western medical practice through "diagnostic urine charts." These were used to link the "colors, smells, and tastes of urine to various medical conditions" (Nicholson et al. 2008:1054) (see Figure 17). As evidence of this, English physician Thomas Willis established the first links between the sweetness of urine and the diagnosis of diabetes in 1674, while Matthew Dobson identified that the cause of this sweetness was sugar in 1776 (Tripathy 2012). These urine charts, researchers in the BMM assert, were the precursors to modern techniques used in the field of metabonomics to "relat[e] chemical patterns to biology" (Nicholson et al. 2008:1054).

Figure 17: 16th century diagnostic urine chart.
A diagnostic urine chart from 1506, which was used by physicians to link the physical characteristics of urine to disease states. Metabonomics researchers draw on this as part of the historical foundation of investigations into metabolism (Nicholson et al. 2008).

[41] The term "metabolism" itself is derived from the Greek word "metabolismos," which stands for "change" or "overthrow" (Titz 2006).

Building on a millennium of investigations into metabolism, researchers in the BMM turn to the technological innovations of the 20th century to further their own history of contemporary metabolism. Central to this narrative is the development of the field of biochemistry—the study of the biochemical processes occurring within living organisms—as well as the technologies for its investigation, such as NMR and MS (see Chapter 2). With the rise of biochemistry, humoral accounts of metabolism were replaced in the mid-19th century by theories of cellular pathology (Silverstein 1989:27). This, as Landecker writes, directed research on metabolism to the inward mechanisms or "singular inward laboratory" of the body (Landecker 2011:172). Typifying this internal perspective were the molecular biology discoveries of the mid-20th century, which established how molecules such as enzymes, and biochemical reactions such as the Krebs cycle, underpinned metabolic processes (Bettelheim et al. 2009). This internal perspective was further developed by a research emphasis throughout the 21st century on the molecular machinery—DNA, RNA, proteins, metabolites—of the human body, which established how "life itself" (Rose 2007) was a molecular process.

Such a biochemical and technological history asserts that at the turn of the 20th century, the study of metabolism was preoccupied with internal rather than external substances and processes of the body. This is similarly argued by Michel Foucault in *The Birth of the Clinic* (2003), as he writes that the transition to 19th-century medicine entailed new forms of medical perception, interaction with patients, and classifications of disease. According to Foucault, previously there had existed a "botanical model," in which diseases were classified according to outward signs or symptoms. However, as 19th century medicine began to address social issues such as epidemics, it moved away from the notion of outward disease homologies and classifications, and towards a notion of disease spread, causation, and mechanism. As such, physiological principles entered into the realm of medicine, and disease concepts became organized around, as Georges Canguilhem (1989) argued, the continuum of the normal and the pathological (see Chapter 4). Thus, physiological

explanations of disease focused on the internal or "hidden" processes underlying disease, such that outward symptoms lost their supremacy in medical practice[42].

Building on this notion that metabolism is an internal process, metabonomics researchers envision contemporary metabolism as something that is inherently probabilistic. Metabolic processes result from a combination of endogenous metabolism—products of the body's genes and proteins—and xenobiotic metabolism—products that are not produced within or are foreign to the body, but which can have major effects on metabolic pathways. With this view of metabolism, researchers emphasize not only that gut bacteria play a key role in human metabolism[43], but also that the elements of endogenous and xenobiotic metabolism interact probabilistically to determine the outcome of disease. They liken metabolism to a "Japanese Pachinko machine," an arcade game in which pinballs flow probabilistically through variously distributed pins and exit holes (Nicholson et al. 2003)[44]. Consequently, they portray metabolism as a molecular processes that is governed by the laws of biochemistry, physics, and thermodynamics, writing that "all molecules interact…in a way that is dependent on their molecular physico-chemical properties and their probabilistic collisions…which are determined by the laws of thermodynamics and kinetics" (Nicholson et al. 2003:674).

Overall, metabonomics researchers draw on this historical account of metabolism, which is certainly cursory and partial, to emphasize how the rise of molecular technologies

---

[42] As this occurred, Foucault argues, the clinic became a site not just for the practice of medicine, but also for the social institutionalization and organization of medical knowledge, resulting in the "clinical gaze"—the institutionalized "perceptual code" employed by medical practitioners working in the clinic—as well as new configurations of disease. The clinical gaze articulated illness as something pathological: it did not just observe visible "self-evident" symptoms, but instead actively calculated and intervened in the body to enact signs of disease. As disease was mapped onto the body—as it was placed within the body and became a "pathological form of life" itself—tissues and organs rather than symptoms became the "perceptible space to which [medical practitioners could] relate the phenomena of disease" (Foucault 2003:129).

[43] Metabonomics researchers in the BMM emphasize the central role that gut bacteria play in regulating the immune system and metabolism of human beings. Gut bacteria carry out metabolic processes that assist in the digestion of certain foods and pharmaceuticals. The types of gut bacteria that inhabit humans, however, are influenced by a combination of environmental and dietary factors.

[44] As researchers describe in an article in the journal *Nature*:

> There is a flow of pinballs (read: drug molecules) through the system after introduction ('dosing'). The destination (metabolic fate) of each ball is not determined absolutely, but is probabilistic and strongly influenced by the distribution of pins and exit holes. The positioning of the pins (enzymes that can transform the compounds), and the exact shape and size of the balls (drug properties), determines the route through the machine (cell or tissue) and which holes in the machine the balls exit (disposition and fate). (Nicholson et al. 2003:674)

throughout the 20th and 21st centuries allowed scientists to correlate biochemical patterns—as reflectors of internal metabolic processes—to states of health and disease. Doing so emphasizes and naturalizes the role that biomarkers, as molecular and biochemical entities, play in assessing biology[45]. Consequently, metabonomics researchers at the BMM use this historical narrative to craft a particular understanding and definition of metabolism that is embedded within biochemical and molecular practices for investigating the body. In doing so, they do away with the concept that single biochemical markers correspond to single bodily conditions, and emphasize that their field does not conform to "simple" or linear ways of thinking about the relationships between bodily substances and disease.

Given such an account of metabolism, one might assume that we are confronted with a story of technological determinism, such that these new ways of investigating metabolism are the result of new and improved molecular technologies. The story, however, is more complex. New ways of seeing metabolism are not only the result of, but also the drivers for, modern molecular technologies. As modern biomedical fields dive into the inner workings of the body, attempting to discover the molecular basis of cells and organisms, they are confronted with increasingly complex relationships and associations. As such, the increasing complexity of this biological world necessitates new technologies—like NMR and multivariate statistics, whose centrality to metabonomics I detail in the section that follows—to decipher complex molecular relationships.

Entanglements Between Biology and Practice

Midway through my fieldwork I sat in on an informal graduate student seminar, in which lecturers in the BMM provided advice on how best to carry out metabonomics experiments. One of the first seminars was on the topic of data analysis practices, and was given by Hector Keun, a senior lecturer whose research focused on the metabolism of cancer. Wanting to provide an easily accessible description of complex statistics to the graduate students, Hector began to develop an analogy between multivariate statistics and a supermarket. He said to the students: "Let's imagine we're gathering data from a

[45] To date, genomic approaches have been the most widely used method for biomarker discovery. Genomic markers are the main type of biomarker approved for commercial and medical use by the U. S. Food and Drug Administration (U.S. Food and Drug Administration 2013), although other "omics" fields such as transcriptomics, proteomics, and metabolomics are becoming increasingly established in biomarker research.

supermarket, right? In the land of, I don't know, metabolomia or something like that. There are some very strange behaviors going on in this supermarket that people want to analyse."

Hector asked his students to imagine an experimental scenario in which individual customers were buying apples, oranges, carrots, and cabbages—which are referred to as "variables" in statistics terminology—in a supermarket. In this particular supermarket, some customers only bought apples and oranges, while other customers only bought carrots and cabbages. Hector then asked the students to consider the pattern or relationship that was occurring among customers in the imaginary supermarket. Drawing diagrams on a white board, he described how this supermarket scenario was ideal for understanding multivariate approaches to data analysis. A univariate approach—which looks at single variables in isolation—would graph each customer's purchases on four different axes, noting the various numbers of apples, oranges, carrots, and cabbages they had purchased. But a multivariate approach—which looks at multiple variables in combination—would graph the customers' purchases in terms of the categories of "fruits" and "vegetables." This approach invents two new variables, which are referred to as "factors," and graphs each person's purchases along two new axes, instead of the original four (see Figure 18).

	Apples	Oranges	Carrots	Cabbages
Person 1	3	3	0	0
Person 2	0	0	3	3

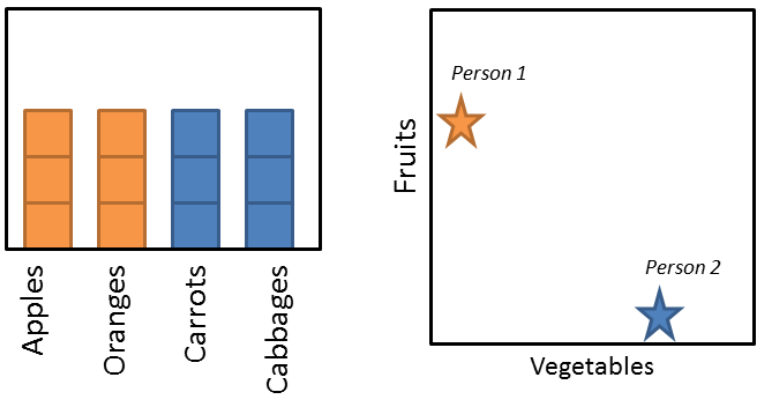


Figure 18: Multivariate statistics, as understood analogy with a supermarket. A diagram showing how multivariate statistics can be understood in terms of a grocery store analogy. The data from two people presented in the table (top) can be visualized as a univariate graph (left) with four variables, or a multivariate graph (right) with two variables.

Using this supermarket analogy, Hector emphasized that multivariate approaches to statistics simplified data analysis by taking four variables (apples/oranges/carrots/cabbages) and swapping them for two factors (fruits/vegetables). Multivariate approaches also tried to capture complex relationships within data, allowing researchers to cope with the large volumes of information generated in metabonomics experiments. They allowed researchers—in a process that is formally referred to as "dimensionality reduction"[46]—to examine the combined properties and relationships between multiple variables and samples. Looking at "tens or hundreds or thousands of [things] at the same time" was incredibly difficult with univariate statistics. Concluding his seminar, Hector proclaimed: "To me, [multivariate statistics] does something so different from univariate analysis that you can't really compare it."

As the previous vignette indicates, a large portion of metabonomics research is spent developing and using multivariate statistical technologies, which are used to transform biochemical data into biomarker panels. Metabonomics researchers take the information generated in NMR and MS experiments and, applying computational techniques, convert it into mathematical information (Barton 2011). This information is analyzed with advanced computer tools: it is standardized and processed into enormous data tables, and is subsequently analyzed by specialized algorithms and computer programming tools. Though a diverse array of computational methods and technologies are used during this data handling, metabonomics *researchers use multivariate statistical practices—which in reality make up a diverse array of technologies and techniques—to make sense of and interpret biochemical data*. This type of statistics, as I examine in this section, measures hundreds of variables at once, rather than individual variables in isolation.

[46] As Hector described: "So…more formally we say the simplification is dimensionality reduction. Okay, so we can think of each of these [variables] as representing potentially a different dimension in the graph. And as soon as you go beyond three you just can't think about it."

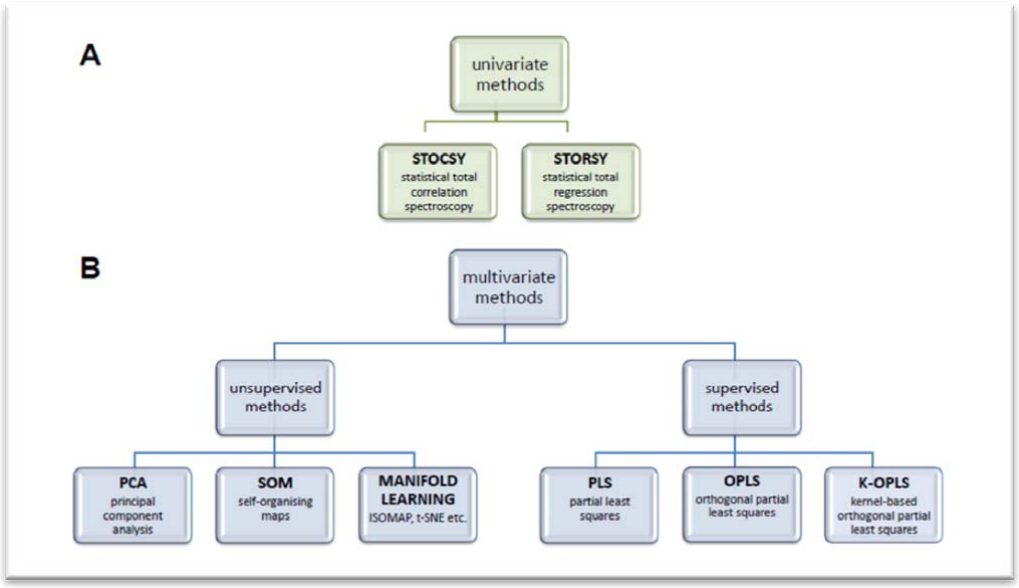


Figure 19: Univariate and multivariate statistical methods.
Figure taken from a doctoral thesis, showing the variety and hierarchy of statistical methods used to analyze metabonomics data. The methods are divided into univariate statistics (A) and multivariate statistics (B), and are further divided into unsupervised (left) and supervised methods (right).

Multivariate statistics are strange and mystifying at first glance. They are subjects of research that defy three-dimensional logic, and are challenging not only for social scientists, but also for many classically trained biologists. Within metabonomics, researchers make the broad distinction between univariate and multivariate forms of statistics (see Figure 19). Univariate statistics are defined as those methods that measure single variables at a time. Multivariate statistics, in contrast, are defined as those methods that make sense of combinations of variables, or in other words, highly-dimensional data[47].

The distinction between the two types of statistics, however, is not so clear cut. Though univariate and multivariate statistics entail different practices and approaches to data

[47] As an excerpt from a metabonomics textbook reads:

> In biology, as well as in other branches of science and technology, there is a steady trend towards the use of more variables (properties) to characterize observations (e.g. samples, experiments, time points). Often, these measurements can be arranged into a data table, where each row constitutes an observation and the columns represent the variables or factors we have measured (e.g. intensities at a specific wavelength, mass-to-charge ratio, NMR chemical shift.) This development generates increasingly complex data tables, which are hard to summarize and overview without appropriate tools. (Lindon et al. 2007:171)

analysis, their evolution and relationship has not been thoroughly explored by historians, who have tended instead to focus on the rise of informational ideas and metaphors (Kay 2000) or computing methods in biology (November 2012). Though the histories of statistical techniques such as PCA and PLS can be traced to specific individuals and historical moments (see Chapter 2), the history of multivariate statistics relates broadly the rise of computing and to the development of mathematical techniques after World War II. As these developments generated increased volumes of complex data, researchers were faced with increasing "numerical and quantitative complexity" (Esbensen et al. 1990:390) of information, and turned to computing techniques that enabled them to tackle such problems with increased computational scale and power (Edwards 2010:171-72). Multivariate statistics as a whole, therefore, is interwoven with many fields, people, and techniques.

To generate biomarker panels with multivariate statistics, metabonomics researchers must first create biochemical information. To do so, they analyze biological samples—urine, blood plasma, or stools—using NMR and MS technologies (see Chapter 2). Though both technologies are central to metabonomics research, they have different strengths and uses. NMR detects a smaller range of biochemicals than MS, and has the capacity to identify several hundred biochemicals. NMR is typically regarded as more reliable and robust, producing consistent results and information across different machines and repeat experiments. MS, in contrast, detects a significantly wider variety and resolution of biological information—identifying up to 5000 biochemicals—and can provide more nuanced data about the differences between healthy and diseased samples. MS machines, however, are difficult to standardize, and struggle to consistently replicate experiments across different locations and time scales, making their use in large or long-term studies challenging.

Throughout my fieldwork, metabonomics researchers explained that they were "comfortable" and "familiar" with one analytical technology over the other. One researcher claimed that he could glance at NMR data and immediately recognize certain biochemicals, or say whether the data was "good or bad or interesting." However, if he glanced at MS data, it was no longer intuitive, and he quickly became overwhelmed by the large volume of information MS experiments generated. Another researcher claimed, in contrast, that analyzing NMR data was "just a mystery really," and that MS was more interesting and easy to do. She said: "There are groups of molecules you do get to know. If you run serum, you kind of know what it should look like, and that there's the phospholipids and other lipids." Though NMR and MS data conveyed similar biochemical information, their appearance—

and thus their familiarity to and use by researchers—was different. Researchers became enskilled in the use of particular machines, learning to manipulate and see data in expert ways (Cohn 2007; Grasseni 2009). Thus, NMR and MS were institutionalized technologies for generating biochemical data in metabonomics, but were used in situated and interpretive ways.

Overall, the multivariate statistical methods used to interpret biochemical data exist within the metabonomics laboratory as a diverse set of technique and tools. Some form the core of laboratory work, while others are custom-made by those researchers who are more adept at computer programming. Researchers broadly carry out two types of statistical analysis: supervised or "hypothesis testing" analysis, or unsupervised or "hypothesis generating," analysis (Keating et al. 2012) (see Figure 19). These approaches are exemplified in two common statistical methods, the unsupervised technique PCA, and the supervised technique PLS. In addition, researchers use many other statistical methods, which include self-organizing maps and cluster analysis (unsupervised), and non-linear methods like orthogonal partial least squares regression (supervised)[48]. As one researcher commented: "Because the datasets are so complex, different [methods] focus on different aspects of the data. And it's like looking at a three-dimensional object from different two-dimensional perspectives."

To carry out multivariate statistical analyses, metabonomics researchers rely on a variety of computer-based computational and visualization tools. Researchers typically use either the commercially-available software SIMCA-P+ (see Chapter 2), or the computer programming tools R and MATLAB[49]. Like NMR and MS technologies, these computerized tools are highly institutionalized, but are also used in situated and interpretive ways. Those researchers with a limited knowledge of computer programming use the "push button" functionality of SIMCA-P+, which is effectively a black-boxed and commercially available statistical analysis program, to carry out basis forms of multivariate statistics like PCA and PLS. Most researchers in the laboratory also carry out several basic commands in MATLAB,

[48] Researchers acknowledge that these tools are also used for data analysis in other post-genomic fields like proteomics and genomics (Nicholson et al. 1999; Lee et al. 2008)

[49] R was developed in the early 1990s by researchers at the University of Auckland as a statistical programming language and environment. It is considered an open-sourced version of the program "S," which was developed at the Bell Labs using early versions of Unix and FORTRAN. Similarly, MATLAB was developed by researchers at the University of New Mexico as an alternative to FORTRAN, and became a commercially available software package in 1984 through the company MathWorks.

often with the aid of local "experts" in the BMM who have a background in computer science, and who can use the full functionality of MATLAB and R to produce a variety of personalized, highly-tailored algorithms. With these computer-based tools, researchers produce visualizations of multivariate data in two- and three-dimensional graphs, in which each point on the graph corresponds to a sample rather than a variable[50]. Doing so makes it possible for researchers to visualize data which typically involves hundreds to thousands of variables, and in which each variable can be described along one hundred axes or dimensions.

Once metabonomics data is transformed into workable multivariate statistical tables and visualizations, researchers use another series of tools and techniques to convert data into knowledge of metabolic compounds and pathways. To do so, researchers try to identify, largely without the use of commercially-available software, which variables correspond to known biochemical compounds. This involves matching the chemical and physical properties of variables—chemical shift in the case of NMR, and mass fragmentation patterns in the case of MS—to information available in metabolite databases such as the Human Metabolome Database (HMDB) (Patti et al. 2012). Because a database match is only a likely match, researchers must carry out subsequent experiments to confirm the presence of particular compounds, often involving commercially-available biochemical standards. Consequently, once researchers identify those metabolic compounds that are present in their experiments, they try to relate them to metabolic pathways, with the help of either published papers or metabolic pathway databases like the Kyoto Encyclopedia of Genes and Genomes (KEGG) (Metabolomics Society 2013).

This process of identifying metabolites and pathways is not without its challenges and quirks (see Chapter 5). For example, one researcher named Hannah explained her struggles to find biomarkers for various diets and foods, despite repeated experiments and data analyses. Hannah's research focused broadly on techniques to measure individual responses to food, and involved numerous experiments to identify those biochemical compounds that corresponded to the metabolism of—and could be used to monitor the ingestion of—certain

[50] As an article in the journal *Nature* describes, multivariate statistics within metabonomics research is:

> A collection of techniques in which the intensities of peaks in a spectrum are used as coordinates in multidimensional plots of metabolic activity. This allows distinctive patterns in the data to be identified more easily than by looking at the original spectra. The multidimensional plots can even be reduced to two- or three-dimensional graphs, to help visualize any clustering of points that might be used to characterize the data. (Nicholson et al. 2008:1055)

food items including fruit and fish. For her experiments, Hannah recruited volunteers within the BMM for a small study, in which she prescribed a series of regimented daily diets involving the ingestion of large quantities of fruit or fish. While she was easily able to find biomarkers for fruit—after prompting, to her participants' dismay, the ingestion of six apples per day—finding biomarkers for fish was not so simple. Despite giving her participants strict instructions about the preparation of fish, which involved frying salmon in a frying pan with particular ingredients, Hannah was left with a mystery compound that she was unable to identify. To investigate the compound further, she conducted a number of follow-up experiments in which she tested, for example, the residue left by Teflon pans during frying. In the end, however, Hannah could not identify the biomarkers for fish, and was left without insight into the biochemical processes involved in fish metabolism.

Ultimately, researchers use combinations of biochemical , statistical, and computational technologies to draw biological conclusions about their work through the production of biomarker panels, which exist in specific and situated forms. Generating iterative multivariate statistical models of biochemical data, metabonomics researchers identify biomarker panels as those entities which are statistically linked to differences between experimental groups of samples. These biochemical markers are identified according to the molecular signals and patterns, and correspond to combinations of specific biochemicals such as lactate, hippurate, and glucose. Thus, the existence of biomarkers panels as combinations of multiple biochemicals highlights the centrality of multivariate statistical practices to the field of metabonomics. Multivariate statistical methods allow metabonomics researchers to grapple and cope with the otherwise unimaginable complexity of metabolism, and to track the changes happening concurrently across hundreds of biochemicals or chemical pathways. Consequently, biomarker panels emerge as a way for metabonomics researchers to cope with the "data deluge" of modern biomedical research.

It should be said briefly that the use of biomarkers in the study of health and disease, as researchers in the BMM themselves concede, is not a new concept. Before the rise of post-genomic science in the late 20th century, biomedical researchers used "mundane" biochemical entities like cholesterol, glucose, and body mass index (BMI) to comment on bodily conditions like high blood pressure, diabetes, and obesity (Metzler 2010). With the rise of genetic technologies, researchers began to associate the term "biomarker" with molecular entities, for example in tests developed for the single-gene mutations, epigenetic modifications, or gene amplifications/translocations underlying rare diseases and cancers

(Wistuba et al. 2011). Many of these "mundane" biomarkers are univariate or bivariate in nature: their relationship to health conditions is determined by the measurement of one or two variables, respectively.

What is new in the use and production of metabolic biomarker panels, researchers in the BMM assert, is their connection to multivariate statistics and holistic views of biology. Because biomarker panels consist not of one but of multiple molecular entities, biomarker panels represent *new forms of and practices for investigating health and disease*. To this end, researchers in the BMM sought to generate biomarker panels for complex diseases like obesity and colon cancer, but were also interested in finding biomarker panels for more esoteric conditions. A group of French students joked one afternoon, for example, that they wanted to generate biomarkers for the best aged and best tasting French cheeses, as a way to investigate "cheese omics." In doing so, they drew on a range of metabonomics research devoted to assuring the quality of commercially-produced beers and wines (Gougeon et al. 2009; Skogerson et al. 2009; Cajka et al. 2011), but also described their interest in finding biomarkers for "gastronomical pleasure" and "tastiness." Though the students were not altogether serious, their comments reflected the ways in which metabonomics gives rise to new technologically-mediated ways of investigating complex and embodied conditions, which have not previously been considered as lying within the realm of biological research.

Thus, the multivariate statistical practices that produce biomarker panels represent interesting and potentially novel ways of investigating biology, metabolism, and "life itself" (Rose 2007). They are entangled with notions of metabolic complexity, but cannot be analyzed in separation from the statistical practices with which they are entangled. As such, in the sections that follow I explore how biomarker panels and multivariate statistical practices are co-produced, such that *biomarker panels do not just represent a biological reality of molecular complexity, but rather intervene into and create one*. Consequently, situated ways of visualizing and understanding metabolism must be understood in relation to the practices which facilitate views of metabolic complexity.

Naturalizing Multivariate Approaches to Biology

Multivariate statistical practices, as I discussed in the previous section, form the core focus and productive effort of research in the metabonomics laboratory. Scientists use multivariate statistics to grapple with complex biochemical information, looking for patterns and differences among their data. Multivariate statistics, however, represent one of many

approaches to data analysis. Researchers in the field of metabonomics deliberately choose to carry out statistical practices that are multivariate in nature, as these are deemed capable of accounting for multiple biological processes and endpoints. Ultimately, this "decision" to conduct multivariate statistics is reflective of the data, claims, and values with which metabonomics researchers work. Multivariate statistical practices, therefore, carry *symbolic and epistemological status*: they are technologies not only associated with—but also integral to the functioning and identity of—metabonomics as a field.

Here, I discuss how multivariate statistical practices are deemed the "correct" way of engaging with the data of metabonomics research. To do so, I discuss several instances in which metabonomics researchers highlight the value they place on multivariate, rather than univariate, statistical approaches to metabonomics data. In doing so, I explore the context in which particular practices and approaches to scientific work operate. It should be noted, briefly, that univariate statistics are used by metabonomics researchers for particular experiments or in particular instances of data analysis. While multivariate statistics are highly valued, researchers are self-reflective and scientifically rigorous with their work, such that they judge the appropriate use of different kinds of statistical methods. Thus, by examining the juxtaposition between univariate and multivariate statistics, I show how statistical practices are imbued with and reflective of values and qualities, which are embedded within the broader context and agenda of metabonomics research.

During the latter half of my fieldwork, I attended a scientific presentation given by a biochemist named Margaret, who was involved in a project trying to characterize the molecular composition of breath samples. Working with a machine that carried out selected ion flow tube mass spectrometry (SIFT-MS), Margaret analyzed the biochemicals present in various gases, including breath condensate. Though she worked on experiments that would typically fall under the realm of "metabonomics," she was not considered to be a metabonomics researcher by other members of the BMM. Because Margaret was employed by the Department of Anaesthetics within St. Mary's Hospital, she belonged to a different "epistemic culture" (Knorr-Cetina 1999) than the BMM. She did not have access to the customized programs and algorithms used to analyze data in the BMM, and instead interpreted her SIFT-MS data using the instrument manufacturer's black-boxed commercial software, which included a database of pre-given standards and compounds.

As Margaret gave her presentation, the metabonomics researchers in the room became increasingly agitated and asked a number of challenging, borderline aggressive, questions. I watched as Margaret showed slide after slide of bar graphs, with which she attempted to explain health and disease in terms of two, maybe three, biochemical compounds (see Figure 20). On the one hand, the research that Margaret described reflected generally poor scientific practice: she had taken the experimental apparatus that produced her results for granted, and had not sought to probe how the various parameters of the SIFT-MS machine had influenced her data. But on the other hand, the research Margaret described reflected a different set of experimental practices, which conflicted with the BMM's values about the correct forms of data and biological thinking. Ultimately, Margaret had presented her data as univariate, as involving a limited number of biochemicals that she was looking at in isolation. In response, the metabonomics researchers in the room asked her why she had not used multivariate statistics to analyze her data, which to them obviously involved complex pathways and multiple biological endpoints.

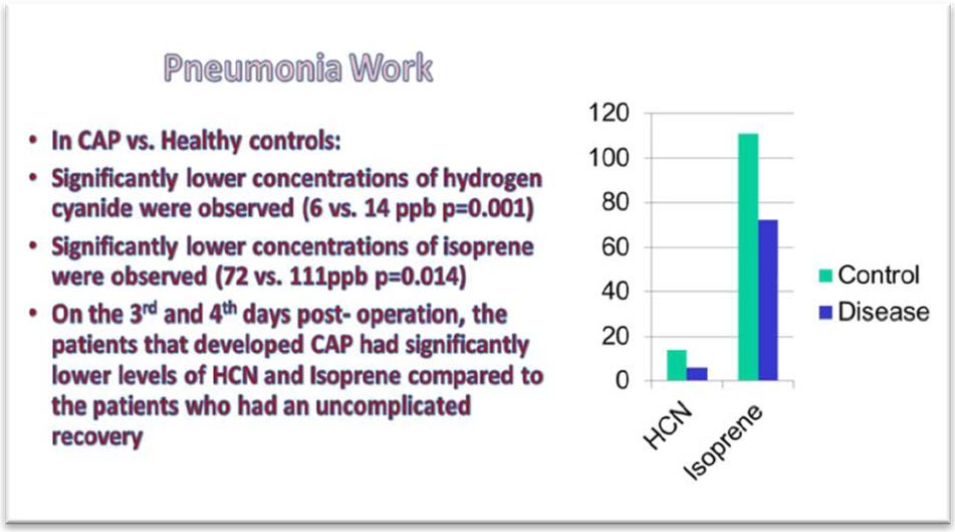


Figure 20: Margaret's presentation on SIFT-MS.
A Powerpoint slide taken from Margaret's presentation to the metabonomics researchers, showing a univariate rather than a multivariate analysis of data. Note how only two metabolites (variables) are examined, and how p values—methods common to a univariate approach to data analysis—are used.

Several days later, I asked a post-doctoral researcher Sarah—whose work also features in Chapters 4 and 5—what she had thought of the presentation. Without hesitation, she said that she had been shocked by how Margaret had presented and explained her data.

Sarah was surprised that Margaret had not sought a more in-depth understanding of the instruments that she had used, or of the mechanisms underlying her work. Margaret had made too many assumptions, Sarah said, modelling only one or two compounds at a time. In order to evaluate the significance of Margaret's work, Sarah needed more details: she needed to see all of the chemicals present, the timescales on which they were being detected, and the accuracy of the signals produced by the machine. She said to me: "I want to see the masses, I don't just want to see things going up and down [on graphs]!"

More importantly, Sarah was appalled that Margaret had analyzed her data using univariate statistics. Approaching biochemical data in this way was unheard of among metabonomics researchers. It provided a view of the data which, Sarah felt, was not true to the complex biological processes that the research was trying to describe. Compelled to address this fundamental flaw in Margaret's research, Sarah met with Margaret for several hours after her presentation to go over her data analysis, and to show her how she could improve her work by using the multivariate statistical models that were common to researchers in the BMM. Sarah explained to me that it was imperative Margaret learn to determine the full range of biochemicals contained within her samples, and to conduct multivariate statistics on her data. Otherwise, she said: "If Jeremy Nicholson [the head of the lab] saw this research, he would fire Margaret on the spot." Multivariate analysis, according to Sarah and other researchers in the laboratory, was a "good" and "correct" type of analysis.

As this example shows, univariate statistics are seen as inappropriate methods and technologies for metabonomics research. On the one hand, the rejection of univariate methods reflects the ways in which metabonomics researchers are indoctrinated into the "correct" ways of doing metabonomics research. Although metabonomics researchers enter the laboratory from a variety of disciplines—microbiology, analytical chemistry, statistical methods, veterinary medicine, to name a few—they are taught to think about and practice biological research in particular ways. In seminars, workshops, interactions with peers, and meetings with supervisors, they are introduced to particular technologies and techniques, all of which emphasize the importance of viewing metabolism in terms of multivariate biochemical information. As such, multivariate statistical practices invoke "previous dealing and cumulative practical know-how" (Alac 2008:503), as well as notions of authority and control that come with understandings of the "right" ways of doing and performing experiments (Shapin et al. 1985). These highly regarded ways of seeing, therefore, are the result of laboratory values and standards that operate within distinct socio-technical networks.

Multivariate statistics, however, are not just a product of the "social construction" of scientific research, in which sociocultural and political elements creep into scientific practice (Latour et al. 1986). Rather, the embrace of multivariate over univariate methods is also an indication of the complex nature of metabonomics data, and of the ways metabonomics researchers feel they must engage with the complexity of metabolism. As Jane Calvert has documented in her work with systems and synthetic biologists, many researchers recognize that biological complexity "is inescapable" (Calvert 2008:393), and that, rather than being reduced, it should entail the development of new methods to embrace this complexity. In a similar way, metabonomics researchers' insistence on using multivariate statistics could be seen as the desire to engage with and capture—rather than reduce or trivialize—metabolic complexity.

To this end, I spoke to Judy Fonville, a chemometrician from the Netherlands who was finishing her PhD at the time I was doing my fieldwork, about the entanglements between multivariate statistics and metabonomics. As Judy explained her research, she reminisced about the initial project she had been given when she arrived at the BMM in 2007. Its goal was to explore how molecular techniques could better establish the biological mechanisms of obesity, revealing whether people were "fat on the inside and fat on the outside." To do so, the project examined if magnetic resonance imaging (MRI) could be used to determine the fat composition of inner tissues, and whether these molecular measures corresponded to external indicators of fatness such as body mass index (BMI) or waist-to-hip ratio.

As part of the project, Judy had been tasked with using multivariate statistical methods to examine the data generated by MRI scans. MRI scans are commonly used to examine injuries in tissues, or to understand the physiology underlying bodily conditions, and produce complex images of internal tissues and structures. Beyond their imaging capabilities, MRI scans can also be used to generate low-resolution NMR spectra of the tissues they are scanning. Their relatively low field strength—the resolution of the signals that an NMR spectrometer can show—makes them suitable for identifying macromolecules like water and fat, but not for identifying the typical number and range of biochemicals seen in metabonomics experiments.

Judy recounted to me that, after one year of research, she had been forced to abandon the project and find another focus for her PhD. While MRI scans generated biochemical data

in the form of information about the fat and water content of various tissues, they were not suited to analysis with multivariate statistical tools. According to Judy, the MRI data was too low resolution, in that it did not show the multitude of biochemicals that metabonomics researchers are used to working with. Judy said: "So the data about the fat, basically you have two peaks, a fat and a water peak. So doing multivariate statistics on two peaks is a little bit silly…you can [just] use a classical statistics type thing." Thus, the complexity of the methods Judy used to analyze data did not correspond to the complexity of the MRI data she was attempting to manipulate.

Because of these challenges, Judy had abandoned the project, despite her interest in working with MRI technology. It was not worth her time to work with the data that could be analyzed with univariate statistics—with a simple correlation between levels of fat and water among different people—in an Excel spreadsheet. Thus, the MRI data was not suitable for use in metabonomics experiments, in which researchers used advanced multivariate statistical methods and complex "handmade" computer algorithms. Though the MRI data contained useful and biologically-relevant data, it did not fit with the methods and goals of metabonomics research. Judy's decision to focus on other research projects, therefore, was intricately tied to her desire to focus on data that espoused the complexity of metabolism.

In conclusion, the examples of Margaret and Judy highlight the value that metabonomics researchers place on the use of multivariate statistics in metabonomics experiments. Multivariate statistics, although they represent one of many approaches to the analysis of metabolic data, are highly naturalized—and to a certain extent institutionalized—methods for the analysis of multifactorial biochemical data. The value placed on multivariate statistical practice can be understood in relation to particular ways of seeing and understanding biology, which are situated within networks of materials, people, and standards. Multivariate statistics are not simply a technology or method for coping with large amounts of data, but rather are intrinsic to the methods, ideas, and values of metabonomics as a field.

Enacting Metabolic Complexity

Multivariate statistical practices—in the sense that they are a combination of methods, ideas, and values—entail new and situated ways of understanding and enacting metabolism. As I show in this section, it is not only that multivariate statistical practices create new understandings of metabolism, but also that these visions of metabolism drive the creation

and use of complex and multifactorial informational practices. *Metabonomics needs to be based around multivariate statistics*, as notions of complexity cannot function independently from multivariate statistics. Metabonomics researchers simultaneously draw upon and produce multivariate statistical practices, and with them complex biological meanings. Ultimately, just as the linear relationship between biomarkers and metabolism is being revised through the emergence of biochemical and statistical practices, so is the linear relationship between notions of metabolic complexity and multivariate statistics.

In the first half of my fieldwork, I worked closely with Claire Merrifield, a former immunologist turned metabonomics researcher, conducting experiments and learning about the practices common to metabonomics research. Claire's project examined the relationship between the metabolism and immune function of pigs, and was situated within a growing arena of biomedical research that seeks to understand the ways in which gut bacteria influence health[51] (Nicholson et al. 2005; Nicholson 2006). In her experiments—which involved analyzing the urine, blood, and liver samples from pigs—Claire was trying to identify biomarker panels that corresponded to "healthy" gut bacteria. These experiments built on the idea, long established in biology and recently popular with food companies, that positive changes in gut bacteria can be induced by the ingestion of probiotics[52]. As a result, Claire was detecting, measuring, and modelling several biochemicals—named hippurate, lactate, and p-cresol—in urine, which were produced by the combined metabolism of pigs and bacteria. Ultimately, she hoped to use these biochemicals to devise a biomarker panel that corresponded to "healthy" gut bacteria.

To determine the biochemicals involved in the biomarker panels for "healthy" gut bacteria, Claire spent many hours in front of her computer. To investigate potentially important biochemicals, she generated a series of multivariate statistical models—using, for example, PCA—that took the myriad of numbers from her experiments and transformed them

[51] Using a systems biology approach, and involving studies in diverse post-genomic fields, scientists have been researching the composition of the microbial colonies living within our bodies, their role in immunity and disease, and the ways that they can be modified to promote health (Nicholson, Holmes et al. 2005; Nicholson 2006).

[52] The probiotics that Claire was studying were produced by the company Nestle, with whom the laboratory had a research alliance. Claire, along with several other graduate students and academics in the laboratory, was part of the "Nestle-ICL Research Alliance." Since 2009, the Alliance has been investigating how metabolism—including human-microbial interactions—can be used to understand the ways in which nutritional solutions (such as probiotics) might be used to combat obesity (Imperial College London 2012c).

into meaningful information. Such unsupervised methods were "hypothesis generating," in the sense that they sought to generate or discover meanings, rather than to prove preconceived relationships, through statistical analysis[53] (see Chapter 5). In essence, Claire used these statistical tools and techniques to investigate which factors were responsible for the differences between the samples. She wanted to see if the differences corresponded to experimental factors like the sex or age of the pigs, or to particular biochemicals.

One afternoon, as Claire sat at her computer flicking back and forth between multivariate algorithms and the two dimensional graphs they produced, she discussed the contrasting use of statistics in the more "classic" field of immunology versus the "newer" field of metabonomics. Due to a collaboration with another university, Claire was analyzing not only metabolic data, but also immunological data in the form of cytokine assays. Because these assays measured the levels of only ten different immunological molecules—in contrast to the hundreds of biochemicals detected during metabonomics experiments—Claire could use simple, univariate statistical calculations such as the t-test or analysis of variance (ANOVA) to make sense of her immunological data (Genser et al. 2007).

These univariate statistics, Claire explained, were used to confirm relationships and meanings that were visually apparent in data. In other words, univariate statistics were used to verify that the patterns observed via other experimental means were statistically significant. Remembering back to my four years of experience as an undergraduate in an immunology laboratory, the researchers in my laboratory had also used univariate statistics in this way. We had conducted statistics on our experimental data—which showed obvious visual differences between experimental and control groups—to prove that our findings were significant and had not occurred by chance. As such, univariate statistics had been a tool to confirm, rather than explore, the meanings that were otherwise obvious in the immunological data.

Claire, in her discussion of univariate statistics and immunological data, illuminated the nature of the multivariate statistics in metabonomics research. She said: "All you're doing [in immunology] is using [statistics] to confirm something that you can see visually.

---

[53] There is a strange commonality between unsupervised statistical analysis and ethnographic participant observation methods, in the sense that both eschew a logic of "hypothesis generation." In this sense, both typify bottom-up approaches. While unsupervised statistics let mathematical relationships emerge from the data, ethnographic methods let social conclusions and key areas of investigation emerge by being open to the richness and variety of actions and interactions within data.

Whereas when you go into metabonomics you can't often see visually what's happening, so you have to use the statistics to then be able to see what it is to look for." Claire could not approach metabonomics data, with its hundreds of variables and measurements, with simple univariate statistics. Instead, she needed to use advanced multivariate statistical tools, without which she would have been drowning in her data, and unable to make sense of the myriad of biochemical information she had generated in her metabonomics experiments.

However, even with the aid of sophisticated multivariate statistics, Claire labored to understand how to interpret—and translate into biomarker panels—the biochemicals going up and down over time in complex patterns (see Chapter 5). In engaging with huge quantities of data, she struggled to understand what information was meaningful. Nonetheless, Claire was able to evolve an understanding by working through multivariate statistical and computerized approaches to data in varied and interpretive ways. Consequently, Claire's active involvement in the process of statistical analysis and the production of biomarker panels highlights the dynamic element of "play" involved in metabonomics research. Researchers develop and forge understandings of their data by producing, amending, and revising statistical analyses and visualizations (Cohn 2007). They participate in "a dynamic interaction between trying to…generate an image that best fits what is being expected, and discovering an image that in some sense tells its own story" (Cohn 2007:99).

But this process of data analysis is also governed by the standards and protocols that pervade metabonomics research. These direct researchers to use some methods, and to produce some statistical forms of knowledge, over others (Timmermans et al. 1997). For example, one researcher who was tasked with cleaning up and standardizing the data from the molecular epidemiology INTERMAP project (see Chapter 6), a multi-year endeavour that involved the data analysis efforts of many different researchers, explained that choices over which approaches and algorithms to use were highly political. Each researcher, he said, had their own way of doing data analysis, and felt threatened that alternative "improved" data practices might invalidate their own results. Thus, the production of biomarker panels through multivariate statistical practices involved a combination of protocols and standardization, as well as interpretation and improvisation.

Like many other researchers in the lab, Claire used multivariate statistics to explore and create the meanings that were "hidden" within her metabonomics data. The relationship she found between healthy gut bacteria and biomarker panels was not pre-given or waiting to

be "discovered" as a scientific fact. Instead, it was enacted through the multivariate statistical practices of the metabonomics laboratory, from which biomarker panels emerged as complex biological entities (see Chapter 5). As Natasha Myers (2006) has similarly noted in her work on protein modelling, researchers use their experiences developing computer models to shape their understandings of biology. She writes: "X-ray crystallographers' extensive assemblage of machines – including metaphors and interactive digital visualization media – [are] collectively geared to produce and interpret atomic resolution models of proteins as molecules" (Myers 2006:11). Biomarker panels do not simply present an objective view of biology, which tries to make sense of the molecular information generated in metabonomics experiments. They are also entangled with and co-produced by the multivariate and biochemical practices inherent in metabonomics research.

Thus, Claire used multivariate statistics to enact a view of metabolism that was dynamic, multiple, and interconnected. She viewed metabolism as the combined effect of not only genes and environment, but also gut bacteria. As Claire commented:

> It's the result of the fact that biological processes are multivariate in nature. You can never take something in isolation, really…when you look in biological systems, everything is dependent on each other…It's like saying my leg hurts, and somebody saying did you eat a banana this morning, and then not looking at the fact that they're actually trapped under a rock. It's just, if you only look at one thing, you're limiting yourself by the information you've got.

Accordingly, the metabolites she detected in her experiments had resulted not only from the metabolic processes under the control of the human genome, but also the metabolic processes under the control of the diverse colony of bacteria occupying the large intestines of pigs. Moreover, these metabolic processes were interdependent rather than independent: they acted in unison to coordinate the digestive and overall health of pigs, and to regulate the metabolism of the organism as a whole.

Ultimately, this section shows how metabonomics researchers like Claire use the active process of working through multivariate statistical analyses to develop multifactorial understandings of metabolism. As such, *researchers use, need, and reproduce multivariate statistics* to make sense of a biological world that is viewed in terms of complex systems, instead of single genes or isolated pathways. Their practices allow them to explore the complex relationships in metabolic pathways, but their practices also enact an interconnected and multiple view of metabolism. It is not only multivariate statistics, but also "biomarker panels" that emerge as the natural and obvious ways for understanding health and disease.

Biomarker panels, therefore, are not neutral or natural objects, but are instead are the result of local practices and values. Researchers emphasize the need to use multivariate statistics to explore certain types of metabolism, but such views of metabolism do not exist independently of the methods used to explore them.

Discussion

Biomarker panels, and the multivariate statistical practices with which they are entangled, entail new ways of interacting with and enacting metabolism. Biomarkers, as one of the ways that metabonomics researchers grapple with health and disease, are entangled with specific, situated biochemical and statistical research practices and ideologies. On a broader level, this reflects the ways in which biological meanings and values are dynamically entangled with technological, and in particular statistical, practices. Viewed in this way, researchers both draw upon and produce multifactorial, complex understandings of metabolism through biochemical and statistical practices. Echoing Peter Chow-White and Miguel Garcia-Sancho's (2011) observations of the relationship between genomics and computer research, biology and statistical technologies interact and reciprocally shape each other. This challenges not only the conceptual and disciplinary boundaries between metabolism, statistics, and biochemistry, but also linear accounts of biomedical development and innovation.

In particular, this chapter highlights the conceptual centrality of multivariate statistics to metabonomics laboratory practices, and also to biomedical practices more broadly. Statistical practices have largely remained absent from the anthropological and STS record on biomedical science, and yet play an increasingly large role in the way that biological information is explored, reasoned, produced, and translated into clinical settings. Thus, many of the statistical methods and techniques that are discussed in this chapter and in the dissertation as a whole are not specific to the field of metabonomics, but rather apply more generally to post-genomic research in fields such as proteomics, epigenetics, and systems biology. The entanglement between data practices and complex understandings of biology can therefore be generalised more broadly to other areas of research in which multivariate statistical practices are also central.

On the whole, multivariate statistical techniques allow researchers to cope with the increasing amount of data—and therefore the increasing complexity—that is produced through metabonomics research. But at the same time, statistical techniques are integral to

the ideas and identities of metabonomics research. They are normalized ways of investigating and making sense of biology, and are supported by historical narratives and institutionalized practices. Thus, multivariate statistics exist as both a tool and an endpoint in and of themselves. They cannot be bracketed as the "technological" component of metabonomics research, but must instead be seen as an intrinsic amalgamation of techniques, ideas, and values.

Multivariate statistical practices, therefore, showcase metabonomics' *conceptual, methodological, and epistemological* commitment to complexity as "life itself." Several decades ago, Michel Foucault (1990) argued that technologies and forms of organization transformed human bodies and biological processes of life into the workable objects of biomedicine. Human vitality, referred to as "life itself," became a new site for control and power, as biomedicine began to focus on birth and death, the size and quality of populations, and health and disease. Taking up this argument, Nikolas Rose (2001; 2007) has argued that concerns with life itself have extended to the molecular level, such that processes and knowledge of genes and biological systems constitute new sites of control and intervention.

Consequently, this chapter argues that contemporary data practices make up new sites for conceptualizing and intervening into biology. Metabonomics researchers are investing in this multivariate view of the biological world as something new and exciting. In doing so, they are not only carrying out multivariate practices and producing multivariate facts, but also eschewing multivariate values about what constitute the best ways of enacting the biological world. A notion of metabolism as a multiple and complex process does not exist "out there" as an objective or "independent life world" (Mol et al. 2002) waiting to be discovered. Instead, it is actively created and enacted by scientists, and in specific, situated ways. Thus, it might seem that "life itself" is being construed as something emergent, temporal, multiple and dynamic (Rose 2013)—or is it?

Thus the question remains, of *what is this complexity with which metabonomics researchers are grappling*? As a conclusion to this chapter, I want to argue that multivariate statistical practices enact a world in which biology is not seen as a linear, one-dimensional, or "flat" set of information, code, or text (Kay 2000: as cited in Myers 2008). In contrast, multivariate statistical practices give life to metabolism by generating non-linear, multivariate, and dynamic ways of engaging with the molecular biological world. These emerging practices shift away from a linear and univariate view of biology, to embrace the

"multi-dimensional structures" and dynamic nature of biological data. In doing so, they allow researchers to grapple with complex metabolic pathways and meanings. With multivariate statistical practices, "the objects of molecular biology are becoming tangible and workable in new ways" (Myers 2008:164).

But although multivariate statistical practices allow metabonomics researchers to enact biological complexity, they do so in controlled and calculated ways. Metabolic complexity calls for new tools, technologies, and practices, such that complexity is "continually encountered; but then also seamlessly bracketed" (Wynne 2005:69). Thus metabonomics engages with complexity, but on specific—statistical and biochemical—terms, and within the controlled environment of the laboratory. Echoing Hannah Landecker's (2011) observations about epigenetics, environmental factors such as age and diet, and bodily conditions such as kidney functioning or fat absorption, are quantified and inserted into mathematical equations. Thus, the biological claims made by researchers do not always match with, or are not entirely realistic about, the capabilities of biological practices. Metabonomics researchers' use of multivariate statistics represents attempts to *engage with and enact complexity in pragmatic ways.*

This resonates in many ways with Karen Throsby and Celia Roberts' (2010) work on gene/environment interactions in studies of early puberty and childhood obesity. Such gene/environment paradigms, Throsby and Roberts claim, highlight the role that not only genetic elements but also environmental exposures play in the development of puberty and obesity. But instead of leading to a more holistic understanding of both conditions, Throsby and Roberts suggest that such approaches lead to an understanding of the environment that is flattened and quantified. "Genes" and "environment" are likened to the natural and social, respectively, such that the environment—in all of its diversity—becomes everything but genes. Consequently, studies of gene/environment interactions do not lead to a renewed focus on the role of environment in the development and prevention of disease, but rather reinforce a focus on interventions that target the genetic, molecular, and metabolic aspects of disease (Throsby et al. 2010:80). As Throsby and Roberts write: "while both puberty and obesity are seen as processes of becoming, and while the environments that are seen as creating both of these 'problems' are social…biological and metabolic processes remain central" (Throsby et al. 2010:76).

Thus, metabonomics recognizes and engages with the vitality of living systems (Canguilhem 1989; Cohn 2004), but does so in controlled and standardizable ways, and with the aid of situated laboratory tools, techniques, and practices (Hacking 1992:59). What is at stake, then, is the extent to which emerging biomedical research projects like metabonomics can capture the inherent complexity and vitality of life—the essence of what makes us human—through advanced multivariate statistics. Returning to Foucault and Rose's question of "life itself": have fields like metabonomics rendered linear and reductionist approaches to biology obsolete, or have they simply displaced or rendered them into different forms? How far does the complex thinking espoused by and enacted in metabonomics research extend, and what are its potential consequences for medical practice? Given the importance and ubiquity of multivariate statistical practices in the modern world, what consequences might a new world of biological complexity—of networks and interconnected processes—have for the ways that biomedical practitioners diagnose disease or think about medical treatment in the near future? I address this question in the chapter that follows, as I examine how metabonomics researchers grapple with the task of assigning categories of "health" and "disease" to the complex entity that is metabolism.

Chapter 4: Finding "Health" in the Metabonomics Laboratory

Introduction

I am sitting inside of the sixth floor "prep lab"—the small laboratory space where samples are prepared and primed for experiments—watching Ryan, a PhD student, handle test tubes full of cerebrospinal fluid. Wearing a laboratory coat, goggles, and gloves, he picks up tests tubes, fills them with a tiny amount of solution, pipettes up and down to mix, and repeats the process many more times. Ryan is preparing samples that he will use later that day in NMR experiments, in which the goal is to find biomarkers for the degenerative disease multiple sclerosis. His aim is to use metabonomics to detect changes in brain biochemistry that are coordinated with, or even precursors to, the physiological symptoms of the disease. In his experiments, and using biomarkers, Ryan hopes that he can establish what is "normal" and what is "abnormal" for multiple sclerosis physiology and experiments.

As Ryan describes his research, I ask if I can tag along for the rest of the day and watch him conduct NMR experiments. Ryan finishes mixing and labelling all of the tubes, puts them in a Styrofoam box full of ice, and motions for me to follow him. We exit the SAF building, which houses the majority of the BMM's activities, and Ryan takes me to the 800 MHz[54] NMR facility, an old brick building that is co-managed with the Chemistry department. He leads me through a series of sterile, fluorescent corridors, until we emerge

[54] The phrase "800 MHz" refers to the proton resonance frequency of the NMR spectrometer, which approximates the overall resolution of the data generated from the NMR spectrometer. Higher frequency spectrometers—which contain more powerful magnets to produce large magnetic fields—are able to resolve a greater number and set of features of biochemical signals. A field strength of 800 MHz is relatively powerful for a commercial machine, with the most powerful NMR spectrometer in the world operating at roughly 1000 MHz (Bhattacharya 3 February 2010).

into a large room full of expensive looking state-of-the-art computers that are used for data analysis.

Ryan motions me into an adjacent room, which houses the biochemical machinery of the 800 MHz NMR facility. Inside, the massive 800 MHz NMR spectrometer provides high resolution data about the chemical properties of solutions, and so can be used for a wide range of research activities, including investigations into the biochemical properties of proteins, and determinations of unknown biochemical structures. The machine—a ten foot tall round metal tub that looks like it is propped up on metal stilts—is surrounded by metal walkways and signs indicating the potential danger of high-strength magnets. Ryan grabs the Styrofoam box full of samples, walks up a metal set of stairs to the top of the machine, and gives me a thumbs up. I watch as he picks up a long, thin test tube about the size of a pencil, which holds a miniscule 50µl of cerebrospinal fluid, a scarce resource because it must be procured during a lumbar puncture. Attaching the tube to a “spinner,” a piece of plastic that holds the test tube, he carefully places it inside of the NMR spectrometer. Standing next to the ten foot tall apparatus, Ryan’s six foot frame looks impossibly small (see Figure 21).

Figure 21: Ryan conducting metabonomics experiments on multiple sclerosis samples using an 800 MHz NMR spectrometer.

After Ryan climbs down from the machine, he goes over to a computer and begins to set up the experiment. Typing in various computer commands, he prompts a thin green undulating line to appear on top of a black background. The line is a real-time representation of the NMR spectrum of his sample, and is focused around the region that shows "lactate," a common biochemical molecule with two peaks, which metabonomics researchers use to standardize their data. As Ryan runs his experiments—moving back and forth between the computer screen, the box of test tubes, and the enormous NMR spectrometer—he flips through the various spectra. He is comparing them by eye, quality-checking the experimental data as it comes out of the NMR spectrometer.

Despite the enhanced technology, which impresses me with its grandeur and sophistication, Ryan complains that all of his samples' spectra appear the same. Looking at the computer screen, he shakes his head and says, "They look like carbon copies of each other." Ironically, in some ways the samples are already carbon copies. Though each sample represents an individual patient, the samples appear as identical test tubes full of clear liquid, which can only be distinguished by numerical codes. They are anonymized, miniaturized representations of disease. As Ryan comments, he has no idea which samples are from patients versus controls, or whether the samples represent early or late stage disease.

Worried that the differences in the data do not seem significant, and consequently that the data may not be able to provide insight into the pathology of multiple sclerosis, Ryan says, "If you can't see that big of a difference, there's probably not much of importance." Such comments provide insight into the practices and nuances of data analysis in the metabonomics laboratory. Though researchers use various steps to analyze their data, their first steps almost invariably involve assessing experimental data by eye. Researchers scan the lines and peaks of NMR data, looking for differences that stand out—as an extra set of peaks, as a strange grouping of signals—among the various samples. By examining the spectra manually, and by relying on their eyes and learned sensibilities rather than complex computer algorithms, researchers examine the *qualitative* differences between data. But because such qualitative differences are not measurable, researchers turn to statistical methods to explore the *quantitative* differences among samples.

Continuing to look through his data, Ryan explains that in cases where NMR experiments do not yield immediate and obvious differences, he turns to a variety of multivariate statistical data analysis methods to shed light on the nuances within the data.

Showcasing his own achievements in the lab, Ryan explains the merits of a statistical technique that he developed during his first year of research in the laboratory, and which is based on the principle that differences between samples can be understood relative to a biochemically defined "biological baseline." Statistical techniques like this, Ryan explains, are used to quantitatively measure biological variation among samples. This, in essence, allows researchers like Ryan to classify samples as either normal or diseased.

As Ryan describes the complex statistical and data practices that form the core of his work, I am struck by his language and choice of words. He describes multiple sclerosis in terms of the biochemical processes that go awry, and in terms of the physiological changes or "perturbations" that lead to a state of disease. For Ryan, multiple sclerosis exists as an "abnormal" state relative to a "normal" biological baseline. It is something that can be quantitatively measured as a statistical variation between normal and abnormal samples. Expressing a similar understanding of disease, another researcher commented:

> The problem [is] how you can put a number, how you can attach a number, to physiological and pathological disruptions in metabolism. A number…so basically, rather than saying qualitatively that this is a physiological disruption and this is a pathological disruption, it's a continuous process in a sense…so put it on a continuum where you can measure it.

Echoing this line of reasoning, Jeremy Nicholson also proclaimed at a meeting: "Always we say, you have to understand normal and abnormal variation pretty well before we can say what abnormal is, right." Similarly, another researcher remarked to me during an interview: "Metabonomics is all about defining the 'healthy state.'" Taken together, these comments reflect how researchers must first establish what is physiologically normal in order to understand the biochemical processes that lead to disease. With this knowledge, researchers can quantitatively establish the point along a biological continuum at which normal turns to abnormal—or more importantly, disease.

Evaluating Health and Disease as the Normal and Abnormal

Ryan's focus on normal and abnormal as quantifiable, measurable entities forms a key starting point for this chapter. It highlights how metabonomics researchers conceptualize health and disease—the embodied and dynamic states for living and breathing organisms—in terms of the "normal" and "abnormal," and in a way that is highly biochemical and statistical. In this sense, the work of Ryan and the other researchers mentioned in this chapter highlights the ways in which the normal and abnormal form the building blocks of metabonomics laboratory life. They are the concepts and common language that researchers use, sometimes

overtly and sometimes more subtly, to articulate issues in health and disease. By examining the normal and the abnormal, therefore, this chapter explores the practices, technologies, and ideologies that metabonomics researchers use to evaluate and grapple with health and disease.

Overall, this chapter explores how the concepts of normal and abnormal—as discussed in Chapter 1—are conceptualized, operationalized, enacted, and grappled with in the metabonomics laboratory. As the works of Georges Canguilhem (1989), Ian Hacking (1990), and Michel Foucault (1990; 2003) establish, normal and abnormal play a key role in determining health and disease in contemporary biomedical practice. Normal and abnormal not only encompass a quantitative means for thinking about and measuring health, but also establish the ideals to which society and biomedical practice strives. Within the metabonomics laboratory, the concepts of normal and abnormal are highly naturalized. They appear as objective, but in reality are linked to specific and situated practices. Moreover, their conflation with health and disease is not pre-existing, but results from the active work of researchers. Thus, the importance of the categories of normal and abnormal does not lie in their overarching use as categories, but rather in the logics and practices that they involve.

Taking the centrality of normal and abnormal in biomedicine as a starting point, this chapter explores how researchers transform the inherently unstable concepts of normal and abnormal into standardized, naturalized, and quantifiable laboratory entities. The stabilization of laboratory objects and the removal of contingency within scientific research has long been a topic in science and technology studies (STS), for example in Steve Woolgar and Bruno Latour's (1986) discussion of the "inscriptions" that conceal the hard work and controversy of scientific knowledge. Similarly, the navigation of uncertainty in disease diagnosis has been an important area of inquiry for medical anthropology (Buscher et al. 2010), with scholars focusing on the ways that patients mobilize agency to deal with diagnostic uncertainty in issues of reproduction (Lock et al. 1998), or on the ways in which biomedical uncertainty is performative and generative (Whyte 1998; Petryna 2002; McGoey 2009; Street 2011).

By focusing on the stabilization of the normal and abnormal as particular scientific objects, however, my analysis adds to this line of inquiry. Beyond examining the inscription of laboratory objects or the performative aspects of uncertainty, *I question how metabonomics researchers' engagement with the contingency of the normal and abnormal*

affects their ability to comment on issues of health and disease. To do so, I focus on the moments in which researchers actively question and explore the practices underlying their definitions of "normal" and "abnormal." I show how the positioning of normal and abnormal as naturalized and objective ways for defining health and disease is something that is actively achieved rather than pre-existing or self-evident.

In the first section of this chapter, I unravel the forces at play within metabonomics research in the making and unmaking of the normal and the abnormal. I explore how the normal and abnormal are enacted relative to particular configurations of ideas, technologies, and values, such that biochemical and statistical concepts emerge as "objective" and "good" forms of knowledge about health and disease. In the second section, I explore how metabonomics researchers grapple with the inherent multiplicity—and therefore contingency—of normal and abnormal, both containing and drawing upon contingency in their attempts to understand health and disease. To conclude, I discuss the challenges that arise when metabonomics researchers cannot easily conflate notions of "normal" with "health." I consequently reflect on the extent to which metabonomics research, in conceptualizing normal and abnormal as biochemical and statistical entities, succeeds in grappling with the dynamic concepts of health and disease.

Stabilizing Molecular Definitions of Health

In this section, I discuss a case study of the work of Sarah, a post-doctoral researcher in the BMM, in order to explore how *the concept of normal is enacted in practice.* Her work, which is explicitly concerned with the statistical boundaries between the normal and abnormal, reveals how the stable and natural "facts" in formal accounts of metabonomics research are continually negotiated through laboratory practices (Adelswärd et al. 1996). By focusing on the everyday practices and technologies used in the metabonomics laboratory, I show how researchers work to measure and define the point at which the normal turns into the abnormal. Researchers, despite emphasizing the objectivity of normal and abnormal, encounter these concepts as inherently ambiguous and problematic entities, and actively work to transform them into stabilized, standardized, and naturalized entities.

Finding "Objective" Definitions of Normal

Over the course of my fieldwork, I observed research on the application of cutting-edge metabonomics technologies to the study colon cancer tissue. This research—which was

one of the translational research projects sponsored by the Imperial College National Institute for Health Biomedical Research Council (NIHR-BRC) (see Chapter 5)—attempted to determine the metabolites that could differentiate between healthy and cancerous metabolic processes in tissue. It used a technology called magic angle spinning nuclear magnetic resonance (MAS-NMR), a type of NMR spectroscopy that analyzes solid tissues[55]. Using MAS-NMR, metabonomics researchers examined tissue taken from various distances from a tumor—1mm, 2mm, 5mm, and 10cm—in order to determine those biochemical processes unique to cancerous tissue. This work attempted, somewhat problematically, to find the differences between healthy and cancerous tissue by quantitatively and "objectively" defining the boundaries between the normal and abnormal.

Sarah, the post-doctoral researcher whose work forms the focus of this section, was charged with setting up the initial machinery of the "Clinical Phenome Centre" (see Chapter 2), which was located in the Surgery and Surgical Technology wing of St. Mary's Hospital. Recently publicized in a *Nature News* article (Bhattacharya 14 December 2009), this Centre housed a £300,000 MAS-NMR spectrometer that had been custom-fitted with a robotic tissue auto-sampler by the NMR manufacturer Bruker Biospin. The overall goal of this MAS-NMR research, according to Jeremy Nicholson, was to provide a quicker and more "robust" alternative to histopathological analysis, a technique involving the manual analysis of tissue with a microscope. The MAS-NMR experiments were therefore attempting "to half or quarter" the usual 40 minutes it took histopathologists to cut, stain, mount, and interpret a sample.

Sarah had joined the laboratory at the beginning of 2011, and like many of the other researchers in the lab, had been introduced to metabonomics in an applied and international setting. After completing a PhD and post-doctoral fellowship at a research institute in Spain, she had carried out a brief internship at the BMM in 2008, with the goal of returning to Spain to set up a "metabonomics platform," a set of instruments and procedures that would allow researchers to conduct metabonomics experiments, at the institute. Several years and jobs later, Sarah had come to Imperial permanently in order to conduct clinically-relevant research

[55] "Magic angle spinning" refers to the fact that the solid samples are spun via an air turbine mechanism within the NMR spectrometer, such that an average set of chemical shift values are obtained for the solid sample regardless of its molecular homogeneity.

in metabonomics. She said: "So I said, if I have to do metabonomics, I better learn how to do it properly…the best place to go…is London."

Figure 22: MAS-NMR spectrometer (right) and robotic autosampler (left) in St. Mary's Hospital.

On a warm morning in late April, I met Sarah on the tenth floor of St. Mary's Hospital for a tour of the new MAS-NMR facility (see Figure 22). Swiping into the Surgical Wing, she brought me to an open room filled with the white noise and ticking of machinery. Handing me a lab coat, Sarah guided me past a tank-like MAS-NMR spectrometer, which stood next to a robotic "tissue autosampler," a series of trays, tubes, and flashing lights connected to a computer interface. In the adjacent room she showed me the "prep-lab," a fume hood containing the various tools and NMR parts used to load tissue samples into the MAS-NMR spectrometer.

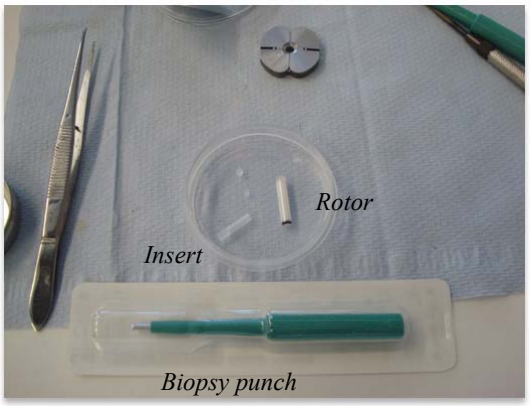


Figure 23: Tools used to collect a MAS-NMR tissue sample, including a biopsy punch, insert, and rotor.

After Sarah explained the various aspects of the facility, she invited me to watch her carry out an MAS-NMR experiment. Retrieving a sample from the -20C freezer, Sarah used a biopsy tissue punch to remove a 1mm piece of tissue, which she placed inside of a plastic disposable “insert” about the size of piece of macaroni (see Figure 23). After placing the insert into an MAS-NMR rotor, a hollow piece of plastic the size of paperclip, Sarah walked across the room to place the rotor into the tissue autosampler. This consisted of a 48-well refrigerated sample tray, which was kept at -16C in order to inhibit metabolic processes in tissue (see Figure 24). After she configured the MAS-NMR spectrometer, she stepped back to watch as the rotor shot up and away from the autosampler tray through a vacuum tube and into the NMR spectrometer. With this, the indicator light changed from red to green to signal that the experiment was underway. Together, these technologies were designed to make the MAS-NMR spectrometer suitable for use in a clinical setting, in which samples needed to be run quickly and efficiently, and with minimal experimental error.

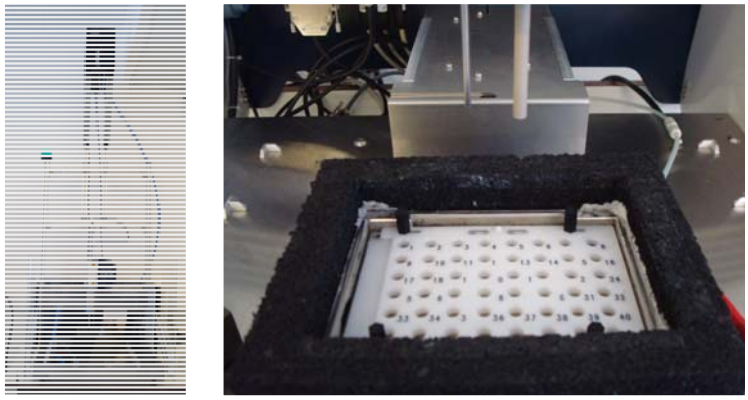


Figure 24: Autosampler (left) and refrigerated sample tray (right) adjacent to the MAS-NMR spectrometer.

Over a series of informal discussions, Sarah explained to me that the goal of her research was to help cancer surgeons improve the classification of healthy and cancerous tissue. In clinical practice, surgeons received advice by histopathologists—specialists trained to examine frozen and stained slides of tissue—as to the cancerous nature of tissues. Examining tissue, histopathologists looked to grade and stage tumors, as a measure of the appearance and spread of cancer, respectively. This process, asserted Sarah, was inherently

problematic: while some tissues were easily categorized as healthy or cancerous, others were ambiguous, making it difficult for surgeons to decide which and how much cancerous tissue to remove. As a colon cancer surgeon involved with the project commented:

> Often in… surgery, we're talking about margins, it's all about margins for cancer surgery. The margin has to be clear for you to be able to say that you've given the patient a curative operation. And sometimes during surgery, you're not sure…At the moment we're almost entirely reliant on the sampling of frozen sections, where you do it in operation, take a little bit and put it in a jar, send it off to the lab. They freeze it, slice it, stain it, and have a look at it, and try to give you an answer. And usually you get an answer within half an hour. Unfortunately the reliability and reproducibility of frozen section has come under increasing scrutiny over the last five to ten years. Its sensitivity and specificity has been repeatedly questioned.

Though histopathology was the "gold standard" of cancer diagnosis, Sarah claimed that it was "arcane." It relied on subjective, and therefore unreliable judgements by the trained eyes of the histopathologist. To illustrate her point, Sarah made an analogy between cancer diagnosis and colors: histopathologists sought to determine if tissues were red—cancerous—or blue—healthy. But the trouble was, many of the tissues ended up being purple—borderline. Thus, in relying on qualitative judgments, histopathology rendered the differences between healthy and cancerous tissue as ambiguous, making histopathology's use in medical practice problematic.

Consequently, Sarah was adamant that metabonomics technologies could provide a more "precise" means of tissue classification. Metabonomics technologies like MAS-NMR could detect and measure aberrant biochemical and physiological processes within cells, allowing researchers to compare the metabolites in healthy versus cancerous tissue. As Sarah said: "So the idea of the project is [to use] NMR to identify…the profile of healthy tissue. Because one thing is what you see, and another thing is…the actual biology of the tissue that you are going to be cutting." According to Sarah, the best way to detect cancer was to establish a *quantitative difference* and a *measurable continuum* between normal and abnormal tissue, rather than relying on the trained judgment of histopathologists.

In some ways, Sarah was cognizant of the difficulties that a metabonomics-based classification of healthy versus cancer tissue might pose. She acknowledged that metabonomics—like histopathology—was likely to encounter borderline cases in which the classification of normal or abnormal proved difficult. However, Sarah suggested that such borderline cases could be resolved by more technologically advanced NMR-machines and statistical models, which would collect and analyze large volumes of data. Painting a technological picture of medicine, Sarah described how high-resolution NMR spectrometers

would have the ability to show a huge range of biochemicals, while enormous sample sizes would enhance the statistical power of data models. Together, these technological advances would facilitate a clearer "separation"—in other words, a clear delineation of boundaries—between healthy and cancerous tissue. As Sarah said: "the more data you have, the more accurate you can be with analysing the problem." In doing so, she placed value on the collection and analysis of large volumes of data as a means to the diagnosis and treatment of cancer (see Chapter 5).

To support this claim, Sarah referred to a well-known experiment conducted by the NMR manufacturer Bruker Biospin on the classification of fruit juices (Bruker Biospin 2013). Bruker had collected NMR spectra for thousands of fruit juice samples, such that the company could use statistical models to determine the fruit type, variety, and geographic origin of the juice. The experiment was widely regarded throughout the BMM as a successful implementation of large-scale, robust NMR research. Extending this logic, Sarah asserted that with enough samples and quantitative data, her research would one day be able to provide a more precise classification of healthy and cancerous tissues. To this end, another researcher involved with the project commented:

> We probably just know the normal ranges of physiological measurements much better after we're doing them for hundreds of thousands of millions of times. So we're probably just at that stage where we don't know what's normal or what's not normal yet. I think we've got an idea, but we've not done thousands of samples where we can just draw the curves and go, that's the cut-off and that means cancer.

As Sarah explored the boundaries between the normal and the abnormal—physically at the margins of tumors, and conceptually within the biochemical composition of tissues—she emphasized the objectivity of her research. She said: "The reason why [MAS-NMR] is useful is because it's an objective answer you are going to obtain, because you know the biochemistry of the tissue. While now, it's a little bit biased, because it depends on the [clinician's] experience." But by asserting that "objective" metabonomics evidence was superior to "subjective" histopathological judgements, Sarah obscured the contingent nature of metabonomics practices and data. The statistical and biochemical basis of metabonomics research did not necessarily imply objectivity, and to assert so was to obscure the fact that metabonomics practices themselves involve the interpretative use of technologies[56].

[56] Another researcher commented to me during an interview that the main challenge with metabonomics practices is the ways that "subjectivity creeps into experimental design." By claiming that subjectivity was a problem, he stated the implicit value of objective knowledge gained through metabonomics practices, and yet

In asserting the objectivity of metabonomics experiments, Sarah eschewed particular "epistemic virtues" (Daston et al. 2007:40), norms and values about how knowledge should be produced. She placed value on biochemical and statistical measurements of normal and abnormal, rather than on physiological and visual assessments of tissue. Thus, as a typical example of metabonomics research within the BMM, Sarah's work was embedded within a particular notion of "objectivity," and within a particular configuration of biochemical and statistical technologies, practices, and values. This enabled a quantitative and measurable normal to emerge as the "correct" means for understanding and studying health and disease.

Shaping the Boundaries of Normal

Several weeks after Sarah showed me the MAS-NMR facility, she presented the results of her experiment in a small seminar to other researchers in the BMM. Sitting around a table in a small room adjacent to the sixth floor laboratory, she showed a Powerpoint presentation of the results of her NMR-based classification of healthy and cancerous tissue. Displaying her results as a two-dimensional graph of the supervised statistical technique partial least squares discriminant analysis (PLS-DA), Sarah showed a classification of tissues in which normal samples appeared as black dots and the tumor samples appeared as red dots (see Figure 25). In graphing the results of her MAS-NMR experiments, Sarah had assumed that there would be a clear difference between the normal and abnormal. Instead, her analysis showed that some of the normal tissue samples were grouped with, and therefore biochemically similar to, abnormal tissue samples. This presented an unclear boundary, and inherent ambiguity, between normal and abnormal tissue.

---

also acknowledged the ways in which metabonomics knowledge of health and disease is tied to practices and is therefore fundamentally contingent.

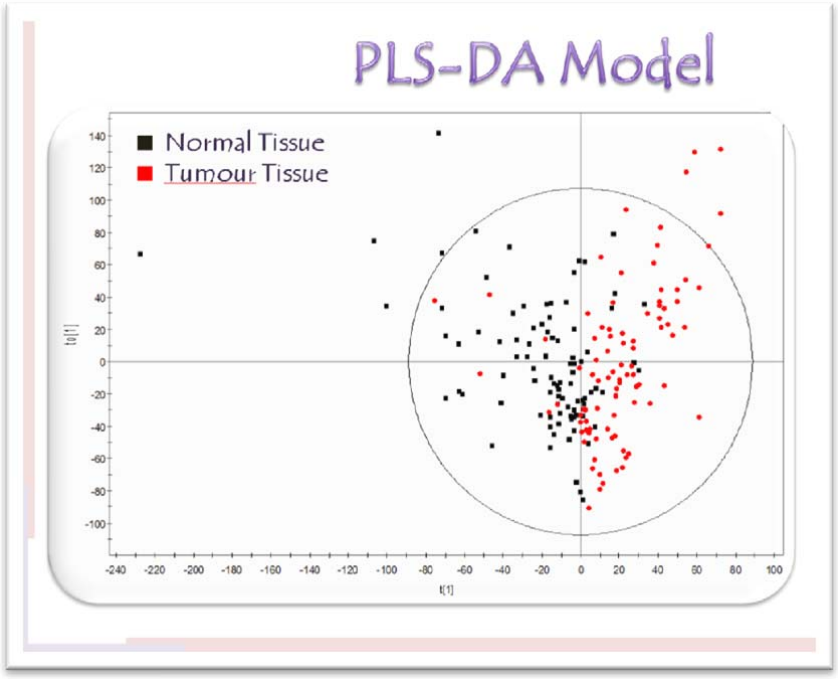


Figure 25: Statistical analysis of normal and abnormal tissue.
Powerpoint slide showing Sarah's multivariate statistical analysis of normal and tumor tissue, in which some normal tissue (black) clusters with tumor tissue (red).

That it is difficult to draw a clear-cut line between two seemingly very different types of samples is a common feature of metabonomics research. Researchers grapple with ambiguity in their daily practices as they work to find, or create, nuanced solutions to biological problems. In this sense, metabonomics research is an iterative and active process. It involves "playing" with data—looking for outliers, looking for interesting biochemical features—and also engaging with a variety of statistical tactics—generating models using new parameters and different types of algorithms. Researchers' first attempts to understand differences in data typically involve a confusing muddle of samples, and ambiguous boundaries between experimental groups. Through an on-going process of refining data and statistical analyses, metabonomics researchers work to resolve this ambiguity, and to produce stable biological concepts. Dealing with ambiguity and uncertainty, therefore, is an inherent feature of metabonomics research, and involves the active negotiation of research practices.

As Sarah discussed the ambiguous classification of healthy and cancerous tissue samples, several other researchers asked Sarah about the possible causes of ambiguity in her data. In doing so, they focused on experimental factors—those laboratory elements that were

conceptualized as "interfering" with the meaning of the data—that could have impacted the quality of the results. Firstly, the researchers asked Sarah about the technologies that she had used to conduct her experiments. She explained that the MAS-NMR spectrometer in the St. Mary's Hospital laboratory was relatively low resolution, and involved automated "shimming," a process in which the parameters of the NMR spectrometer are standardized to ensure that the resulting data is comparable. These aspects of the machine had been implemented to facilitate the effectiveness of MAS-NMR experiments in clinical settings, as lower resolution machines with automated and standardized features would ensure that experiments could be run quickly and reliably.

However, Sarah claimed that these "streamlined" aspects of the MAS-NMR, though they were valuable for clinical experiments, had negatively impacted the quality and clarity of her data. The low resolution machine had identified a small number of biochemical signals for the tissue classification, while the automatic shimming had provided data that was less "clean"—that had less clearly identifiable, sharp NMR peaks—than she would have expected from manual shimming. Sarah likened the machine's capabilities to a person's ability to identify human beings within a room: a person can be more accurately identified based on a combination of information about their height, eye color, hair color, and ethnicity, as opposed to an identification made on the basis of gender alone. In the same way, Sarah explained, the identification of the tissue samples as normal or abnormal would have been more accurate had there been more information—more biochemical signals, and of a higher clarity—within the MAS-NMR data.

Secondly, the researchers asked Sarah about the kinds of biological material she had used for her experiments. She explained that the MAS-NMR spectrometer at St. Mary's Hospital had required her to analyse solid tissue samples, rather than fluids like urine or serum. This presented a limited set of analytical capabilities: an analysis of tissue provided only a "snapshot" of the local bodily environment, and could not comment more holistically, as could urine and serum, on the state of the patient from which the tissue had come. But, as Sarah explained, this lack of information was also complicated by the fact that the experiment was not accompanied by clinical data. Because her experiments were from a pilot study, Sarah did not know the diagnosis, prognosis, and outcome of patients, and therefore had no information about the severity or outcome of their cancer. In this way, the lack of additional biological and clinical data made it difficult for Sarah to explain potential outliers—for

example those normal samples that clustered with tumor samples—or sources of ambiguity in her data.

Working with Sarah, I had the overwhelming sense that she believed the meaning and classification of normal and abnormal tissue was a pre-existing entity lying within her samples, waiting to be discovered. But as I have argued, Sarah's notions of normal and abnormal tissue was not discovered, but rather was *enacted through experimental practices*. Thus, the practices with which Sarah examined biological materials performed a particular vision of the normal and abnormal as metabolic and statistical entities, giving them meaning within the broader content of metabonomics experiments in the BMM. This is not to say that normal and abnormal were socially constructed and did not exist as things in the world, but rather to say that they were described and investigated by metabonomics researchers—perhaps at the expense of investigating other biomedical or social objects—in particular biochemical and statistical ways.

Ultimately, this section treats Sarah's moment of experimental ambiguity as a critical insight into the notions of objectivity and experimental values that pervade metabonomics research. It shows how the normal and abnormal do not simply exist as *a priori* scientific objects, but rather are enacted in situated ways. In the BMM, biochemical and statistical practices constitute the normal and abnormal as "good" and "more precise" forms of knowledge than histopathological analyses of tissue. Metabonomics researchers value the ability to sense health and disease not as microscopic images of tissue, but rather as biochemical signals and statistical patterns (see Chapter 5). As such, the statistical and biochemical notions of normal and abnormal that are enacted through metabonomics research represent one approach among many to the enactment of health and disease. Researchers actively work to stabilize and naturalize the normal and abnormal as the "correct" ways of engaging with health and disease. However, this obscures the fact that all forms of biomedical knowledge, regardless of how technologically advanced they may be, are fundamentally contingent and embedded within configurations of people, ideas, and technologies.

Grappling with the Contingency of Health

Several days later, I asked Sarah about her experience with the laboratory seminar. She expressed her frustration that the experiments had not turned out as she had hoped, and described how she had used statistical techniques to further investigate the complexities of

classifying tissue. This involved, on the one hand examining sample variability—how closely the samples in each group resembled each other—and on the other hand examining what might be causing certain samples to appear as “outliers”—as significantly dissimilar to other samples.

Sarah explained that her use of statistical models was an attempt to “get the most out of [her] data,” in order to remove or correct the variability caused by “external” factors in the experiment. These external factors included the ways that the samples were handled by surgeons, stored by technicians in freezers, or dealt with by metabonomics researchers during NMR experiments. Conceptualizing the tissue samples as separate from the experimental setup, Sarah hoped that she could use statistics to correct for the influence of external factors. This would allow her to find out more about the characteristics or essence of normal or abnormal tissue, rather than the intermingled effects of the tissue and the experimental practices that were used to examine it.

In the previous section, I discussed how metabonomics researchers enact particular versions of the normal and abnormal through biochemical and statistical practices, stabilizing the normal and abnormal as the primary way for engaging with health and disease. In this section, I discuss how the conflation between concepts of “normal” and “healthy,” or “abnormal” and “disease,” is not always straightforward, but rather is actively achieved. I draw attention to moments when researchers are faced with multiple versions of the normal and abnormal, and must therefore confront the instability of disease objects. As such, this section focuses on two cases—research on community acquired pneumonia (CAP), and research on pancreatitis—to examine how researchers grapple with the multiplicity and contingency of disease objects.

A Multiplicity of “Normal”

Several months into my fieldwork, I sat in on an informal research seminar given by Ian Wilson, a Senior Principle Scientist at the pharmaceutical company AstraZeneca. Ian Wilson had known Jeremy Nicholson—and has consequently collaborated with the laboratory—for nearly thirty years, completing a series of research projects on drug toxicity and metabolism. Given his background in pharmaceutical work and his long-standing collaboration with the laboratory, Ian Wilson had garnered years of knowledge of and experience with a wide range of metabonomics experiments, as well as the challenges associated with them.

Sitting around a table filled with eagerly listening PhD students, Ian Wilson spoke about several projects that AstraZeneca was involved with, some aimed at diagnostic or drug development, and others involving collaborative work with a metabolomics research group in Manchester. On a broad level, he highlighted the importance of selecting the appropriate "controls"—in other words, normal samples that represented the absence of disease—for metabonomics experiments. The selection of controls in experiments with living, breathing humans, Ian Wilson emphasized, was a tricky thing. It involved studying a disease in an environment replete with confounding factors that could influence the outcome of experiments. He remarked: "One of the other things you have to worry about is control samples that you've got…versus real samples that you've got from the clinic or from some other lab. Because there's all sorts of things that can go wrong."

To illustrate his point, Ian Wilson discussed a project he had been involved in on Community Acquired Pneumonia (CAP), a respiratory disease that is caused by a number of pathogens such as bacteria and viruses. To investigate those aspects of CAP that were causing patients to become unwell, researchers had looked at breath samples from diseased CAP patients and healthy controls—abnormal and normal samples, respectively—in hospital settings. The importance of this work, according to Ian Wilson, was that it would clarify whether CAP was bacterial or viral in origin, and would therefore aid clinicians in selecting an appropriate course of treatment. He said: "Most doctors can't tell the difference [between types of CAP], so they give you antibiotics anyway. And by the time they get the results back from the lab, you're either cured or you're dead."

When the researchers had compared the diseased and normal samples, explained Ian Wilson, the resulting data had shown an extraordinary difference between the two experimental groups. He said: "We got community acquired pneumonia people, and we got controls. And when you did the [statistics]…you could drive a tank through the middle. They were really well separated! …I tell you it look[ed] beautiful. And the student that got this thought she had died and gone to heaven." A statistical analysis of the CAP data, in other words, had showed a significant difference between the biochemical composition of samples from CAP patients and normal controls.

Upon closer inspection, however, the researchers had realized that the difference between the diseased and normal samples was problematic. Ian Wilson's dialogue with the PhD students at the seminar went as follows:

Ian Wilson:	Sorry, what's the difference between healthy people and sick people in hospitals?
Student 1:	Drugs?
Ian Wilson:	Yes, exactly! They're all on drugs, amoxicillin mostly, but paracetemol as well. So this group you could separate basically [because of the] drugs. Now when you took the drugs away, they still separated, which is good. But what sort of person gets community acquired pneumonia?
Student 2:	An old person?
Ian Wilson:	And what sort of person do you usually grab as a control in a hospital?
Student 1:	Students…not old people?
Ian Wilson:	Well not just students. Worse than students, doctors and nurses, who tend not to be ill… So immediately you look at these data and yes they look fine. But it's tricky…if you got your control samples from healthy animals or from your colleagues, and you're getting clinical samples from somewhere else, there's a good chance the matrix will be difference.

Ian Wilson explained how the researchers, as they investigated their data further, had attempted to reason through their selection of diseased versus normal samples. They had realized that the difference between experimental groups was not caused by disease processes, but rather by the medications that diseased patients took as part of their hospital treatment. As Ian Wilson explained, the healthy controls had been problematic. Healthy doctors and students had not been a good choice for controls: they were normal in the sense that they did not suffer from disease, but they were not normal relative to the diseased CAP patients. For this particular experiment, a better "normal" would have been patients who had taken medications but not suffered from CAP.

Continuing with his discussion of the CAP experiment, Ian Wilson commented:

> So I mean, somebody comes in with CAP, you can't say, can we rewind your body so we can get the urine before you get sick? Now you could say, would you mind coming back in three weeks' time when you're feeling better? And the response rate [would be] zero, because people only tend to go the hospital when they are sick. So in fact the real experiment we should have done there, and we know this now…we should have got another bunch of sick people as the controls, preferably on the same medication. …We needed a bunch of people with infectious disease, or controls of amoxicillin…I mean it really is immensely complicated.

Thinking about the experimental design, the researchers had realized that the difference between experimental groups was caused not only by medications, but also by the age difference between the CAP patients and their younger control counterparts. Examining the cohort of patients in the study, the researchers had determined that those with CAP were significantly older, as CAP disproportionately affected older populations whose immune systems were compromised. In this instance, an even better definition of normal would have been older patients who had taken medications but not suffered from CAP.

Overall, Ian Wilson's discussion had established multiple notions of the normal in relation to CAP. The first was represented by students and doctors, the second by patients who were receiving medications, and the third by patients who were elderly and also receiving medications. Listening to Ian Wilson, however, I was struck by the fact that the notions of "normal" embodied in the metabonomics researchers' selection of controls did not align perfectly with biomedical definitions of "health." Ian Wilson had, for example, established that older patients taking medications were more "normal" than younger students and doctors[57]. In doing so, Ian Wilson had provided a *statistical and biochemical definition of normal*, which could be measured in the controlled environment of the laboratory, but which—as I discuss later in this chapter—did not necessarily correspond to definitions of health as the absence of disease (Canguilhem 1989). Such reasoning through the "appropriate" selection of control and disease samples, therefore, is an inherent feature of much laboratory experimentation, but also highlights the instability of the relationship between metabonomics' statistical and biochemical definition of normal and a broader understanding of health.

Ultimately, this sub-section demonstrates the particular and strategic ways that metabonomics researchers engage with contingency of the normal and abnormal. When the goal of experiments is to shed light on the biochemical processes causing disease, metabonomics researchers envision the multiplicity of normal and abnormal as a problem to be solved. Consequently, metabonomics researchers attempt to control the contingency of the normal and abnormal by bracketing "experimental factors" from the "essence" of disease (Will 2007; Olson 2010). This involves attention to the role that medications and age play in determining the biological processes of disease, and attention more broadly to the fact that sampling practices can interfere with experimental results. By bracketing diseases from contexts in which they are enacted, researchers attempt to transform normal from a multiple and contingent object into a standardized biochemical and statistical entity. While this makes

---

[57] In a similar way, Jeremy Nicholson commented during an interview:

> Healthy is…[*laughs*]…healthy is actually making a statement about their biology. That they are… in an optimal or close to optimal physiological state. You could have a lot of people that are normal that are not very healthy, and if you take the American population, there's a very good example. Where, you know, maybe 40% of the people are clinically obese. [Maybe] it will be 50% of the population, and you could argue that with 50% of the population that becomes normal. But they're bloody unhealthy. Right. So…it can be normal to be unhealthy, in a particular population.

sense within the controlled environment of the laboratory, it supposes that experimental factors are independent from the disease objects to which they give meaning. Thus, in carefully selecting control samples as the “normal” and unhealthy samples as “disease,” metabonomics researchers conceptualize normal and disease as pre-existing entities lying dormant within samples, waiting to be discovered.

Thus, as researchers engage with the contingency of normal and abnormal, they do so in fundamentally limited ways. Attention to experimental practices is restricted to the technological aspects of sampling, and does not encompass the broader values and judgements involved in metabonomics research. While researchers enact normal in relation to internal metabolic signatures and statistical variation, they do not consider the external environmental context in which disease occurs. Similarly to the multivariate statistical practices I have discussed in Chapter 3, considerations of health and disease are limited to those factors that can be measured and quantified. In this way, researchers avoid confronting the challenges that arise from equating a statistical and biochemical normal with a more holistic notion of health.

A Multiplicity of “Abnormal”

I encountered a similar problematization of multiple versions of the “abnormal”—rather than the “normal”—during my interactions with Alma Villaseñor, a visiting PhD student from Spain. Alma was working on a six month clinical research project to diagnose acute pancreatitis, an inflammation of the pancreas that causes symptoms of abdominal pain, nausea, and vomiting. Though pancreatitis often resolves without treatment, severe forms of the condition must be treated with hospitalization in the intensive care unit (ITU) due to the potential of multiple organ failure. This is problematic, however, because pancreatitis is difficult to diagnose correctly, as it is characterized by a rapid onset and a diffuse set of symptoms that mimic other causes of abdominal pain. Alma’s research, therefore, was an attempt to discover several “biomarkers,” biological or metabolic indicators of disease (see Chapter 3), that might distinguish acute pancreatitis from conditions with similar symptoms. Ultimately, her research sought to uncover the biochemical characteristics that defined pancreatitis as an abnormal biological state.

To find biomarkers of pancreatitis, Alma was comparing clinical samples from people hospitalized with pancreatitis, with clinical samples from people diagnosed for a range of other conditions that cause abdominal pain, such as appendicitis, gallbladder stones, and

diverticulitis. To do so, Alma was working with a surgeon named William, who had collected the clinical samples several years prior and who now supervised her project. Over the course of several months, I watched Alma conduct—in motions similar to Ryan's and Sarah's—metabonomics experiments on urine and serum samples from a cohort of more than 150 patients. This involved preparing samples into dilute concentrations in test tubes in the "prep lab," as well as spending long hours and weekends in the sixth floor laboratory running samples in NMR and MS machines.

The majority of Alma's efforts, however, had gone into the process of analysing her data with multivariate statistical models. Every day that I came into the laboratory, I saw Alma sitting at her desk, clicking through the data analysis program SIMCA-P+, trying to find meaningful patterns in her data. Alma repeatedly told me that the challenge of metabonomics research was not in generating data with NMR and MS machines, but rather in making biological sense of the biochemical and statistical data (see Chapter 5).

As Alma attempted to analyze her data, she lamented that she was having a difficult time. Echoing Ian Wilson's discussion of the CAP experiment, the first time Alma had carried out statistical analysis on the NMR data, she had seen a significant difference between samples from control patients and pancreatitis patients. But when she investigated further, she had realized that these differences were due to the drug paracetemol, rather than to the disease mechanisms of pancreatitis. She said:

> Pancreatitis [patients] had taken much more paracetemol than the others. So that doesn't help me for anything, you know?. . . I don't have the doses they have taken. I mean, I suppose they took more paracetemol, because I saw that they have more paracetemol signals. Which is because…for example, if you hurt, okay, you ask for paracetemol. But if you still ha[ve] that pain, maybe they give you more…I don't know the doses.

After removing those biochemical signals that were related to paracetemol metabolism, in a process that involved the painstaking manual review of the data from each of the 150 patients, Alma began to take another look at her data. What she saw, however, was an even more confusing result. Her data was not split into two groups, one with pancreatitis patients and the other with controls. Instead, it was split into three groups, with the pancreatitis and control samples mixed within each group (see Figure 26). Alma's data, therefore, appeared to be inconclusive: it showed no patterns or differences between patients with pancreatitis and patients without the disease.

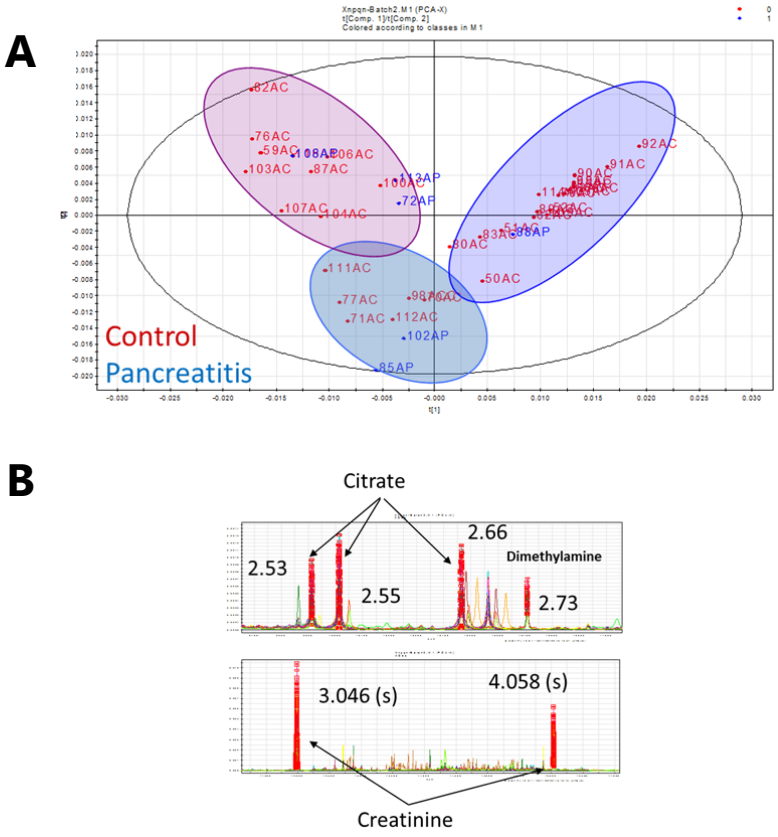


Figure 26: Statistical analysis of pancreatitis data.
Alma’s pancreatitis data, in which (A) is a PCA plot of NMR data, showing how the samples had separated into three distinct groups. (B) is an analysis of the chemical shifts of various biochemical compounds, showing the differences among the groups were due to several biochemical compounds, whose chemical signals are indicated by arrows. The images are the results of Alma’s data analysis of her pancreatitis data, ad were taken from an informal presentation she gave to her supervisor.

Several weeks later, Alma met with William to discuss her problematic results. She showed William the results of her analysis, which she had tabulated into a series of two-dimensional graphs and charts within a Powerpoint file. While some graphical representations of the data showed that the samples had clustered into three groups, made up of numbered samples in red (disease) and blue (control) writing, others showed those metabolites that were responsible for making the three groups of samples cluster together (see Figure 26).

Discussing the results with Alma, William helped her to reason through explanations for why the data had separated into three—rather than two—different groups. Alma had double-checked that her experimental methods were sound, so she had ruled out the influence

of technologies, such as NMR and MS machines or statistical algorithms, on the quality of her data. William, as a result, had begun to reflect back on the potential issues he had encountered with his collection of the pancreatitis samples from hospital patients. As he flicked through Alma's graphs, he noted that all of the samples in one group had high levels of creatinine, a metabolite which is the by-product of muscle breakdown, and which in clinical practice indicates that patients are severely ill. He said: "Isn't that absolutely extraordinary how different that is…See [Alma], you should be encouraged…Because that's not an [NMR or MS] effect. That's just clearly completely abnormal [physiology]." For William, the presence of creatinine indicated that Alma's experimental methods and data were sound, and that her confusing results had a "hidden" biological explanation.

After discussing the results with Alma for some time, William concluded that the odd grouping of the data reflected a problem with the sampling methods, rather than a biological effect. The clustering of the data into three groups, said William, corresponded to the fact that the samples had been collected by different clinicians and from different groups of patients. One clinician had focused on patients in the intensive care unit (ITU), another clinician had focused on patients in the emergency room (A&E), while William had gathered samples from patients in a variety of locations. This was a strategy, William explained, that was intended to maximize the number of patients who could be incorporated into the study. But as a side effect of this strategy, each group of patients had suffered from different severities of illness, and had consequently been characterized by very different physiologies.

Upon reflection William realized that this sample collection strategy had had unintended consequences on the quality and content of the pancreatitis data. Each group of pancreatitis patients was biologically and metabolically different. Pancreatitis patients in the ITU, for example, had been placed there because they were critically ill, and during their stay in the ITU had received a concoction of medications. Looking back on the clinical data that he had collected along with biological samples, William saw that patients from the ITU had been "really sick," while the patients from the A&E had exhibited "abnormal physiology," and the patients from William's sample group had been "not terribly unwell." Thus, Alma's data showed three different version of pancreatitis, and ultimately three version of the "abnormal."

I have discussed Alma and William's research on pancreatitis because—similar to Ian Wilson's research on CAP—it highlights how metabonomics researchers confront and

engage with the contingency of disease objects. To this end, Alma and William not only bracketed contingency, but also mobilized it as something productive. They used their uncertainty to ask important questions about the meaning and form of pancreatitis as a disease. According to William, Alma's data provided insight into the fact that pancreatitis—while clinically a single disease—was made up of multiple biological entities that arose from combined genetic and environmental factors. Accordingly, *pancreatitis emerged as something that was inherently multiple*: it was associated with a diverse set of biological and metabolic pathways, as well as a diverse range of symptoms and treatments. As William said: "[The data] reinforces the fact that…diseases are not homogenous or easy to separate…And it shows that people behave differently when they get really sick. And actually, we can see that. We can measure it." William acknowledged, based on his clinical experience, that individual patients had unique experiences and responses to disease.

To this end, Alma's work provided the researchers with a chance to revisit the symptom-based clinical classification of pancreatitis, and to "draw distinctions and see underlying unities that would not have been possible on a clinical basis" (Navon 2011:207). Overall, William and Alma's case highlights how metabonomics researchers do not always view multiple disease objects as problems to be solved, but instead draw upon contingency to ask important questions about the status quo of biological knowledge. As Julie Sommerlund (2006) has described with the classification of bacteria, multiplicity and contingency can act as productive forces, by prompting researchers to ask questions and carry out new experiments. Thus, William and Alma used multiplicity of pancreatitis to redefine disease as something molecularly diverse[58].

In grappling with pancreatitis as something molecularly diverse, however, William still described disease in fundamentally biochemical and statistical ways. In his efforts to redefine disease, he focused on the variability of biochemical signals present in his samples. He did not ask, in contrast, how the contexts in which the patients had developed disease, or the initial medical diagnosis they received, had affected their outcome. Consequently, this sub-section shows that as metabonomics researchers engage with the multiplicity of the abnormal, they do so within a limited biomedical framework. Though researchers

[58] This was largely an attempt to align the reclassification of pancreatitis with "precision medicine," the notion that physiologically similar conditions like cancer and obesity can be defined according to specific molecules and pathways, and subsequently treated with personalized approaches to medicine (see Chapter 6).

acknowledge that organisms can display a wide range of states of health and disease, they still value and naturalize the fact that disease should be studied molecularly, and with biochemical and statistical technologies.

Ultimately, the examples of Ian Wilson, and of Alma and William, shed light on the various ways that metabonomics researchers grapple with the multiplicity—and therefore the contingency—of concepts of normal and abnormal. By confronting the problematic aspects of their experiments, metabonomics researchers articulate multiple versions of the "normal" and "abnormal." These are embodied in the multiple possible controls in Ian Wilson's CAP experiment, as well in the multiple types of disease in William's and Alma's pancreatitis research. This speaks to Annemarie Mol's (2002) work on the multiplicity of biomedical objects, in which researchers engage with the "practicalities of enactments" (Mol 2002:160-61) and do active work to promote experimental stability and coherence. Thus, the multiplicity of disease objects in metabonomics experiments highlights the active work that goes into establishing the relationships between normal/abnormal and health/disease.

This section also shows, however, that as metabonomics researchers grapple with the multiplicity and contingency of biomedical objects, this process is itself not uniform. Though metabonomics researchers' engagement with contingency is fundamentally strategic, the strategies that researchers use are also varied. While in some cases contingency is seen as problematic, in other cases contingency is seen as something generative, and is used to develop new forms of biological knowledge. But in the end, both ways of dealing with contingency involve calculated—and inherently limited—frameworks for engaging with health and disease. Thus, although metabonomics researchers actively grapple with contingency and question the conflation between "normal" and "health," or "abnormal" and "disease," they do so with explicitly biomedical—and therefore inherently limited—ideas and values.

Disjunctures Between "Normal" and "Health"

Though normal and abnormal provide guiding concepts and ideologies in everyday metabonomics practice, they also present, as I have discussed in the previous section, at times fundamental challenges to researchers' abilities to comment on health and disease. By relying on biochemical and statistical practices to define and engage with biology, researchers struggle to capture the dynamic and individual nature of disease. The binary categories of normal and abnormal do not easily translate into the multiple and holistic

concepts of health and disease. For example, during the time I spent with Sarah, she openly discussed some of the challenges she had encountered in establishing the boundaries of normal in her research on cancer tissue. She had encountered a variety of issues with the experimental setup of the MAS-NMR spectrometer: the custom-made robotic autosampler had a tendency to break, and the statistical modelling within the computer program MATLAB had a tendency to crash. But she also encountered issues with her acquisition of normal tissue. Echoing Ian Wilson's discussion of the CAP experiment, she described how her normal samples had been taken from patients who were being treated for cancer, as the ethical constraints of the experimental design had precluded sampling tissue from healthy volunteers.

In discussing those ambiguous samples that were continually proving a nuisance to classify, Sarah seemed uneasy or troubled by some aspect of her work. She questioned her methods, and above all, wondered if a normal sample taken from someone with cancer could be classified as healthy. As we chatted, she paused for a moment and said: "You know, the normal is not always healthy." It was at this moment that I truly realized the centrality—and challenge—of concepts of normal and abnormal to metabonomics research. In my fieldnotes from that day, I circled Sarah's quote in red pen, and scribbled possible explanations for the seemingly strange disjuncture between the normal and the healthy.

What Sarah had articulated was not only the contingent nature of the concepts of normal and abnormal, but also their inability, in some cases, to comment on issues of health and disease. Sarah, through deeply questioning her selection of experimental samples, acknowledged that the normal that metabonomics experiments enacted did not necessarily correspond to health. Because Sarah's normal samples came from diseased patients, they could act as controls for experiments, but not as holistic representations of health. In acknowledging this, Sarah—like Ian Wilson, Alma, and, William—had engaged with the contingency and practicalities of her work, but in a way that made her fundamentally question the validity of metabonomics experiments.

For example, throughout my fieldwork metabonomics researchers frequently expressed concern that their experiments highlighted the metabolic effects of medications rather than the metabolic processes underlying health and disease. Ryan, the PhD student whose work I mention at the beginning of this chapter, explained how his efforts to identify inborn errors of metabolism had instead identified medications that were not written in

patient records. Likewise, another PhD student working on vein and artery disease had encountered problems when his work identified common hospital medications like anaesthetics and pain killers, rather than vascular physiology. These research projects had been constrained by the logistical and ethical practicalities of obtaining biological samples from human beings. Metabonomics researchers were often forced to select control and experimental samples from existing populations in hospitals or studies. This meant, more often than not, that experimental samples came from patients who were sick in multiple ways, and whose conditions were already being intervened upon at the time of the study. It also meant that control samples rarely came from healthy patients, and instead had to be obtained from sick patients who were already undergoing hospital procedures. As such, the designs for clinical metabonomics experiments were rarely ideal, in the sense that they struggled to isolate the "facts" of disease from the "facts" of medical interventions.

Thinking back to Ian Wilson's seminar, I found myself wondering further about the ability of metabonomics research to comment on health and disease. After describing the CAP study in its totality, Ian Wilson had discussed another study that had attempted to establish a normal baseline for the health of residents of a town in the northern United Kingdom. Describing the study, he said:

> Again one of the things the [laboratory] group has done is really quite ambitious, they've tried to define normal rather than define disease. They said, let's get loads of samples from normal people. Now defining normal is pretty difficult, especially if the population you've chosen is males and females from [a town in the north of the UK]…So they're normal, but they'll be [full of] hidden disease, they'll be pre-diabetic, they'll be people who have got the time to go to their doctor during the working day. . . So it's normal for a given definition of normal…And some of those normal may not be normal at all. They may be diabetic or alcohol abusers or smokers. I mean, that's another thing, is someone who smokes normal or not?

Ian Wilson, like Sarah, outwardly questioned whether the "normal"—which in this case encompassed alcohol abusers, smokers, sufferers of diabetes—was healthy. He questioned whether the normal established by metabonomics experiments corresponded to the broader health of populations, and to their ability to exist without disease or the need for medical interventions. In doing so, Ian Wilson questioned whether metabonomics practices, in trying to statistically establish what was "normal" in a heterogeneous and unhealthy population, could shed light on the broader patterns of health and disease that occurred beyond the confines of the laboratory.

Such comments speak to Georges Canguilhem's (1989) insistence that the "normal" is a decontextualized, evaluative, and frozen concept of health (see Chapter 1). For

Canguilhem, health was not normal, but rather the normative capacity to continually establish new "normals" in response to the environment. To this end, as Sarah emphasized her desire to see clinical data—which consisted of five years of radiological findings, biopsy and histopathological results, blood tests, body mass index measurements, and medication records—she acknowledged that metabonomics measurements of the normal were not sufficient for the grasping the complexities of cancer. Sarah felt that this clinical data would shed light on the long-term status of the patients, who, unlike the samples that they had provided, could change over time. As Sarah said: "With tissue, what we will be able to answer is yes, this is healthy this is not. But if you want to see the degree of change, or how the patient is going to evolve or things like this …you need loads of samples…and you need samples for a long period of time in order to see how [things] change."

Sarah's comments, therefore, articulated the contrasts between statistical notions of the normal, which reflect decontextualized snapshots of an organism, and dynamic notions of health, which encompass the way an organism changes with the environment and over time. Sarah's desire to pair her own metabonomics work with clinical data revealed her realization that the normal, while it was the operating concept of health in the laboratory, was not necessarily so everywhere. Ultimately, her comments raised fundamental questions about the ability of statistical and biochemical experiments to capture health and disease in relation to the normativity and vitality of life.

Sarah was not the only metabonomics researcher to express this concern. For example, I spoke with another researcher named Carrie about her work on the metabolic characteristics of rodent models of obesity. This involved research on Zucker rats, an inbred strain of laboratory rats with mutations in the leptin gene, which leads to impaired glucose tolerance and obesity. Though her work had established key metabolites responsible for the abnormal physiology of obesity in rats, Carrie expressed concern that this research only touched upon particular scales and aspects of obesity. It did not address issues of unhealthy food environments or choices, which Carrie acknowledged were implicated in the development of health issues pertaining to obesity. Thus, Carrie believed that her findings would elucidate the metabolic effects of leptin mutations, but was sceptical that her research would influence the treatment of obesity as a societal problem. She said to me: "The answers are ultimately that people need to eat better and exercise better."

In doing so, Carrie acknowledged—like the majority of researchers I interacted with in the BMM—that the results of her experiments were contextual and valid only within certain parameters of laboratory practice. Her comments emphasized that modern biomedical research, despite its contributions to medical knowledge and practice, cannot exist independently from the clinic or society (Mol 1998). Continuing on, Carrie emphasized that to make her work applicable to clinical issues of obesity, she would need to examine the reasons and mechanisms for variation between obese individuals. Carrie stressed that human beings were very different from Zucker rats: they did not have identical genetic backgrounds, and thus did not develop obesity in a near identical fashion. There was a great degree of variation in the ways human beings responded to diets and developed obesity. Illustrating her point, Carrie joked that she ate poorly, smoked, and did not exercise. Yet she still managed to maintain a healthy weight, which showed that some people could treat their bodies poorly and never develop obesity.

Overall, such comments expressed a concern with relating issues of individual and population variability to concepts of health and disease. As another researcher commented during an interview, metabonomics experiments rarely provided a conclusive definition for health and disease because there was such a range of variation between individuals. As the researcher said:

> What's healthy for one person isn't necessarily healthy for another person…The differences between people depend on what you're looking at…So urine will directly reflect what you take in, so a free living population is going to be massively different depending on what they take in. Japanese people are very different from American people because they have a lot of fish in their diet, so they excrete a lot more metabolites derived from fish, so when you look at the metabolic profile they're totally different. But does that mean that the Japanese are healthy and Americans aren't? Well not really.

Thus, as metabonomics researchers attempted to engage with normal and abnormal as variation along a statistical continuum, they also struggled to reconcile this variation with the binary nature of the normal and abnormal. The quantitative nature of the normal and abnormal, in other words, did not always match up with the qualitative nature of health and disease. Like the enactments of complexity I have discussed in Chapter 3, metabonomics researchers' enactments of "normal" engaged with health and disease in specific and limited ways. They did not take into account those aspects of life which influenced health and disease, but which could not be quantified through laboratory experiments.

Ultimately, this section argues that metabonomics researchers encounter fundamental challenges in relating laboratory norms to clinical ideas of health and disease, a point that I

elaborate upon in the chapter that follows. As metabonomics researchers attempt to define the boundaries of the normal and abnormal, they have an intrinsic awareness that such boundaries are inherently unclear, fuzzy, and overlapping. As one clinician proclaimed: "The thing is, you cannot draw a line in [disease], whatever it is, where there is not overlap. Even in the most discriminatory models there's always overlap." Thus, in medical practice, the statistical categories of normal and abnormal are not as binary as they seem, or as researchers want them to be. In light of the variation that exists within and between individuals, metabonomics researchers struggle to define the hard and fast point at which the normal is distinct from the abnormal. Questions of the dynamic and normative nature of health, or of the individual differences in disease, confound the conflation between normal/abnormal and health/disease.

Discussion

Within the metabonomics laboratory, the normal and abnormal are enacted as quantitative, biochemical, and statistical objects. They are naturalized as the "precise" and "objective" ways of engaging with issues of health and disease, and are positioned as superior to histopathological forms of knowledge. However, such claims to objectivity obscure the ways in which normal and abnormal—as the products of actively negotiated laboratory practices, technologies, and values—are fundamentally contingent and unstable objects. Consequently, metabonomics researchers must constantly grapple with the contingency and multiplicity of the normal and abnormal. Reasoning through their experiments, they deploy multiple strategies not only to contain contingency, but also to mobilize it as something productive for further research. But although metabonomics researchers ask important and often self-reflective questions about experimental design and the status quo of biological knowledge, their statistical and biochemical practices provide an inherently limited framework for engaging with health and disease. Ultimately, the binary concepts of normal and abnormal struggle to engage with the dynamic and individual nature of health and disease.

In its totality, metabonomics research reveals the challenges faced by those experimental sciences—metabonomics, genomics, and other post-genomic sciences—that are attempting to extrapolate decontextualized measurements of normal and abnormal to derive holistic and dynamic understandings of health and disease. With the fast-paced knowledge production of the post-genomic sciences, the practices of biomedicine around which Georges

Canguilhem and Michel Foucault based their observations about the normal and abnormal have undoubtedly changed. Sociologist Nicholas Rose (2009) has argued, for example, that large-scale genomics projects have done away with the binary of normal/abnormal to establish that "variation" and "risk" are the modern normal and pathological, respectively. Drawing upon these questions, I have argued that metabonomics, instead of articulating new concepts of normal and abnormal, enacts concepts of normal and abnormal that are not always compatible with the individual normativity and vitality of life.

This speaks in many ways to the work of doctor and biologist Ludwik Fleck (1986), who argued in the 1920s that the realm of human experience and perception is not always congruent with the exactness—and the strict boundaries and logics—of modern science. According to Fleck, medical and scientific reasoning were fundamentally different, as scientists looked for the typical or normal, and clinicians looked for the atypical or abnormal. Through the practice of medicine, claimed Fleck, it became apparent that there was no fixed boundary between health and disease, and that such boundaries were different in individual patients. Though Fleck was writing about early 20th century biology and medicine, his observations still apply to contemporary times. They speak to the struggles of metabonomics to "find a law for irregular phenomenon" like disease (Fleck, as quoted in Seising 2008:1236), and to superimpose strong logical reasoning onto realms of practice which—as I discuss further in the chapter that follows—require human intuition and judgment.

What is perhaps ironic about the challenges facing metabonomics, is that the field, as I have discussed in Chapter 3, is engaged in a holistic rhetoric of metabolism that proclaims to capture the dynamic character of health and disease. Metabolism, a term which encompasses the sum total of the processes that maintain the life of an organism, is an inherently dynamic process. Through metabolism, the body interacts with and adjusts to its environment, consuming, assimilating, and reproducing the materials that it needs to live, and also challenging many bodily boundaries such as self versus other, internal versus external, and body versus environment. Metabonomics is therefore characterized by a rhetoric of normativity, as researchers continually emphasize that their field, in contrast to genomics, can engage with biological complexity and gene-environment interactions.

Metabonomics practice, however, speaks to a different reality. As I have argued throughout this chapter, by transforming dynamic metabolic processes into mathematical, statistical patterns, metabonomics research fixes and freezes the processes of metabolism,

preventing health and disease from being expressed in dynamic terms. In some ways, metabonomics researchers are aware of the potential limitations of their own work. As they struggle to make sense of their data, they emphasize the importance of collecting large volumes of biochemical information, and of carrying out complex multivariate statistical analyses. In this effort to capture more information, there is the sense that the limits and problems of metabonomics research can be transcended with the capture and analysis of large volumes of data, a point to which I return in the chapter that follows. Ultimately, this begs the question of whether the challenges inherent in metabonomics’ attempts to define health and disease can be eclipsed by the aggregation and interpretation of more information (Pollack 1 December 2011), or whether there are fundamental barriers in the need for human interpretation and judgment.

Chapter 5: Interpreting and Valuing Data in Translational Research

Introduction

On a sunny summer morning, I am standing in a laboratory on the tenth floor of St. Mary's Hospital watching Sarah—the post-doctoral researcher whose work I discuss in Chapter 4—interact with a surgeon-in-training named Joseph. Wearing pristine white laboratory coats that reach down to their knees, Sarah and Joseph are working on metabonomics experiments on colon cancer tissue in the newly-installed magic angle spinning nuclear magnetic resonance (MAS-NMR) facility. The facility consists of a large MAS-NMR spectrometer, several computers to analyse the data generated from experiments, a large freezer in which samples are stored, and a separate room in which samples are prepared.

Joseph is one of several clinical practitioners carrying out research in the BMM as part of his training to become a surgeon. After completing the clinical component of his training, he elected to carry out a PhD in the BMM on the metabolic properties of colorectal cancer. He has planned to draw on his training in colon cancer surgery to collect samples, and to gain insight into the integration of laboratory and clinical approaches to disease diagnosis and treatment. Despite being through more than a decade of surgical training, however, Joseph's laboratory experience is minimal. Joseph has spent little time in the metabonomics laboratory, as the first few months of his PhD have been spent collecting colorectal cancer samples from the surgical operating theatre. As a result, Joseph is a self-proclaimed "complete beginner in the laboratory." He is adept with surgical tools, but he has none of the skills required to carry out NMR experiments or analyse metabonomics data.

Throughout the morning, Sarah has been reprimanding Joseph for his improper handling of tissues and samples, and her frustration is obvious. Joseph has just exited and re-entered the laboratory while still wearing used gloves, and Sarah is upset that this has potentially contaminated the laboratory environment. Joseph, Sarah exclaims, has spread bits of tissue across the computer, freezer, door handle, and anything else he has touched while wearing gloves. Sarah feels that this is a reflection of Joseph's lack of care and concern for the rigour of metabonomics experiments. Joseph's response is to try to defend himself—by explaining that in his clinical work he is not used to changing gloves—but Sarah is too flustered to listen to his reasoning.

Such challenges speak to the fundamental differences between laboratory and clinical professions, which scholars have referred to as "styles of reasoning" (Hacking 1999) or the "professionalization" of biomedical groups (Wilson-Kovacs et al. 2011). Clinical researchers, though they constitute a diverse range of specialities and skills, are indoctrinated into certain ideas and practices by biomedical institutions (Löwy 1996). Consequently, clinical researchers learn to "see" and engage with bodies in particular ways: their everyday contact with patients impacts how they think about and approach biomedicine, and dictates that their research is focused on the diagnosis and treatment of human health and disease. Therefore, it could be argued that clinical researchers and scientists represent different "epistemic cultures" (Knorr-Cetina 1999), and consequently carry out "boundary work" to create professional forms of expertise and knowledge (Gieryn 1983; Burri 2008). But the differences between clinical researchers and scientists are not only limited to different cultures and perceptions. Their differences are also ontological, in that they entail fundamentally different practices for thinking about and interacting with disease.

Several days later, when I speak to Sarah about her interactions with Joseph, she explains her frustration with Joseph's lack of interest and effort in experimental laboratory work. Because Joseph is so busy doing surgical training and collecting clinical samples, says Sarah, he is not able to fulfil his duties as a doctoral student-in-training. Problematically, he prioritizes his patients over his experiments, and does not spend enough time learning how to do experiments from Sarah. As a result, Joseph has made critical mistakes handling the tissue samples and machines in the tenth floor laboratory. "He doesn't even know how to pipette," Sarah says angrily, "And he doesn't actually know what research *is*."

Sarah's comments speak to the fact that clinical researchers and scientists are different not only in their cultures of professional training, but also in their very notions of what constitutes biomedical practice and its objects of investigation. Sarah and Joseph's conflicts over sample handling embodied what other researchers described as a "gulf of understanding" between clinical researchers and scientists[59]. Though Joseph struggled with the practicalities of laboratory research—of learning how to use NMR spectrometers, of attempting to balance the time demands of clinical work and laboratory research—he struggled on a more fundamental level to understand the ideologies and values of molecular, post-genomic research. As Joseph's supervisor—a surgeon named William, whose work I discuss in Chapter 4—commented, Joseph was "not versed in the language of basic science, much less biochemistry and multivariate statistics." He was not expected to have the same skills and knowledge about metabonomics experiments as Sarah, because his everyday clinical work entailed thinking in terms of patient needs and disease treatment. As William said:

> A clinician will never look at a [an NMR spectrum] and go 'Oh yeah, that's a 1.18 doublet, yeah, alcohol.' Forget it…and the clinician doesn't have time to look at a PLS-DA plot, and go 'Yeah, that's quite pretty, but what's the Q2 and what's the R2?'[60]

In contrast to Joseph's struggles with metabonomics experiments, Sarah found it challenging to understand the ideologies and values of addressing the treatment of living and breathing patients. William described, for example, how Sarah had struggled to relate MAS-NMR data to data for clinical treatment and outcomes. He said:

---

[59] Another researcher described his struggles to interact with clinicians who were providing samples for his metabonomics experiments. Because he was examining the metabolic properties of calcified arteries, he needed to conduct experiments on biological tissue that was kept on ice, which prevented degradation and extra metabolic activity. However, the vascular surgeon with whom he was collaborating did not understand the need to preserve tissue in this way. As such, he consistently delivered arterial samples that had not been placed on ice, which made it difficult for the researcher to explore the metabolic basis of calcification.

[60] This technical terminology refers to the supervised multivariate technique partial least squares discriminant analysis (PLS-DA). Q2 is an estimate of the predictive ability of a multivariate statistical model, and is calculated by cross-validation of data within the model. A lower value of Q2 indicates a higher predictability for the model. R2 is an estimate of the residual, or the difference between the model and the original data. It represents unexplained variation or errors. A low R2 indicates a large amount of noise or irrelevant information in the data.

> [Sarah] looks at the data as a scientist. So she's looking at the patterns, and she's trying to figure out why people are clustering. But she's not looking beyond that into the clinical data, and trying to understand what's actually happened to the patient while they were in hospital, and what is the pathology and what was their outcome.

William emphasized the ways in which clinical and metabonomics researchers had different practices and ways of thinking about biology. Issues of communication and collaboration arose, therefore, from the combined fact that Sarah did not carry out work in terms of patient bodies and treatment, and that Joseph did not carry out work in terms of the production and analysis of biochemical and statistical data. As another metabonomics researcher commented during a seminar: "Fair enough, [clinicians] have a patient to deal with…I believe that the life of the person is more important than the research project. But why do they choose to do research as well as medicine?"

Overall, issues of understanding and language highlight the challenges faced by researchers like Sarah and Joseph, as they attempt to collaboratively apply metabonomics technologies and practices to clinical problems. Taking Sarah and Joseph as a starting point, this chapter argues that the research challenges at the clinic-laboratory interface arise through different and overlapping realms of practice, and the different disease objects that they enact. Conflicts and issues of collaboration arise not only because of different ideas about how to diagnose and treat disease, but also because of *different ways and practices of doing biology*. From this perspective, clinical work and laboratory research represent different realms of material practice, and consequently produce or "perform" different disease objects and values (Moreira 2006) (see Chapter 1). They struggle, in other words, to achieve coherence in their approaches to biomedical research and practice (Mol 2002).

Translational Research as Informational Practice

This chapter begins with a vignette of Sarah and Joseph, because their interaction highlights some of the fundamental and recurring challenges faced within "translational research," the movement of knowledge from the laboratory to the clinic. Commonly referred to as "bench to bedside" research, translational research has become an increasingly important concept in the biomedical sciences over the past decade (Kohli-Laven et al. 2011), and encapsulates a wide range of research practices, locales, disciplines, funding strategies, and ideologies (O'Malley et al. 2007a; Rajan et al. forthcoming)[61]. Though definitions of

[61] The term translational research originated in the 1990s in the United States as part of the National Cancer Institute's (NCI) attempts to develop cancer therapies (Kohli-Laven et al. 2011). Soon thereafter it spread to

translational research are multiple and heterogeneous, the term broadly encapsulates attempts to mobilize the knowledge generated in the life sciences towards the advancement of human health. The movement of knowledge between the laboratory and the clinic, moreover, is bi-directional: laboratory research is not simply applied to clinical issues, but rather involves the co-production of research practices and agendas (O'Connell et al. 2006; Wainwright et al. 2006; Morgan et al. 2011).

During my fieldwork, the BMM was undertaking a series of clinically-oriented research projects funded by the Imperial College National Institute for Health Biomedical Research Council (NIHR-BRC), as discussed in more detail in Chapter 2. These involved not only the participation of clinically-trained researchers in metabonomics laboratory experiments, but also the application of metabonomics technologies and approaches to clinical issues. This was part of the increasing impetus to bring academic medicine into contact with laboratory research through the figure of the "clinician-scientist." It articulated the growing expectation within the UK that research occurs concurrently with clinical practice, and that clinical research is the "essential conduit" for the translation of laboratory research from bench to bedside (Wilson-Kovacs et al. 2011). Spurred by Jeremy Nicholson's attempts to use metabonomics research as a tool to develop translational medicine, clinical researchers and surgeons with a wide range of specialities and backgrounds entered the BMM to complete long- and short-term projects. They did so primarily as a way to advance their clinical careers, both to be at the "cutting edge" of their field, and also to promote a deeper understanding of the biological processes underpinning their clinical practice. Thus, through the increasing incorporation of clinical researchers into research projects, funding structures, and daily activities, the BMM began to establish "translational medicine" as an important research endeavour.

This chapter explores the concept of translational research as an informational practice: as an often problematic attempt to create, shape, and move data between the

---

other areas of the biomedical sciences, such that it was clearly articulated as a cornerstone of research in the National Institutes of Health (NIH) 2003 Roadmap, as well as in the 2006 Clinical and Translational Science Awards (CTSA) scheme. Concurrently in the United Kingdom, translational research has been supported through funding streams in the National Institute for Health Research (NIHR) and the Medical Research Council (MRC), for work on large scale clinical trials and early-stage discovery, respectively. Specifically, the NIHR has committed more than £800 million to the establishment of Biomedical Research Centres (BRC) and Biomedical Research Units (BRU) within the NHS and university partnerships. Likewise, the MRC has committed more than £300 million to translational research between 2012-2016 (Science and Technology Committee 2009; Homer-Vanniasinkam et al. 2012).

realms—conceptual and physical—of laboratory research and clinical practice. It focuses on the practices that render knowledge at the interface between the laboratory and the clinic, examining how disease objects are enacted and problematized by researchers in everyday practice. Rajan and Leonelli (forthcoming) have written, for example, how within translational research the integration and aggregation of different types of data is promoted as the key to "translating" biological understandings into medical practice. This eclipses, as Paul Edwards (2011) writes, issues of "data friction," in which the movement of data between people, organizations, and machines creates costs in time, energy, and human attention, and ultimately compromises the ability of disciplines to collaborate.

This informational view of translational research, however, entails *particular definitions/articulations of what constitutes data, and of what value data has in medical practice*. Acknowledging that such definitions and values are highly dependent on the context in which data is developed and used, this chapter explores how data in translational metabonomics research is something inherently statistical, molecular, moveable, and relational. Here, I follow Sabina Leonelli's definition of data[62] as "mobile pieces of information, which are collected, stored and disseminated so as to be used as evidence for claims about specific processes or entities" (Leonelli in press:3). Thus, this chapter explores how data, rather than improvements in clinical technologies or practice, is valued as vital to the advancement of human health. It asks how we have come to see data as something central to investigations of health and disease (see Chapter 4), and how this renders the information-based treatment and diagnosis of disease problematic.

In contrasting the laboratory with the clinic, my aim in this chapter is not to essentialize different realms of practice, by claiming that there are fundamental differences between clinical work and laboratory research. Nor is my aim to examine the perceptions and discourses that surrounds translational medicine, without examining them as diverse and material practices. My aim, rather, is to examine how the changing objects of biomedical research are articulated at the interface between clinical and laboratory practice. It is, in other words, to examine how the increasing contact and hybridization of the laboratory and clinical

[62] Working with this definition of data as information that is made meaningful, it should be noted that information is never "raw." Through the experimental practices that lead to its production, information always has value before it is analysed and transformed into data (Gitelman 2013). Thus, the distinction between information and data is rhetorical rather than ontological, and serves to distinguish between information-at-large and information that has been transformed into a naturalized, normative form.

sciences is resulting in changing technologies, practices, and approaches to the understanding and treatment of disease.

Ultimately, in this chapter I explore the everyday work, and the challenges that arise, as researchers engage with and make sense of data that is generated in the application of metabonomics practices to clinical issues and questions. First, I highlight the centrality of data and informational practices to translational research: I explore how multivariate statistical techniques enact biological tissues as statistical patterns and thresholds, and how this differs from the tissues enacted by histopathological practices. Secondly, I discuss how researchers, in their search for statistical patterns and thresholds, place value on collecting and analyzing data in particular statistical forms as the best way to diagnose and treat disease. Thirdly, I discuss how researchers articulate the main challenge in conducting translational research not as the generation, but the interpretation of data in light of biological pathways and systems. To conclude, I suggest that the challenges faced in translational research lie not only in different cultures and technologies, but also different values surrounding human interpretation and judgment.

Data Practices and Disease Objects

Throughout my fieldwork, I came to know Judy Fonville, a PhD student in the BMM whose work I discuss in Chapter 3, as she navigated the final stages of her doctoral research. Her work exemplified research in the field of chemometrics, the science of modelling chemical data using multivariate statistics to examine problems in biology and chemistry. Judy was one of several researchers within the BMM carrying out work with matrix-assisted laser desorption/ionization mass spectroscopy imaging (MALDI-MSI), which she used to find and make sense of the biochemical patterns contained within tissue slices. Though Judy's MALDI-MSI data came from a collaborating laboratory at the University of Birmingham, this specialized technology was also being developed in parallel in the BMM as a "clinical platform"—a translational medicine technology that would be used in clinical settings for disease diagnosis—in collaboration with the Waters Corporation, one of the premier developers and manufacturers of mass spectrometry (MS) machines.

Using Judy's work as an example of the ways in which translational research revolves around the use of complex data practices and statistics, this section explores the centrality of informational practices and statistical knowledge to translational medicine research in the BMM. Firstly, it examines how through specific practices, in particular multivariate

statistics, metabonomics researchers enact tissues as complex data and patterns of information. It argues that researchers create and learn to see the differences between tissues as statistical thresholds and boundaries through the use of multivariate statistical techniques. Secondly, it examines how contrasting realms of practice—between histopathologists and metabonomics researchers—lead to the enactment of different biological objects.

MALDI-MSI was originally developed within the field of proteomics to determine the peptides present on the surfaces of tissue sections (Moody 2004). Within the BMM, it involves the application of MS, which was developed in the early 20th century and used in fields like physics and chemistry to sense the masses and structures of molecules, to studies of imaging within tissues. MALDI-MSI enables the simultaneous analysis of hundreds of molecules within a tissue sample, and in particular allows researchers to determine their spatial location. Consequently, its development provided a molecular complement to imaging techniques like immunochemistry and fluorescence microscopy, all of which study the spatial arrangement of molecules within biological tissues (Stoeckli et al. 2001:493).

For her project, Judy used MALDI-MSI to analyse the biochemical composition of thin cross-sections of rat brain tissue. With MALDI-MSI, Judy analyzed data in the form of "molecular maps" of the tissue, which transformed each tissue slice into a grid of data points: each point or "pixel" along the grid represented a physical area of tissue, and contained information about thousands of biochemicals[63] (see Figure 27). In essence, MALDI-MSI provided a microscope-like view of the biochemical composition of tissue, allowing Judy to generate images that showed the identity and spatial location of the biochemicals within tissues.

[63] As one scientific paper describes MALDI:

> Each laser-irradiated spot (pixel) gives rise to a mass spectrum that is correlated to a discrete X,Y coordinate location on the tissue. Thus each spot or pixel contains a dataset having thousands of channels (m/z values) with each channel having its own brightness (intensity). The intensity of each m/z value can be expressed over the array of pixels as a 2D ion density map. Commercial or custom software can be used to generate images depicting the localization and relative intensities of hundreds of ions in a single acquisition from a tissue section. (Seeley et al. 2008:18126)

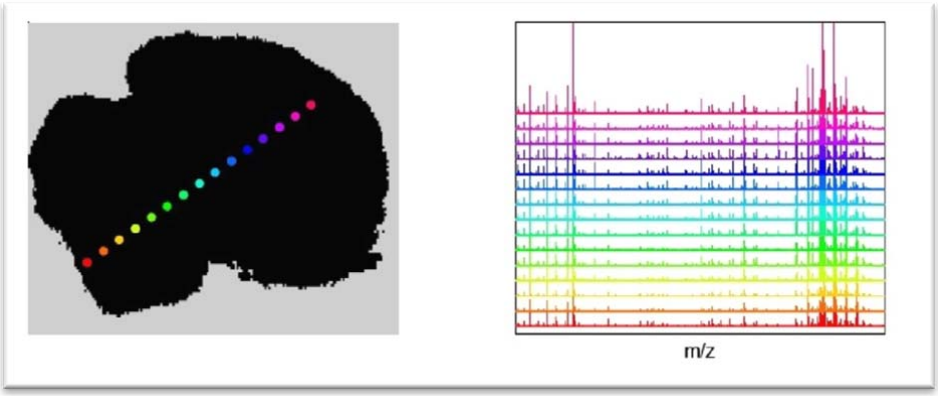


Figure 27: MALDI-MSI of rat brain tissue.
A "map" of MALDI-MSI data, taken from Judy's thesis. The left picture shows several data points (pixels) in the grid of data points along the tissue slice, while the right picture shows the biochemical spectra that are contained within each data point.

Enacting Tissues as Statistical Patterns

I heard Judy present her work on MALDI-MSI, which also constituted the final chapter of her thesis[64], at a departmental seminar given to graduate students, lecturers, and clinician-researchers. Judy described her efforts to understand the data generated by MALDI-MSI, and how she had explored the implementation and use of various data analysis techniques. With these data analysis techniques, Judy hoped to draw biological conclusions from complex biochemical information. On the one hand, she hoped to use statistical techniques to understand the structure and function of rat brain tissue. On the other hand, she hoped to determine whether MALDI-MSI data could be correlated with—or could perhaps improve upon—histopathology (see Figure 28), a clinical technique that involves the visual analysis of stained cells under a microscope.

[64] This chapter was titled "Multivariate analysis of mass spectrometry imaging data," and was later published into two articles in the journal *Analytical Chemistry* (Fonville et al. 2012a; Fonville et al. 2012b).

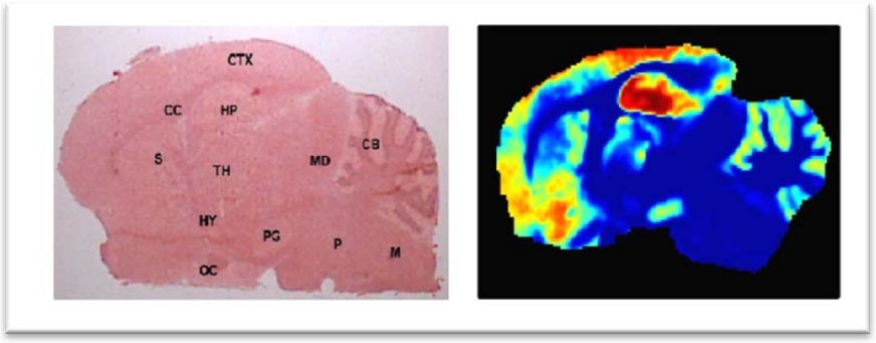


Figure 28: Comparison between histopathology and MALDI-MSI.
Pictures showing a histopathological analysis of brain tissue (left), in which a histopathologist has identified the functional regions of the brain, and a multivariate statistical analysis of brain tissue (right). Taken from Judy's thesis, and later published in Fonville, Carter et al. (2012b).

Before Judy could draw biological conclusions from MALDI-MSI data, she had to experiment with a variety of multivariate statistical techniques to model and visualize molecular information. Highlighting the advantages and disadvantages of each technique, she generated a set of strangely beautiful pictures that depicted colourful regions and patterns within the rat brain tissue. In some ways, Judy's research was an attempt to cope with the enormous amount and complexity of data that the MALDI-MSI analysis generated. She explained to me that MALDI-MSI generated files so huge that the data analysis programs on her computer often crashed. I later learned that state-of-the-art, high-resolution MALDI-MSI machines could generate file sizes greater than two gigabytes per tissue slice. This large volume of data required scientists like Judy to find new ways of working with biochemical data, which could reduce the size of data without compromising its complexity.

Similar to the researchers I discuss in Chapter 3, Judy emphasized the importance of using complex and exploratory multivariate approaches to the MALDI-MSI data. Such techniques were crucial to the success of her work, but were not a common practice among researchers working with MALDI-MSI data. Many researchers, Judy claimed, "discarded all of the data": they only looked at small pre-determined anatomical regions of interest within the tissue, or focused exclusively on a handful of preselected biochemicals. Such approaches limited the amount of information researchers could glean from the data.

Such multivariate methods, Judy explained, allowed her to "extract" the meaning from her data by comparing many variables, rather than just one or two, simultaneously.

Judy's emphasis on complex multivariate statistics analysis highlights the values that metabonomics researchers place on certain knowledge forms and practices (see Chapter 3). In asserting that more "simple" forms of data analysis were poorly-executed or problematic, Judy placed value on "unsupervised" multivariate statistics, methods used to find relationships within data without relying on prior knowledge about those biochemicals or features of the data that were worth examining. This might entail, for example, comparing animals labelled as "healthy" and "diseased," and allowing the statistical analysis to find differences among the animals without prior knowledge of their health. Judy asserted, moreover, that unsupervised multivariate statistics provided an "objective" and "unbiased" means for researchers to explore those relationships within the data that were not readily apparent. In doing so, she obscured the fact that experiments can never be without the influence of values, world views, or the bias of researchers (Räsänen et al. 2013). Data and the techniques through which it is produced are "always structured according to somebody's predispositions…and value choices all the way through" (Filler 2009).

For Judy, multivariate statistics not only allowed her to process large data files, but also gave her the ability to make sense of data that was too complex, held too many data points and patterns, to be interpreted by eye. An analysis of MALDI-MSI data was impossible to do by hand, because each tissue slice contained twenty thousand pixels and tens of thousands of chemical peaks. Judy relied on multivariate statistics, therefore, to find patterns and meanings that were hidden within biochemical data, and which would otherwise be inaccessible through visual analysis. She said:

> [Simple statistical methods are] always inadequate, because as soon as you do manual things, you have to go through 5000 images, or you're looking at lots of individual peaks. Then you should think to yourself, hold on, there is way more information than I can cope with in one glance. So it's probably worth doing [multivariate statistics], so I'm not missing things. Because you're very likely to miss things if you do it by eye.

Judy asserted that multivariate techniques like principal components analysis (PCA) and self-organizing maps (SOMs) fundamentally enabled researchers to see the relationships and differences between various regions or substructures within rat brains. Like the researchers I have discussed in Chapter 3, she valued the capability of multivariate statistics to show the combined effects of multiple biochemicals, or "the associations between different [parts of the data]" (Fonville et al. 2012b:B). As another researcher working on MALDI-MSI commented, "It's [all] about generalized pattern recognition. We don't want to see one

image by itself, it doesn't mean anything. Because we want to compare, otherwise it's been done already. It's a comparison between images [that matters.]" Judy commented:

> You would not have seen the structures with any individual [point along the tissue], simply because it's a combination of different effects that…shows the structure. That's why you should do multivariate methods, and how [they] can show you extra information.

As Judy experimented with statistical practices, she described the differences between regions of brain tissue as mathematical patterns, comparisons, and juxtapositions, using the statistical language of "pixels," "clusters," and "m/z values" (Fonville et al. 2012b:B). By experimenting with different unsupervised multivariate statistical techniques, Judy could influence and visualize the boundaries between substructures in the brain based on biochemical similarities and differences. To this end, PCA showed the overall differences between gray and white matter, while SOMs showed the differences between anatomical substructures like the corpus callosum and hippocampus (see Figure 29). Thus, in using different statistical techniques, Judy envisioned biological tissues and processes as a series of molecular patterns. Her concern was not with identifying the biological composition of the tissues, but rather with showing their statistical relationships and meanings. As a clinician collaborating with Judy commented, MALDI-MSI was surely complex, "but you don't actually need to know what you're looking at, it could just be patterns clean and simple."

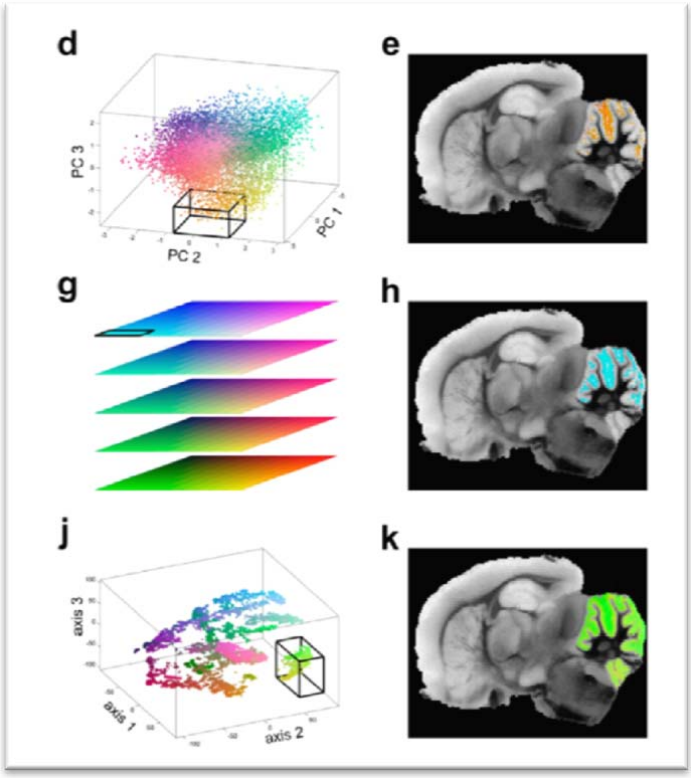


Figure 29: Different multivariate statistical methods showing different brain structures.
Pictures taken from a paper published by Judy on her research, showing how three different multivariate statistical methods can reveal different brain substructures (Fonville et al. 2012b). PCA (top left) shows gray matter (top right), while SOMs (middle and bottom left) show the corpus callosum (middle right) and hippocampus (bottom right).

In summary, Judy's understandings of molecular patterns and brain substructures were intricately tied to her use of multivariate statistical methods. Judy articulated the value of multivariate statistical analyses in multiple ways: she claimed that they provided her with "hidden" or visually inaccessible molecular information, and allowed her to conceptualize rat brains as being made up of statistical patterns and relationships. She therefore enacted tissues as patterns of statistical information, reducing the differences between brain structures to statistical thresholds and boundaries. Thus, this section shows how metabonomics researchers' ability to "see" substructures in the brain—to glean meaning and value from MALDI-MSI data—are intricately tied to multivariate statistical practices and technologies. Multivariate statistical practices are, as Evelyn Fox Keller writes, "models of and models for"

(Keller 2000): they are the means by which researchers conceptualize tissue, but they are also the material basis for experimental activity[65].

Ultimately, Judy's research on MALDI-MSI articulates how *the forms of data that are considered interesting and useful to metabonomics researchers are both created and made meaningful through multivariate statistical practices*. Multivariate statistical techniques give rise to numerical calculations and complex visualizations, but also enable these calculations and visualizations to have meaning and value within the realm of metabonomics practices. As Anne Beaulieu (2001; 2004) writes about bioinformatics-driven research on the human brain, new objects of knowledge are intertwined with new ways of knowing: data practices like multivariate statistics entail not only new materials and technologies, but also new types of knowledge. Thus, Judy's research on MALDI-MSI highlights the centrality of data and informational practices to translational research, and shows how they are not only supportive but also integral to the production of certain biomedical objects.

Metabonomics Meets Histopathology

At a later point in my fieldwork, Judy presented her MALDI-MSI research to a group of clinicians based in St. Mary's Hospital. This meeting of researchers and clinicians occurred under the banner of the NIHR-BRC funded projects in the BMM, which encouraged practicing clinicians to provide input on the development of laboratory technologies that were being translated into clinical practice. Encouraged to present her work as a tool that could be used by clinicians in everyday research, Judy had taken the statistical methods that she used in the analysis of rat brain tissue, and packaged them into a computerized interface called a "toolbox." This was in an effort to present the complex methods used in metabonomics research in a simplified fashion, such that clinical researchers could apply Judy's data analysis methods to their own research endeavors. In essence, Judy presented her research, which I had seen as a work-in-progress as she completed her PhD, as a finished and refined product.

[65] The use of Evelyn Fox Keller's work to interrogate the practices surrounding models in biology raises philosophical questions about the differences between representations of statistical data and models of biological research. Although I see statistical practice as both a means to and a source of information allowing for intervention into biology, my concern is not with the ways in which statistical visualizations are deemed as *representative* of a type of biological reality (Ankeny et al. 2011:315-16).

Standing at the front of a dimly lit seminar room, Judy contrasted the benefits of "modern" MALDI-MSI technology with "dated" histopathological approaches. Histopathology, as discussed in Chapter 4, plays a central role in the diagnosis of disease, and has been the gold-standard of tissue analysis since the early 20th century. It is carried out by highly-specialized professionals who examine stained cells under a microscope, and who look for morphological differences between normal and abnormal tissues. As Ilana Löwy writes in *Preventive Strikes: Women, Precancer, and Prophylactic Surgery* (2009), histopathology has been involved in the understanding and diagnosis of cervical and breast cancer since the early 20th century. Histopathologists "learn how to recognize forms and patterns and to classify shapes" with skills that are based in individual experience. Their skill emerges through trained judgments, which try to transcend the variability of human biology to produce biomedical knowledge that can be used to diagnose disease.

In her presentation to the clinicians, Judy asserted that MALDI-MSI could provide a more objective view of biology, because it relied on molecular and statistical technologies rather than the subjective decisions of histopathologists. Echoing researchers' comments in Chapter 4, Judy explained that MALDI-MSI could reveal the "hidden" abnormal processes happening in visually normal tissue, as large quantities of molecular information could be used to quantitatively measure the extent and nature of disease. Thus, Judy's work was one of several attempts by researchers in the BMM to find a molecular alternative to histopathology, in an effort to compare the biochemical composition of the tissue with the inferences of the histopathologists.

Comparing MALDI-MSI and histopathology, however, was not without difficulties. At a basic level, researchers struggled to compare the format and resolution of MALDI-MSI images to the format and resolution of histopathological images. While MALDI-MSI images were generated and analyzed by computers, histopathological slides were analysed by eye and were therefore not commonly digitized. And when the researchers were able to examine the digitized slides, they saw that there were issues of scale. While histopathology resolved images of individual cells, MALDI-MSI resolved images with "chunks of cells in each pixel." This presented key problems to the comparative analysis of the two techniques, as the researchers could not easily overlay the two images. Such problems were reflective, in the end, of the different goals that MALDI-MSI and histopathological analysis entailed. While MALDI-MSI sought to biochemically characterize tissue slices as a whole, histopathologists looked for specific biological markers like vascular invasion and

angiogenesis at the cellular level. As one of the metabonomics researchers commented in frustration: "We're unclear whether the histopathologists understand what we need to do or are able to do."

Despite these challenges, embedded within Judy's presentation was the suggestion that MALDI-MSI could one day provide a superior alternative to histopathology. Though most researchers working on translational projects acknowledged that their work would likely operate in parallel to existing clinical practices, Judy insisted: "You would want to show that you can do *more* than histopathology." In response, the physicians to whom Judy was presenting began a heated discussion. They wondered: how could a statistics-based judgement of biology replace a time-honored practice like histopathology? Would MALDI-MSI be able to inform disease diagnosis and treatment with the same success as histopathology? Comparing Judy's molecular methods to histopathology, one clinician commented:

> I think the main difference here is that you're effectively looking at the metabolic signature of anatomy. And in our [work] we're effectively looking at cancer, which is an entirely different ballgame.

According to the clinician, MALDI-MSI visualized tissue as a "molecular signature of anatomy," while histopathology visualized tissue as a disease process. In doing so, the clinician articulated the difference between examining biology as a set of localized statistical and biochemical signals, and examining biology as a holistic process affecting a patient. To this end, another clinician commented:

> Also, for me, it's even that…of course you're going to add a whole lot of information that we simply don't have. But the real thing is to take the information and go back to the tissue, and say what is, I mean, why do you get this. Because one thing would be to make a machine which would detect cancer or not cancer, which could be really clever. But the real question is, what is this telling us about the pathogenesis that we would not know in any other way.

Though both clinicians acknowledged that MALDI-MSI could provide a new perspective on the biochemical composition of tissue, they were skeptical that MALDI-MSI could help them reason through the diagnosis and treatment of disease. In many ways, this skepticism was well-placed: metabonomics researchers continually emphasized that their main challenge was not in generating statistical information, but rather in making sense of it, a point to which I return later in this chapter. Thus, for the clinicians, histopathology provided a window onto the biological processes that ultimately affected the diagnosis and treatment of disease. It was valuable not because it shed light on tissue structures, but rather because it provided insight into the dynamic nature of biological outcomes.

Consequently, MALDI-MSI and histopathology enacted fundamentally different biomedical objects. While Judy's research enacted tissue structures as *statistical thresholds*, histopathological practices enacted tissues as *disease processes.* As one clinical researcher described, histopathologists looked for specific visual features—at both the microscopic and cellular level—as a way to comment on health and disease. They used microscopy to look for markers of vascular invasion and angiogenesis, not only to assess the stages and grades of tumors, but also to formulate recommendations for disease diagnosis and treatment. In this sense, MALDI-MSI and histopathology entailed not only different logics about disease, but also different approaches to disease diagnosis and treatment. While Judy sought to maximize information and make sense of tissues in terms of statistical thresholds and patterns, histopathologists sought to relate tissue morphology to disease processes and to maximize the health of the patients from which the tissue had come.

Ultimately, this sub-section shows how different realms of practice—different configurations of technologies, ideas, people, and values—lead to the enactment of different disease objects. It contrasts the statistical patterns enacted in metabonomics research with the disease processes enacted in histopathological tissue analysis. Conflicting ideas about how to diagnose and treat disease arise not simply because of different "cultures" of biological research, but also because different realms of practice enact and perform different versions of disease. Thus, this section argues that informational and statistical metabonomics technologies, though they are applied to clinical issues, are inherently distanced from clinical questions and goals. It speaks to the fact that, as Ilana Löwy (2009) writes, scientists and practitioners have different ethics and values: scientists and clinical practitioners focus, respectively, on the description of biology and the coordination of medical interventions. Consequently, the enactment of different biological objects also entails the normalization of different practices and values. This is not to say that clinical practices are better than laboratory practices, but rather to emphasize that the practices and objects of metabonomics research are aligned with the objective production of biological facts rather than the normative outcome of patients (Löwy 2009:62).

Applying Multivariate Statistics to Clinical Data

The coproduction of statistical practices and knowledge, which forms a key focus of the previous section, highlights the centrality of informational practices to translational medicine research. It emphasizes the increasing role that data analysis methods and statistical

techniques play in making sense of and in "translating" biological data into medical knowledge and interventions. In this section, I examine how researchers place value on increasingly complex and large volumes of data, and perceive data as central to the success of translational medicine research. This data is generated by the application of multivariate statistical techniques not only to molecular and biochemical data, as I discussed in the previous section, but also to existing clinical data consisting of routine laboratory tests and physiological measurements. Thus, this section explores how the increased value that metabonomics researchers attribute to complex and large datasets eclipses the fact that what counts as "data" in the first place is changing.

Noah, a former PhD student who had become a Junior Research Fellow, was working on the application of multivariate statistical methods to clinical data. While many of the researchers in the BMM had experience in molecular biology and biochemistry, Noah's expertise was in chemometrics. His work in the BMM focused on the development of both pattern recognition techniques for the analysis of biochemical data, and also user-friendly interfaces for statistical methods, both of which required computer-science programming skills. At the time of my fieldwork, Noah had begun work on a project that focused on patients diagnosed with sepsis, a life-threatening inflammatory condition. This was based on clinical data from within the Intensive Care Unit (ITU) at St. Mary's Hospital, one of Imperial's main clinical research centres. Overall, Noah's work attempted to combine existing clinical data with new methods for data analysis, visualization, and interpretation.

More broadly, Noah's work was part of a translational medicine initiative within the BMM to integrate existing clinical data with "omics" data derived from metabonomics experiments. This was an attempt to maximize the amount of data that could be used to make diagnoses and predictions about the patients. The researchers involved in the project realized, however, that for metabonomics to work in a clinical setting, it would first have to interface with existing clinical data. They envisioned a future of clinical practice in which metabonomics data would complement and interface with, not replace, existing clinical methods. Thus, the goal of their research was not only to establish the use of new metabonomics technologies within the clinic, but also *to find new and multivariate ways of interpreting existing clinical data*. The goal, as Jeremy Nicholson described, was "to statistically combine the old and the new."

Such comments highlight the ways in which multivariate statistics exists as both a material technology, and also as a means for epistemic and social control. Multivariate statistics, as I have discussed in Chapter 3, exist within the BMM as highly naturalized and institutionalized ways of interacting with and analysing data. As such, the application of multivariate statistics to clinical data represents not only a new form of data analysis, but also a new means of promoting the knowledge practices that form the core of the BMM's research. Just as the use of statistics and quantitative data methods was seen as tools to unify research in the social and natural sciences (Porter 1988), the application of multivariate statistics to clinical data is seen by metabonomics researchers as a means for unifying post-genomic and clinical research (see Chapter 2).

Before implementing metabonomics in clinical settings, Noah wanted to first establish whether it was possible to analyse clinical data in a multivariate way. He attempted to apply the multivariate statistical methods that were commonly used in the analysis of metabonomics data—such PCA and PLS (see Chapter 3)—to clinical data. But he also attempted to develop new software and informational technologies to integrate multiple types of data.

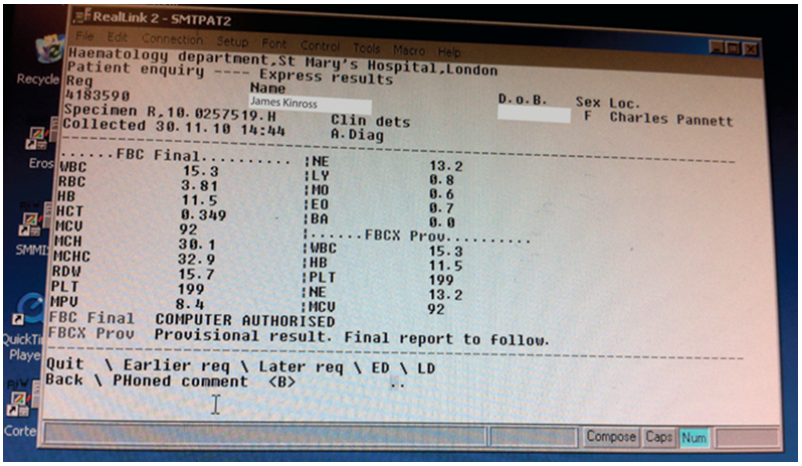


Figure 30: Univariate statistics in the ITU.
A screenshot of the data infrastructure and visualization that was used in the clinical setting of the St. Mary's hospital intensive care unit (ITU). Note how clinical measurements (abbreviated at left) are given in single values in isolation. Taken from a Powerpoint presentation given during a seminar.

To this end, Noah found the existing methods by which clinical researchers examined clinical data shocking. He described the software interface within the ITU as a drab grey

screen that was sparsely populated with numbers, while another clinical researcher who worked in the ITU commented "it's like an excel spreadsheet from hell." Beyond the aesthetic problems with the clinical software, Noah's main criticism of the clinical data was that it was presented in a univariate and "oversimplified" way. This made it impossible, said Noah, to visualize the patterns and complex statistical relationships contained within the clinical data. Describing the clinical software and data, he said: "so from my personal perspective, I'd say it's a kind of stone age technology" (see Figure 30).

In describing his reaction to the existing technological interfaces at St. Mary's hospital, Noah emphasized—like Judy—the importance of using multivariate rather than univariate approaches to data analysis and visualization. Multivariate approaches would allow researchers to explore the complexity within their data, and to uncover underlying and previously hidden patterns. In highlighting the value of statistical relationships within clinical data, Noah emphasized *the generation and analysis of particular—and statistical—kinds of "data."* As such, he placed value on multivariate forms of data—which consisted of large volumes of complex information —as the "correct" and "ideal" way of diagnosing and treating disease. Like Judy, Noah enacted biology and disease as a statistical object, referring to biological features with the statistical language of "parameters" and "features."

Beyond the challenges with the existing clinical software, Noah asserted that the main challenge of his own work was in determining how best to analyse and visualize complex datasets. This involved, on the one hand, the standardization of multiple types of clinical data, and on the other hand, the aggregation of large volumes of biomedical data. In this way, Noah's work embodied metabonomics' idea that the best way to learn about disease was to collect as much data as possible, in a process Noah described as "data mining for improved information recovery." Researchers like Noah believed that creating more powerful tools to look for statistical relationships and patterns, and to collect more data, would eventually translate into the improved diagnosis and treatment of disease.

Jacob, an Honorary Lecturer at the BMM whose work I also discuss in Chapter 3, was involved in a similar project to apply multivariate statistics to clinical data. Jacob was tasked with analysing a large dataset comprised of information from more than 30,000 ITU patients, with the end goal of trying to better predict the occurrence of pneumonia. This dataset, which was open-source and had been collected by researchers at the Massachusetts Institute of Technology (MIT), not only included a much larger number of patients than the St. Mary's

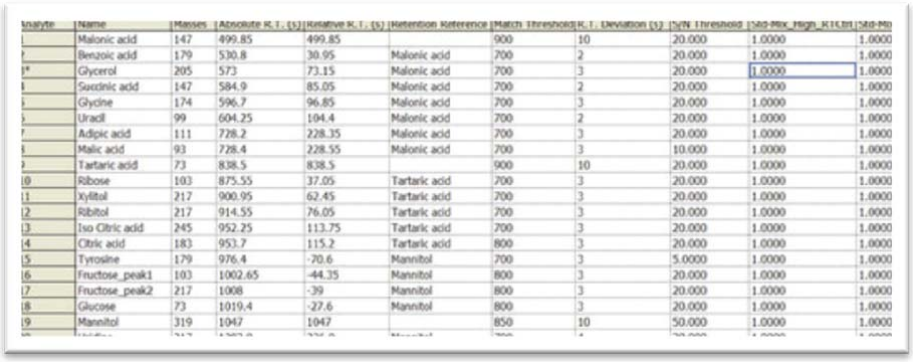
ITU dataset, but also included many additional biological parameters such as five-year follow up information.

When I spoke to Jacob about his work on the MIT dataset, he described the overall project, similar to Noah, as an attempt to analyse clinical data with the multivariate methods commonly used in metabonomics. This involved, first and foremost, formatting routine clinical observations about disease—such as blood glucose, blood oxygen levels, heart rate—into a digital database, which Jacob planned to subsequently analyse with multivariate statistical methods. By carrying out a multivariate analysis, Jacob attempted to "see if [the] part[s] can be linked together…to provide more information." Like Noah, he emphasized that the overall goal of his work was to learn about disease by collecting and analysing as much data, in the statistical sense of the word, as possible.

The translation of clinical information into multivariate data, however, was not a straightforward task. The MIT dataset required significant manipulation, as the documentation of clinical parameters was done in an incomplete and "messy" way. Different clinical parameters had been recorded for different patients, and sometimes with different scales or formats. In this sense, Jacob's work with the MIT clinical dataset speaks to the broad challenge involved in the standardization of complex data from multiple sources. As Berg and Timmermans (1997) have argued, the practices of standardization seek to make data universally meaningful, but do so in a local way. Standardization is the result of active and situated practices, and encompasses a degree of flexibility that allows the same data to have meaning in different contexts. As Jacob described, his work involved "curating, sorting, concatenating…just to get a nice rectangle full of data." Jacob's discussion of standardization, therefore, highlights the creative and innovative practices required to transform and standardize clinical data into a multivariate form.

Speaking to Jacob, the work required to apply multivariate statistical methods to the MIT clinical dataset seemed like a monumental challenge. Jacob asserted, however, that the work would be no different from working with the NMR or MS datasets metabonomics researchers routinely used. For Jacob, analysing clinical data involved the same methods and practices as analysing metabonomics data: it involved building a data "matrix"—a two-dimensional table composed of rows and columns filled with numbers—and looking for patterns with multivariate statistical methods (see Figure 31). Though the type of data contained within the MIT dataset was different, the form it took and the methods it required

for analysis were the same. "You build a table in a consistent way," Jacob said, "And after that, all of your data is always the same."



Analyte	Name	Masses	Absolute R.T. (s)	Relative R.T. (s)	Retention Reference	Match Threshold	R.T. Deviation (s)	S/N Threshold	MS Mix_High_RTLIB	MS Mix
1	Malonic acid	147	499.85	499.85		900	10	20.000	1.0000	1.0000
2	Benzoic acid	179	530.8	30.95	Malonic acid	700	2	20.000	1.0000	1.0000
3*	Glycerol	205	573	73.15	Malonic acid	700	3	20.000	1.0000	1.0000
4	Succinic acid	147	584.9	85.05	Malonic acid	700	2	20.000	1.0000	1.0000
5	Glycine	174	596.7	96.85	Malonic acid	700	3	20.000	1.0000	1.0000
6	Uracil	99	604.25	104.4	Malonic acid	700	2	20.000	1.0000	1.0000
7	Adipic acid	111	728.2	228.35	Malonic acid	700	3	20.000	1.0000	1.0000
8	Malic acid	93	728.4	228.55	Malonic acid	700	3	10.000	1.0000	1.0000
9	Tartaric acid	73	838.5	838.5		900	10	20.000	1.0000	1.0000
10	Ribose	103	875.55	37.05	Tartaric acid	700	3	20.000	1.0000	1.0000
11	Xylitol	217	900.95	62.45	Tartaric acid	700	3	20.000	1.0000	1.0000
12	Ribitol	217	914.55	76.05	Tartaric acid	700	3	20.000	1.0000	1.0000
13	Iso Citric acid	245	952.25	113.75	Tartaric acid	700	3	20.000	1.0000	1.0000
14	Citric acid	183	953.7	115.2	Tartaric acid	800	3	20.000	1.0000	1.0000
15	Tyrosine	179	976.4	-70.6	Mannitol	700	3	5.0000	1.0000	1.0000
16	Fructose_peak1	103	1002.65	-44.35	Mannitol	800	3	20.000	1.0000	1.0000
17	Fructose_peak2	217	1008	-39	Mannitol	800	3	20.000	1.0000	1.0000
18	Glucose	73	1019.4	-27.6	Mannitol	800	3	20.000	1.0000	1.0000
19	Mannitol	319	1047	1047		850	10	50.000	1.0000	1.0000

Figure 31: Multivariate statistical data matrix.
An example of a data matrix, an array of information used by metabonomics researchers to collect and organize data (Chan et al. 2011)

Overall, Jacob emphasized that the study of health and disease could be "optimized" by collecting and statistically analysing as much data as possible. Like Noah, he claimed that it was important to find and generate patterns within clinical data, and believed that this would lead to improved disease diagnosis and treatment. However, Jacob's example shows not only the value metabonomics researchers place on the collection and analysis of large volumes of data, but also how the very nature of what counts as "data" within translational research is changing. While the clinical parameters that are commonly observed and recorded within the ITU have specific meanings and engender certain practices for clinical researchers, within metabonomics research they do not count as data. Data within the post-genomic context of metabonomics research is something multiple and relational: it is enabled through large and specially-formatted datasets, and requires the practice of multivariate statistics.

To conclude, in the previous section I explored how as researchers enact tissues as complex molecular entities and statistical patterns, they highlight the centrality of data and informational practices to translational research. In this section, I have shown the value that metabonomics researchers place on such data, as they claim that the collection and analysis of large volumes of data has the potential to transform disease diagnosis and treatment. However, what counts as "data" is highly contextual, and researchers only attribute meaning to data once it takes on a particular multivariate form, and is analysed with standardized

multivariate practices. Consequently, as this section describes the application of multivariate statistical practices to clinical data, it highlights how metabonomics researchers' attempts to change the form of clinical data is also an attempt to imbue such data with new and "better" meanings.

Making Sense of Metabonomic Information

While the previous section explored the value placed on multivariate data for the understanding and treatment of diseases, this section explores the challenges faced by researchers in the interpretation of such data. In doing so, it builds on the recurring challenge—as articulated by metabonomics researchers themselves—of making biological sense of data. Expressing this challenge in relation to his work on clinical databases, Jacob emphasized: "The need here is actually…to define what is the biological question or the clinical question. So you've got the database. So what do you want to know?…We have to define what exactly is the question, say, what do you want to see."

This expression was a common refrain throughout my fieldwork: it was articulated in many forms throughout my interviews with researchers involved in clinical projects, and discussed as one of the main doubts researchers had about the potential of the work to influence medical practice. Fundamentally, the expression highlights the tensions faced by researchers who struggle to make biological sense of data, and yet emphasize the fundamental role its collection and organization plays in translational research. Despite the overt value metabonomics researchers place on the production and use of multivariate forms of data, they still acknowledge that the interpretation of such data poses a serious challenge to the application of metabonomics practices to clinical issues. This section explores, therefore, how metabonomics researchers struggle to translate their findings into clinical practice, and to make their results meaningful in relation to clinical epistemologies or understandings of the body. It explores how metabonomics researchers struggle to align their findings with the biological pathways and functions that are oriented around patient care and disease outcomes.

Mid-way into my fieldwork, I met a former researcher within the BMM named David to discuss his impressions of metabonomics research. Because he had moved to a different field and university after completing his doctoral research at Imperial College London, he discussed the main strengths and challenges that faced metabonomics research in the BMM:

> And I think that's historically what [the BMM has] done actually quite well, they're really good at the analytical side, and have a really strong understanding of NMR and how to set up these experiments…So the other side, which is more kind of understanding the biology behind an event or a perturbation of the biological system, I think that's trickier in general, in biology.

The BMM, David emphasized, was very successful at the "analytical side" of experiments, at identifying and quantifying the biochemical components within biological fluids and tissues. The laboratory had discovered a large number of biomarkers, the quantifiable end-products of metabolism that could be correlated with health and disease (see Chapter 3). It had also generated an undeniably large volume of papers, which it published in a variety of medium- and high-impact journals. Despite its research productivity, however, David emphasized that the BMM *struggled to understand and interpret the meaning of the biomarkers* its researchers produced. It struggled to relate concrete biological processes to specific genes, metabolic pathways, or bodily systems.

Claire, another researcher whose work I discuss in more detail in Chapter 3, made similar comments about the difficulty of biologically interpreting metabonomics results. When I asked her about the practical challenges she encountered in her everyday research, she said:

> It's not necessarily that it's too much information, it's just that it's complicated to put it all together in a meaningful fashion…We're still at a stage where, okay, x metabolite goes up and y metabolite goes down. And we don't really know what that means, and we don't really know whether that's translatable [or] reproducible. And there's a huge pressure on this from other scientific fields to…actually figure out why these metabolites are up or down.

Like David, Claire emphasized how metabonomics researchers struggled to interpret biochemical data in relation to biological processes or bodily functions. She emphasized that it was not the volume of information, but its arrangement into meaningful patterns and associations that posed a challenge. "You work with a hundred character alphabet," she said, "[And] using a language that no one really knows that well." According to Claire, metabonomics was successful at establishing statistical relationships, or at correlating changes in metabolite levels to disease states. It struggled, in contrast, to relate such results to meaningful biological pathways or disease symptoms. Claire questioned whether the biochemical and statistical methods of metabonomics experiments could be translated, applied to, and used in clinical settings.

In a similar way, a lecturer in the BMM remarked to me during an interview that it was "easy to hide behind the numbers" in metabonomics experiments, especially when working with statistical relationships and outputs that were abstract and easy to manipulate.

Metabonomics research, he claimed, was comfortable and successful with number crunching, but not with establishing how such numbers could relate to the human body's functions. The lecturer asserted, therefore, that metabonomics researchers *struggled not to produce but to translate* their findings into biomedical applications. They struggled to convince medical professionals and researchers in other fields—not to mention the Federal Drug Administration (FDA)[66], the body responsible for approving the clinical use of metabonomics technologies—that their biochemical and statistical methods could succeed in clinical settings[67].

In general, the interpretation of metabonomics data was made difficult for several reasons. Firstly, the same biochemicals tended to recur across multiple experiments and analyses, making their biological relevance unclear. As a doctoral student named Edward commented to me, biomarker panels, the groups of biochemical markers I discuss in Chapter 3, tended to be composed of the same "common" biochemicals. For example, lactate and hippurate, which play a role in cellular respiration and microbial metabolism respectively, were features of almost every experiment. The recurrence of common biochemicals was due to the fact that many compounds were involved in multiple metabolic pathways, or as one researcher called them "metabolic hubs." Thus, many biochemicals were the end result of multiple biological processes occurring simultaneously within an organism. Researchers questioned if common metabolites were detected because of disruptions of cellular respiration, the use of particular medications, or the ingestion of certain foods. In this way, the complex nature of metabonomics data—the fact that it was the end product of many biological processes—made its interpretation challenging.

The interpretation of metabonomics data was made difficult, secondly, because the biological origins of the biochemicals that NMR and MS instruments detected were not always clear. I spoke with a researcher named Thomas about the challenges involved in

---

[66] As of the writing of this dissertation, no metabonomics-based biomarkers have been approved by the FDA for clinical use.

[67] One researcher who was writing an editorial about the successes of metabonomics during the early 20th century received the following feedback from his reviewers:

> Completely missing is the need to say that this type of biomarker discovery is not biomarker validation. The whole field of metabolomics is stalled in the phase of discovering biomarkers that are never subsequently checked by applying them to prospective studies, or even re-analyzing the study samples with specific methods to check the validity of the profiling-derived hypothesis.

making sense of the data generated by a technology called the "intelligent knife." The intelligent knife was a surgical device that analysed the molecular composition of the "chemical smoke", the ionized tissue particles, created during electrocautery, in which the standard surgical blade was replaced by a device that cuts with an electric current and rapidly cauterizes tissue. With the intelligent knife, researchers attached the surgical device used to carry out electrocautery to an MS machine, which detected the chemical composition and patterns within the chemical smoke (Imperial NIHR Biomedical Research Centre 2011).

According to Thomas, one of the main issues with making sense of the data generated by the intelligent knife was in figuring out what exactly—what tissues, what types of biochemicals—the machine was measuring. The intelligent knife was an incredibly complicated device that attempted to make real-time measurements about the composition and nature of tissues. In his work with the intelligent knife, Thomas had difficulty understanding whether the chemical smoke was coming from tissues at the surface or from tissues deep within the surgical incision. Knowing the origin of the smoke was fundamental, because it had implications for the types of cells or biological pathways implicated in surgical treatment. If, for example, the intelligent knife was detecting surface-level cells, it would not necessarily show biomarkers for specific, localized disease conditions.

Interpreting the data from the intelligent knife was made further difficult by the uncertainty surrounding the range of biochemicals that the device was able to detect. The intelligent knife, like other analytical instruments, had inherent capabilities and limitations (see Chapter 3) that made it suitable for the detection of a certain range of biochemicals. This, as Thomas said, raised questions about whether the machine would be able to detect those biomarkers that were implicated in health and disease. Thomas commented:

> The other thing is that we're at the mercy of what we're able to see. Okay, so, when you cauterize a tissue, what we're seeing is phospholipids…cell surface membrane stuff. So, that's one of the biggest challenges right now, is exactly that…we see fatty acids, phospholipids, lysophospholipids, fragments of phospholipids. So that's what we're limited to right now in our characterization of tissues…We're only seeing lipids. There's so much of a metabolome out there, and we're just able to tell tissues apart by lipids because that's what we see.

Thomas emphasized that the intelligent knife could only detect fat-containing molecules that occurred at the surface of cells, whose importance in surgery and disease diagnosis was unknown. Metabonomics researchers were, as Thomas described, "at the mercy" of the machine's technical capabilities. Though they were able to build customized statistical algorithms to analyse the machine's data, they had to operate within the parameters of the

machine's commercially-determined settings. Thus, metabonomics researchers struggled to interpret the biological meaning of the intelligent knife data, primarily because they could not always say whether the biochemicals it detected played a key biological role.

In conclusion, I have argued that the broad challenge facing metabonomics researchers is that of the interpretation—rather than the generation—of experimental data. Researchers continually question how their statistical and biochemical data can be "translated" into biological pathways or bodily functions. The links between biochemical and statistical data and states of health and disease are not pre-given or objective, but rather are enacted through the everyday work of metabonomics research. As such, metabonomics researchers struggle not to produce situated forms and values of data, but rather to make such data meaningful in relation to clinically-relevant understandings of the human body. Ultimately, when metabonomics research is placed alongside clinical practices, the value researchers place on large volumes of complex data is rendered problematic. Metabonomics researchers struggle to determine how biochemical data can be used to ask particular questions, or to see how it can be made meaningful in relation to disease treatments or outcomes.

Translation and the Role of Human Judgment

In the previous sections, I explore the processes and challenges associated with the movement of knowledge between the metabonomics laboratory and the clinic. This section steps back slightly from the realm of everyday metabonomics practices, in order to examine how researchers envision the future "translation" of metabonomics technologies into clinical technologies and solutions. It examines how such visions tell us about the different forms, uses, and values of data that exist at the laboratory-clinic interface. It also examines how such visions overshadow the interpretive issues highlighted in the previous section, by portraying translational research as a technological rather than a practical feat. It prompts us to ask: what limits exist—if any—for implementing statistical visions of health and life in clinical practice?

I spoke at length with William—a surgeon in the NHS who specialized in bowel cancer, and who was also Joseph's supervisor—about the future visions and possibilities of metabonomics technologies in clinical settings. William was the clinical coordinator of the BMM's translational research projects funded by the NIHR-BRC, and thus was involved with a variety of clinically-oriented projects. William had been one of the first clinician-

researchers to spend an extended amount of time doing metabonomics research in the BMM, having completed his PhD in metabonomics several years prior. His current role in the laboratory was as a clinical lecturer with 50% research and 50% clinical responsibilities, an allocation of time which—echoing the conflicts between Joseph and Sarah—he described as "just a lie." He said: "Okay, it just doesn't happen, you do 100% clinical all of the time, and then you do 100% research all of the time. And that's just the way it works. Because it's just impossible, you can't care for your patients and do basic science."

As the leader of the clinical aspects of the BMM's translational research projects—which involved facilitating communication and collaboration between clinical and laboratory researchers, writing funding proposals, and promoting the laboratory's research at public events—William had developed a concrete vision of the translation of metabonomics technologies to clinical settings. For William, the value of metabonomics was in its ability to measure, model, and provide information about complex surgical interventions. Echoing the characterizations of researchers and clinicians mentioned at the beginning of this chapter, he emphasized that surgeons had little knowledge of the biological processes underpinning surgical treatments, or of how patients responded to them. "It's a dense, complex system… and in surgery we have no measure of this system at all, it's totally primitive." Surgeons, therefore, turned to metabonomics for a way to make surgery more "scientific." Echoing previous criticisms of histopathology, William hoped metabonomics would transform surgery from a profession based on subjective human experience to a technological intervention based on objective data.

In asserting the need to make surgery more scientific, William obscured the fact that surgical skills are a combination of technological innovation and bodily know-how (Prentice 2005). Though surgical practices contribute to the production of objective and docile bodies (Hirschauer 1991), they also draw heavily on tactile and sensual knowledge. Similar to histopathology, surgical knowledge is produced through experience and physical interaction with bodily substances. Thus, in asserting that surgical knowledge was subjective, metabonomics researchers placed value on molecular techniques for diagnosing and treating disease, and obscured the contingency of their own practices. They conceived of knowledge about disease as being pre-existing or "housed somewhere inside the surgeon" (Prentice 2005:861), instead of conceding that knowledge was created through the surgeon's interaction with the physical body of the patient.

Consequently, William envisioned a future of "systems biological translational work" in surgery. Though metabonomics technologies required large machinery, as well as the constant monitoring and intervention of human beings, William saw a future in which they were neatly packaged into self-contained boxes. In envisioning a future where clinicians could carry out metabonomics experiments in combination with their treatment of patients, William valued the use of post-genomic data in clinical practice, and also emphasized the importance of standardization and automation. He said:

> It may take my whole career, so that I can walk into an operating theatre, and there can be a machine there that will be a shoebox sized mass spectrometer. And I'll drop the sample in, and the data will come out [as a] lovely, clear data visualization. And it will tell me the information that I need. And [will] use my nanorobot, and I instruct and I'll type on it, and off and it will go. And I won't even have to touch the patient.

Such visions of the future, while they are clearly hypothetical, provide insight into the ideas and values that researchers have about the present and expect for the future (Brown et al. 2003; Wainwright et al. 2006). Consequently, William articulated a future of translational research in which complex multivariate data was simplified into push-button interfaces, such that it could be transformed into a format that would be more easily used and interpreted by clinicians. He explained that part of Noah's work on ITU data, as discussed previously in this chapter, involved not only the reformatting of clinical data, but also the development of streamlined and easy-to-use interfaces for clinical practitioners. In doing so, he emphasized that metabonomics practices in their current form would be too complex for clinicians to understand or apply to patient care. As another researcher named Mary commented:

> William has great visions of mass spectrometers in trauma units or in intensive care unit…And when you design a lot of these experiments, you may not need such a complex instrument… Because with a lot of these, you need an actual surgeon to be able to run it. You're not going to take one of our massive mass specs and shove it in, and expect someone to know how to use it. So you hope eventually it will be a smaller box, and maybe more of a 'yes no' answer to things. Something that's easier to interpret.

By placing value on simplified informational practices, both William and Mary stood in contrast to other metabonomics researchers, who placed value on engaging with the complexity of multivariate metabonomics data. While other metabonomics researchers felt that metabolism and biological life could not be understood without complex informational processes (see Chapter 3), William and Mary emphasized that there were practical limits to engaging with complex information during the diagnosis and treatment of patients in clinical practice. As another clinician-researcher named Andrew commented:

> Clinicians want simplicity, they crave it in their decision making…They all want a simple test, a simple score, that gets them a yes-no answer. And as I see it…the more you simplify something complex, you lose the ability to predict…And what's the balance…at what point does complexity become too difficult as a bedside test?

Though such comments invoke broad generalizations about the capacities of clinical practitioners to carry out and understand certain types of research, they also articulate the central role that human interpretation and judgment play in medical practice. They indicate the centrality of clinical decision-making to patient care, and the ways in which it combines technological information with human intuition.

Such comments not only signal the practical limits to engaging with complex information, but also intimate that clinicians—in their everyday practice with the dynamic human body—deal with and enact complexity in distinct ways. Clinicians, as I discuss in more detail in Chapter 6, rely on their interpretive abilities to judge patients' states of health and disease. In doing so, they draw on different kinds of embodied information (Carmel 2012), which is at times incompatible with biochemical and statistical data. While metabonomics researchers engage with a complexity that is multivariate and biochemical in nature (see Chapter 3), clinicians engage with a complexity that is dynamic and related to biological processes. Thus, within clinical settings, complexity is multiple: it is enacted in different ways, and entails different practices, logics, and disease objects. Seen in this way, conflicts in the realm of translational research arise because the biological complexity articulated in metabonomics practice conflicts with the bodily complexity articulated in clinical practice.

In articulating the differences between laboratory and clinical practices, my aim is not to elevate human interpretation and judgment over the quantitative measurements and inferences that characterize metabonomics research. As clinicians place value on human intuition and leverage their working knowledge of patients in hospital settings, they attempt to assert their authority and control over certain aspects of medical practice. Clinicians see the influx of medical technologies as a threat to medical institutions and realms of power. As Jenny Reardon writes: "To teach the computer to do what life scientists once did is to deskill physicians…and others whose professional life is devoted to interpreting biological life" (Reardon 2011:104). This prompts us to consider not only whether metabonomics technologies and clinical practices entail different values and logics, but also whether some values and logics are better suited to certain aspects of medical diagnosis and treatment.

Throughout my fieldwork, it was not only clinicians but also researchers themselves who articulated a reliance on human interpretation and judgment, and—to a point—a distrust of statistical automation. As I have discussed throughout this thesis, researchers emphasized that multivariate statistics revealed otherwise hidden aspects of biochemical data, and therefore that metabonomics technologies allowed them to surpass the limitations of visual analysis. However, researchers also conceded that handling and inspecting their data manually was critical for assuring the quality of their experimental methods and conclusions. Claire, whose work I discuss in Chapter 3, emphasized that it was important not to completely rely on computers to carry out data analysis, as she said "I'm not sure how much I really trust the data." Many researchers, she asserted, used statistical analysis as an initial means to explore their data, and then used manual inspection to look for interesting differences. Likewise, Ryan, whose work I discuss in Chapter 4, emphasized that it was important not to "let yourself be fooled by the data." For him, statistics were merely a tool, rather than an end-all-be-all for determining if experimental conclusions were obtained by chance. As he said: "People tend to use statistics [if] they don't see anything. So they want to use the statistics to show them something they don't see."

In discussing the future of translational metabonomics research, William made similar appeals to the importance of human interpretation and judgment. He emphasized that clinical practices were inseparable from the decision-making capacities of highly trained and skilled medical practitioners:

> You will have to understand that surgery is quite often always going to come back down to the judgement of the surgeon on the table. And no one is going to take that away. There are very few tests in surgery. By tests in surgery I mean tests that you perform while the patient is on the operating table and you're operating – that will completely determine what the surgeons do.

Despite his visions of a technologically-enhanced surgery, William articulated a future of translational research in which metabonomics technologies complemented rather than replaced clinical practices. As William said during an interview: "All of this is not going to replace what we already have. It's going to augment it." By emphasizing the interdependence of emerging technologies and human capacities, William spoke to Peter Keating and Alberto Cambrosio's (2003) argument that technologies can never truly replace human judgment, and can only serve to rearrange it. As Keating and Cambrosio argue, though technologies attempt to automate biology and transform it into an information science, human judgment is still required to turn "quantitative differences… into qualitative

distinctions" (Keating et al. 2003:59). Consequently, as metabonomics researchers posed technological innovation as a solution to the problem of translation, medical practitioners posed human interpretation as a necessity for the alignment of the laboratory and the clinic. Thus, this chapter shows how visions of the technological future of translational metabonomics research conflict with the inherent appreciation—among both medical practitioners and metabonomics researchers themselves—for the interpretive practices of clinical medicine.

Discussion

Translational research, in attempting to bring metabonomics technologies to the clinic, involves tensions between different research practices, forms of data, and objects of research. Processes of translation between laboratories and clinics are fundamentally problematic, because the laboratory and the clinic entail different realms of practice and enact different biological and disease objects. Thus, metabonomics researchers and clinical researchers have fundamentally different understandings not only of what makes up biology and disease, but also of how biology and disease should be researched and treated. Thus, the divergences and tensions between the laboratory and the clinic should be seen as a "symptom" of the changing practices, objects, and ideologies of biomedical research at the translational research interface (Leonelli 2012). Such divergences raise the question of how—and with what consequences—*data rather than medical practice* is positioned as something central to the diagnosis and treatment of human health and disease.

In doing so, this chapter examines translational research as an informational way of generating and making sense of data, and argues that informational practices are becoming increasingly central to the diagnosis and treatment of human disease. Through metabonomics technologies and practices, biological tissues and disease objects come to be understood as statistical patterns and numerical relationships. As value is placed on the production and analysis of particular kinds—large volumes and multivariate forms—of data, such data is seen as a central aspect of disease diagnosis and treatment. Consequently, metabonomics research typifies what Anne Beaulieu (2001) refers to as "digital objectivity," in which information technologies come to replace the trained judgement of individuals. With digital objectivity, metabonomics researchers make claims to objectivity through the use of statistical calculations and averages, and attempt to eclipse the "manual possibilities" of data analysis or to reveal the "hidden meanings" of data. They emphasize, in other words, the

triumph of data automation and standardization over the trained judgment of interpreters and clinical practitioners.

Such a value on the production, automation, and standardization of data speaks to increasing efforts in the life sciences to transform the handling and interpretation of data into a marketable commodity. As life scientists frequently admit, the costs of generating genetic and other types of data—once a central concern to genetics and genomics following the Human Genome Project (HGP) of the 1990s (Pollack 14 June 2010)—have been outstripped by the costs of turning information into useful data (Patterson 5 December 2011). Advocating for new teams of "algorithms, machines, and people," companies such as "Knome" have emerged whose sole service is number-crunching and interpreting data from other laboratories (Eisenberg 2 February 2013). The costs of making data meaningful, and of interpreting and translating complex information, are a significant challenge for translational research.

To this end, this chapter argues that the value placed on the creation and analysis of data is rendered problematic in the translation of laboratory research into clinical practice. At a fundamental level, metabonomics researchers struggle to interpret and align biochemical and statistical data with biological outcomes, presenting fundamental challenges to the "translation" of biochemical and statistical information into holistic and dynamic understandings of human biology, disease, and biomedical treatment. Thus, as metabonomics portrays translation as a technological feat, it eclipses the challenges involved in aligning the practices and values of the laboratory and clinic. As this chapter shows, data and automation cannot triumph or replace human judgement and interpretation. Such human capacities are still central to the application of metabonomics research to clinical issues, and cannot—at least at this point in time—be overcome with complex types or large volumes of data. Despite the increased availability and value of informational technologies in clinical settings, medical practitioners will not disappear from view. Instead, as Julie Sommerlund (2006) observes about the intersection between naturalist and genetic classifications of bacteria, older practices and meaning systems will continue to work in parallel to new ones.

In conclusion, this chapter is concerned with how we might think about the act and effect of "translation" in metabonomics research, which appears at times as a problematic attempt to implement laboratory technologies in clinical practices. In discussing the challenges that arise at the laboratory-clinic interface, I want to suggest that the translation of

knowledge from one realm of practice to another can never be complete. Translation entails the movement of some types of knowledge over others, as those carrying out the translation select the meanings and values they wish to convey. Thus, as values and practices come into contact and conflict, the process of translation changes its objects of knowledge so that they become workable in new ways or can act as "boundary objects" (Star et al. 1989). In this way, as metabonomics disease objects are translated into clinical practices—or as, in reverse, clinical objects are translated into metabonomics practices—it is not vitalistic but informational ideas and values that become salient.

With this view of translation, the question becomes not whether statistical and biochemical measures of biology can replace human interpretation and judgement, nor whether the enactments of biology are different from those of clinical. The question becomes, rather, if the enactments of biology, along with the techniques for engaging with those enactments, remain the same across multiple registers of practice. In this sense, my discussion of translational research builds on my arguments in Chapter 3 to show how as metabonomics researchers develop new tools to measure biology, they also articulate new modes of thinking about, intervening into, and enacting biology. By developing multivariate and molecular ways of measuring biology, researchers eschew informational, quantitative and statistical values about diagnosing and treating disease. But this presents, as I have discussed in Chapter 4, challenges to engaging with health and disease as inherently dynamic and multiple entities. Consequently, my discussion of translational research shifts considerations, as I discuss in the chapter that follows, to whether multivariate statistics help to improve and standardize medical care, or whether they instead lose track of dynamic patients and instances of disease. As informational changes occur throughout the practices and structures of biomedicine, how does this affect the ability of medical practice to diagnose and treat individual patients and population groups?

Chapter 6: Finding the "Person" in Personalized Medicine

Introduction

I am sitting outside of a pub near Paddington train station in London, chatting informally with several researchers in the BMM. We have just attended a research seminar in the nearby St. Mary's Hospital, at which a recently graduated PhD student named Claire—whose work I discuss in Chapter 3—presented the results of her work to a variety of clinical and laboratory researchers. Though her seminar focused on customized dietary interventions in fetal pigs, it has sparked a lively conversation about the potential for metabonomics research to flow out of the laboratory and translate into healthcare solutions for society.

As we sip on our drinks, a researcher named Anna describes how she participated in a metabonomics study conducted at a laboratory in Italy, in which she worked for one year prior to coming to the BMM. The study was an attempt to determine people's individual "phenomes," a term used by metabonomics researchers to describe the combined effects of genetics and lifestyle, which can be measured with molecules in blood, urine, or tissue. For the study, Anna had contributed several batches of urine over three years, providing a series of phenotypic readouts for her body. From this, the researchers had attempted to determine a unique signature in her urine, which could be used to track Anna throughout the study.

Describing the study, Anna recounts how the researchers had encountered problems with the data from a particular individual, whose biological makeup seemed to be out of the ordinary compared to the other participants. Given that Anna was the only participant who had not been born in Italy, the researchers had assumed that she was the outlier. But after further investigation, the researchers had determined that one of the other volunteers in the study had gotten sick several weeks after her urine was collected, and had been responsible

for an unusual sample with a skewed metabolic readout. This was interesting, says Anna, because the study had inadvertently shown that metabonomics experiments could be used to see "pre-sickness" in individuals. In the future, perhaps researchers could use this research to track a person's metabolism over time, in order to predict[68] and tailor therapies to instances of ill health.

At the mention of this study, the other researchers begin to talk excitedly about the possibilities of using metabonomics experiments to develop implementations of "personalized medicine." This is the idea, widespread in the post-genomic sciences, that medical treatments can be tailored to an individual's biology, and that molecular technologies can be used to direct *the right treatment for the right patient at the right time*. Though personalized medicine is made up of diverse research practices, it represents an epistemic shift away from a "one-size-fits-all" approach to medical care, and towards an approach that evaluates disease risk, treatment response, and safety for unique individuals (European Science Foundation 2012).

Following Anna's discussion of the study in Italy, the other researchers begin to recount how they too have contributed biological samples to studies throughout the world. A researcher named Hannah describes how she established an informal collection of urine around the world by participating in several different studies on personalized nutrition. This is reminiscent of emerging efforts to create "biobanks," repositories for biological materials, in biomedical institutions across the world for the study of health and disease in populations. Consequently, Hannah wonders to her colleagues, could this be used to develop a personal urine repository, from which she could track her metabolism as she moved between jobs at different universities, as she aged, and as she was exposed to different diets and environments? This would allow her to use a collection of her own metabolic measurements to see how she had been healthy or sick at different points in her life, and to correlate periods of sickness with her environmental exposures or biological makeup.

Embedded within these comments, beyond general musings about the future potential of metabonomics technologies, is the assertion that individual biological data is a highly valuable resource for biomedicine. Individual information can be used to determine a

[68] Technically, prediction refers to the likely response of an individual to a particular treatment or intervention, while prognosis refers to the likely overall outcome of individual regardless of treatment.

person's initial metabolic state relative to the combined effect of genes and environment, and consequently to see how he or she will respond to bouts of flu or the ingestion of medicine. Also embedded within these comments, however, is the assertion that individual information about health and disease is only valuable in relation to broader information about populations. Population-wide metabolic information can be used to "calibrate" information about individuals—to determine the ranges for normal and abnormal measurements—and consequently to produce longitudinal models of health and disease. As such, the conversation at the pub emphasizes how metabonomics plays a role not only in the "personalization" of medical diagnoses and treatments, but also in the determination of population-wide markers and patterns of disease.

Personalized Medicine in Historical Context

This focus on personalized medicine forms a central starting point for this chapter, and highlights how metabonomics researchers mobilize notions of normal and abnormal to study disease in particular ways and contexts. Throughout my fieldwork, metabonomics researchers articulated personalized medicine as an attempt to reclassify diseases according to their molecular pathology, and as an attempt to shift diagnosis away from clinical symptoms and towards molecular characteristics. They also articulated personalized medicine as a way to characterize the wide range of individual variation in disease pathology and drug metabolism, and as a way to improve on current clinical methods for diagnosing diseases and patient responses to drugs (Nicholson et al. 2012b).

While the majority of social scientists have characterized personalized medicine in relation to genetic technologies and ideas[69], metabonomics researchers envisioned personalized medicine—like the translational research discussed in Chapter 5—as an informational practice that revolved around the management and interpretation of biological data. By aggregating information about the molecular basis of drug metabolism and disease pathology, researchers hoped to provide more objective medical care based on quantitative biochemical and statistical measurements (see Chapter 4). Consequently, researchers in the BMM largely referred to personalized medicine's association with genetics and genomics as

[69] Arising from the social scientific fascination with "the gene" (Haraway 1997:142; Fox Keller 2002; Nelkin et al. 2004), researchers have explored genetic testing for responses to drugs (Hedgecoe 2004; Rajan 2006) and susceptibility to disease (Rapp 1999; Konrad 2005; Franklin et al. 2006; Gibbon 2007; Lock et al. 2007; Lock 2011), as well as direct-to-consumer genetics (Reardon 2011; Tutton et al. 2011).

a consequence of the Human Genome Project (HGP) of the 1990s. This, they said, emphasized the importance of pharmacogenomics, the development of genetic tests to measure susceptibility to diseases or drugs, particularly for the treatment of cancer[70]. Though metabonomics research was directed at personalized medicine, researchers acknowledged that it was a "buzz word," a term that was associated with particular fashions and trends in scientific research. As one researcher commented: "All of these terms are a bit flexible and they change from week to week…The personalized medicine stuff and the stratified medicine…it means whatever it means to whoever you are talking to at the time."

Personalized medicine, however, is not an entirely new concept. Medicine, with its focus on patient-practitioner relationships, has always been personal (Bates 2010:115). As Richard Tutton writes, the term personalized medicine was used throughout the 20th century to describe patient-centred care, in which physicians practiced medicine as an "'art' of clinical judgment" (Tutton 2012:1721) and provided individualized rather than a one-size-fits-all care for patients. Such individualized care, Tutton argues, is idealized in modern narratives of personalized medicine, and is reconfigured as the achievement of molecular technologies rather than clinical skills (Keating et al. 2003; Löwy 2009). Thus, in contemporary biomedicine the concept of "personalized medicine" has been appropriated by researchers to associate modern molecular medicine with past ideas about individual patient care.

With such a historical perspective, it becomes clear that the "persons" constructed by contemporary personalized medicine are not natural or value-free entities. Instead, as they are investigated in biomedical research, persons become "technogically-assisted categories" (M'charek 2005; Hinterberger 2012), as they are made up and enacted through the technologies that aim to study them. Within personalized medicine, moreover, it is not simply persons but also *populations* that are the object of technological investigation. Despite the discursive focus on "personalized" treatments and interventions, the practices of personalized medicine are fundamentally concerned with comparing groups of people (Rajan 2006; Hinterberger 2012). Through the coordinated practices and ideologies of personalized

[70] Genetic tests have been developed to test for polymorphisms in drug-metabolizing enzymes such as cytochrome P450 isoenzymes that affect drugs like Codeine, Clopidogrel, tamoxifen, and warfarin; the genes KRAS and EGFR genes, BRCA genes, and the HER2 gene that signal susceptibility to treatment with Herceptin (Chin et al. 2011; Tursz et al. 2011; Wistuba et al. 2011).

medicine, individual persons and collective populations are enacted and co-constituted as molecular and statistical entities

Thus, personalized medicine's rhetorical focus on the individual obscures the important, albeit less visible, role that populations continue to play in modern biomedicine. Though personalized medicine places discursive value on individual health, patient responsibility, and personal DNA sequences, it relies on the practices of storing information within biobanks (Mitchell et al. 2010) and of calculating statistics about populations (Holmberg et al. 2012). To this end, personalized medicine typifies Foucault's notion of "biopower" (Raman et al. 2010), as medical institutions intervene into populations through practices of "governmentality," and also into individuals through "technologies of the self" (see Chapter 1). Within personalized medicine, visions of populations strongly influence the epistemic frameworks and research practices used to produce knowledge about individuals. Contemporary discourses on personalized medicine, though they are explicitly focused on individuals, contain "a hybrid of molecular and population categories" (Raman et al. 2010:1722).

Consequently, this chapter explores how individuals and populations are enacted and co-produced in the metabonomics practices that make up personalized medicine within the BMM. While efforts to develop personalized medicine leverage a focus on individual health as something unique and new, I argue that personalized medicine is not solely about the individual. In the first section I argue that within personalized medicine, individuals and populations are co-produced: practices aimed at investigating individuals and populations draw on and require one another to become meaningful. In the second section, I argue that personalized medicine belies an increased reliance on statistical practices to make judgements about health and disease, but under the rhetoric of personalized care. In the practice of personalized medicine, this reliance on statistics makes it difficult for medical practitioners to tailor medical interventions to individual patients.

As the penultimate chapter of this dissertation, my observations here build upon my arguments about the increasing and unstable use of a statistical "normal" to assess health and disease (see Chapter 4), as well as my arguments about the increasing role and value on biochemical and multivariate statistical information to guide approaches to health and disease (Chapter 5). This chapter differs slightly from the preceding chapters, however, in that a significant portion of the data upon which I draw is based on published accounts rather than

ethnographic encounters with metabonomics data. This is because much of the analysis in this chapter is based on developments in the BMM that were still emergent, and were therefore only partially formed, during my fieldwork. Thus, this chapter places the daily metabonomics practices that have formed the focus of this dissertation into broader perspective, showing how they participate in a larger framework of translational and personalized medicine.

Enacting Personalized Medicine in the "Patient Journey"

Several months into my fieldwork, I sat in on a small weekly meeting of researchers involved in the "BRC Grants," a series of translational research projects colloquially named such because of their funding by the Imperial National Institute for Health Research Biomedical Research Centre (NIHR-BRC). At this particular meeting, Jeremy Nicholson made a rare appearance to announce broad changes in the research agenda of the BMM. Gathering researchers around a large sheet of paper, he unveiled grand plans for the future of metabonomics research at Imperial College London. As he pointed to a complex diagram that showed patients and samples flowing through various departments and stages of hospital treatment, he explained how several interlocking projects, funding streams, and initiatives were being directed toward "phenotyping the patient journey" (see Figure 32). This would entail the use of metabonomics technologies to longitudinally model and determine biomarkers for patients' health status before, during, and after their treatment in the hospital. As Jeremy Nicholson described:

> Everybody who is coming into a hospital is undertaking a journey…And what happens is when you go into the hospital, there's all sorts of workup procedures on you trying to figure out what's wrong with you. They do an intervention, and you get better or you don't get better. So what we're trying to do is deploy all the technologies at different stages of the journey to enhance the biomarker information and diagnostics.

The patient journey, therefore, encapsulated the metabolic changes happening within human bodies over the course of a hospital intervention, as patients were given medicines and technological interventions. To this end, drawing on the recent award of £7 million pounds over five years from the NIHR-BRC[71], the BMM intended to develop a series of metabonomics tools—the MAS-NMR analysis of solid tissue mentioned in Chapter 4, as well as the MALDI-MSI imaging and "intelligent knife" project mentioned in Chapter 5—to look

[71] This was part of a larger £113 million pound grant awarded to the Imperial College BRC for work on four key themes. These included biobanking, stratified medicine, genomics and genetics, and imaging (Imperial College London 2012d), of which Jeremy was the leader of the stratified medicine theme.

at the patient journey in cancer and surgery. Referencing the recent publication of a framework for phenotyping the patient journey in the prestigious medical journal *The Lancet*, Jeremy Nicholson proclaimed that this work would provide a revolutionary "new paradigm for hospital medicine" (Kinross et al. 2011).

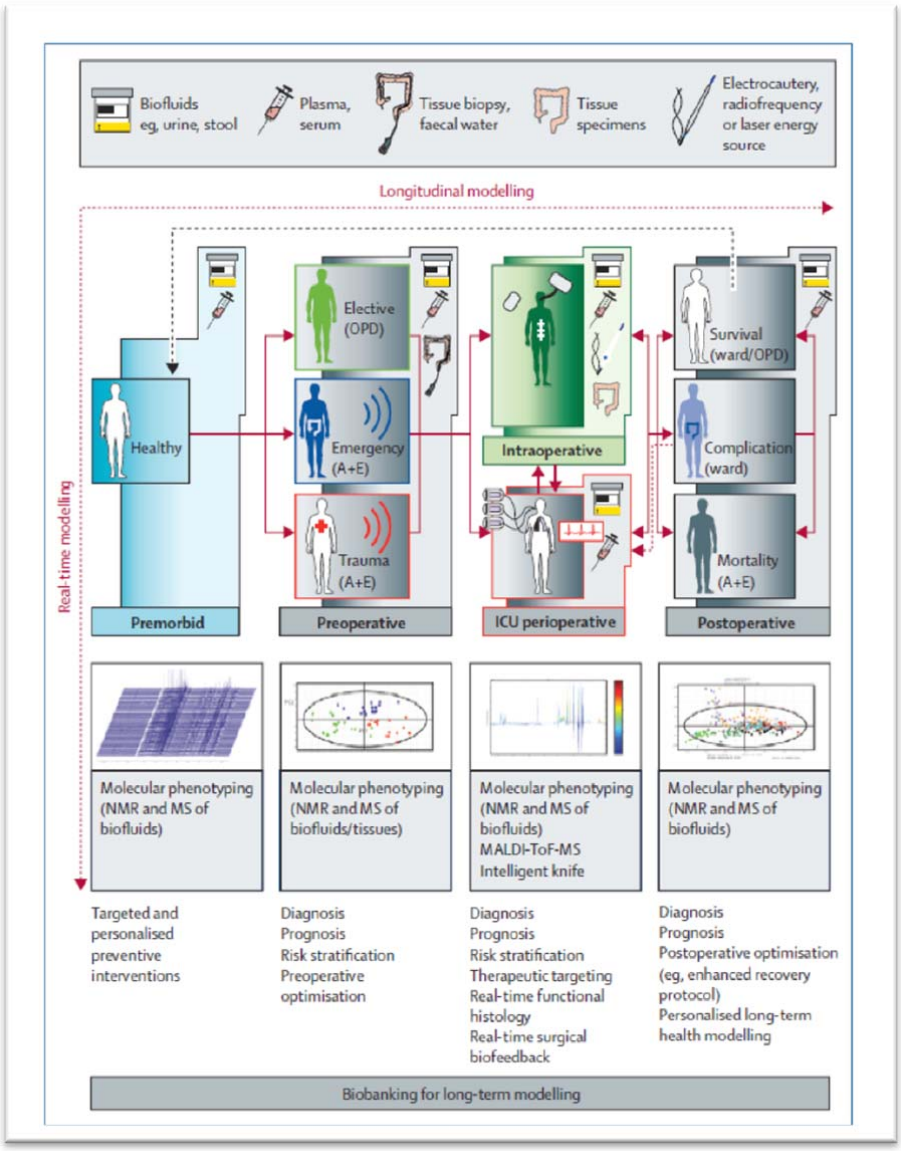


Figure 32: Phenotyping the patient journey.
A diagram depicting the BMM's vision of the "patient journey" within the hospital environment. The patient journey is made up of a combination of longitudinal (x axis) and real-time (y axis) modelling practices, characterized by analyses of biological information over time and during clinical procedures, respectively (Kinross et al. 2011).

Reading between the lines, Jeremy Nicholson's vision of phenotyping the patient journey[72] was essentially a vision of "stratified medicine," a word used synonymously with personalized medicine to describe the study of patient sub-populations rather than specific individuals (Trusheim et al. 2007; Baker 2011). Stratified medicine was considered by researchers in the BMM to be the practical implementation of personalized medicine. It used the same principles—molecular technologies to study individual responses to disease and drugs—but with the recognition that biomedical approaches had to operate within the sociocultural and economic confines of healthcare systems. Personalized medicine, in the sense of developing individual drugs for each patient, would be unrealistic for drug companies, as the cost of researching and developing a single drug was billions of dollars and more than ten years. Thus, as Jeremy Nicholson commented, one of the overarching goals for the BMM's work on personalized medicine was to provide a cost-benefit analysis in terms of measurable improvement in mortality and morbidity. Any research the BMM did should work towards shortening the patient journey and improving overall outcomes. Jeremy Nicholson described in an interview:

> Personalized medicine is really simple. It's getting the right treatment for the right patient based on knowledge of their biology. Alright, now convert that into reality [*laughs*]. Stratified medicine is producing a model based on the biology of the patient that gives you the best guess as the best likely outcome for a therapy. So it doesn't take you or me as individuals…but it puts us within classes of likely behaviour.

For researchers in the BMM, personalized medicine was not only a pragmatic way of improving health outcomes, but also a probabilistic way of measuring and predicting health. It involved statistical calculations and predictions about future disease, and inherently linked to probabilistic estimates of disease "risk" in individuals to broader populations (Mitchell et al. 2010).

Personalized medicine, therefore, was an attempt to make individuals and populations knowable in new and statistical ways. As one clinician-researcher focusing on carotid arterial disease commented, personalized medicine was a way of identifying those patients who did not display outward symptoms, but were still likely to suffer from disease. It was a way of

[72] The surgeon William—the head clinician involved in the BRC Projects—recounted that he had co-developed the concept of "phenotyping the patient journey" with Jeremy Nicholson. Professor Nicholson had spoken to William about a desire to do longitudinal modelling of patients as they moved through a hospital. Consequently, William had used graphic design skills that he had developed during his clinical training to put pen to paper for "a very obvious clinical journey. This is kind of what you do every day [as a surgeon], this is what you see every single day."

focusing not on "the ones with warning signs" who would receive surgery, but on those "that are [going to] have a stroke…[you need to] risk stratify those people somehow." Thus, personalized medicine provided an indication of the most likely outcome and course of treatment for a group of individuals. As Jeremy Nicholson asserted:

> Personalized medicine [is]…here's a new pill, we want to know is this good for you, bad for you, or do nothing for you. If you just knew that information, three levels, three strata, that would be enough. And so you have probabilistic modelling of it, and that's the difference. And for me, it's the essential difference because I think it makes it practical.

Within this vision of personalized medicine, individuals and populations were conceptualised as calculable statistical and biochemical entities. Using the same multivariate statistical practices that were used to make sense of and perform metabolism (see Chapter 3), metabonomics research enacted individuals and populations as the sum total of measured metabolic variables, which were generated from a combination of endogenous metabolism, bacterial metabolism, and environment factors. Individuals and populations thus emerged as multivariate statistical entities, whose relationships to the various metabolites implicated in diet, medications, and disease were expressed as proportions of variance[73]. This stood in contrast to genomic studies, which enacted individuals and populations as statistical frequencies of genetic elements, and focused on the role of the genetic code in the development of disease.

In making individuals and populations calculable and knowable, such a vision of personalized medicine typified notions of "risk" as a technological achievement of modern society (Beck 1992). Within the field of metabonomics, personalized medicine was not based on outward signs, symptoms, or illness experiences, but rather on internal and molecular measurements. It involved the surveillance of individuals and populations from above and below, and also the problematization of the "normal" within biomedical contexts (Armstrong 1995). Thus, as metabonomics researchers sought to develop technologies for personalized medicine, they also gave rise to potentially new ways of intervening into bodies, based on the future probability or chance—rather than the physical manifestation—of disease.

---

[73] In multivariate statistical models like principal components analysis (PCA), the differences within data are visualized according to "principal components." Principal components are factors that represent the maximum variance (multivariate difference) within the data, but do not necessarily correspond to real-life variables such as diet and age. As such, principle components are said to be responsible for certain percentages of the variance in the data. In this way, the multivariate statistical approaches in metabonomics capture "probability" as a combination of factors, rather than a function of a single variable, such as diet or age.

Individual and Population Phenotyping

To examine the patient journey with metabonomics technologies and approaches, the BMM made plans to establish two complementary facilities, which would carry out a combination of real-time diagnostic and epidemiological longitudinal research. These facilities would advance individual-level clinical studies and interventions, and also population-level epidemiological studies of biomarkers and disease risk factors. Together, these facilities represented opposite but inter-related practices and goals in the study of personalized medicine: the implementation of "pharmacometabonomics"—the prediction of drug toxicity and response in individuals—and the implementation of "metabolome wide association studies (MWAS)"—the determination of disease biomarkers and phenotypes in populations. To this end, such developments beg important questions about how the rhetoric of "personalized medicine" relates to its practice, and about how the "person" and "population" are implicated in such medical practices.

At the time of my fieldwork, the BMM was establishing the "Clinical Phenome Centre" (see Chapters 2 and 4) in the St. Mary's Hospital, which was located a 30-minute walk north of the BMM's laboratory in South Kensington (Imperial NIHR Biomedical Research Centre 2011). Equipped with six MS and three NMR machines, the Centre aimed to collect and analyse samples from patients undergoing pharmacological and surgical treatment for cancers, sepsis, and other serious conditions such as gastric bypass and pneumonia. This would involve, as several researchers engaged in setting up the Centre explained to me, extensive coordination and collaboration between laboratory researchers, clinicians, and medical specialists (see Chapter 5). In the end, researchers hoped that the Centre would "help doctors make a more informed diagnosis, choose the best treatment based on the individual characteristics of the patient, and monitor their progress more precisely" (Wong 14 November 2012).

The activities of the Clinical Phenome Centre drew on the concept of pharmacometabonomics, in which measurements of initial metabolic phenotypes were used to mathematically model responses to subsequent drug therapies and interventions (Nicholson et al. 2012b). The concept of pharmacometabonomics emerged in the mid-2000s[74] after

[74] According to John Lindon, the concept of pharmacometabonomics began to take shape in the early 1990s during the work of then post-doctoral student—and currently head of the BMM—Elaine Holmes, whose work examined the trajectory of metabolic responses to renal toxicity (Holmes et al. 1992).

researchers in the BMM conducted a study examining the variability of responses in rats to the chemical compound galactosamine hydrochloride, which causes toxic liver damage. They found that by measuring the metabolic status of the rats before administration with the toxic chemical, they were able to predict the extent of liver damage in individual rats (Clayton et al. 2006) (see Figure 33). This study, which was conducted in collaboration with the company Pfizer during the era of the BMM's collaborative work on the COMET project (see Chapter 2), was later applied to humans to predict responses to the drug acetaminophen, a common over-the-counter drug that can have toxic effects on the liver at high doses (Clayton et al. 2009). Together, these two studies provided a "proof of principle" that metabonomics technologies could be translated into clinical settings to inform medical treatment.

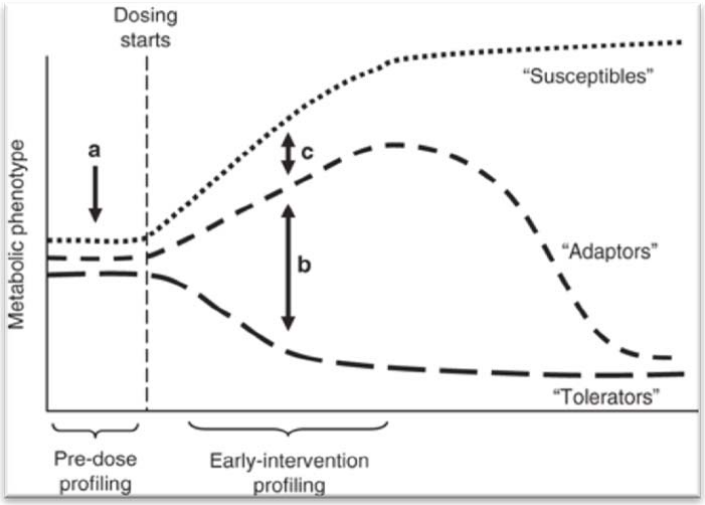


Figure 33: Pharmacometabonomics and individual phenotyping.
Diagram depicting the concept of "pharmacometabonomics": how an organism's starting metabolic phenotype (left) influences the response to and outcome of medical treatments (right) (O'Connell et al. 2010).

Researchers in the BMM emphasized that pharmacometabonomics approaches to the patient journey would enable the prediction of responses to surgical procedures and drug interventions. By generating information about an individual's starting metabolic phenotype—the combined outcome of genes, diet, drugs, stress, and pathogens—researchers would be able to understand future health and disease. As one researcher described to me:

> So can you model something at the beginning of the journey, and give a prediction of the outcome of a particular therapy or use that information to choose that optimal therapy. So in cancer, will that chemotherapy work, will the patient be a good responder or a bad responder based on that pre-intervention profile? So that's a pharmacometabonomic principle that we're trying to apply.

Using the metabolic information contained within individual phenotypes, researchers would generate statistical models to predict an individual's response to medical treatment or intervention. These would be based on the measurement and quantification of each individual's "normal" metabolism, which would serve as a baseline against which to measure disease. With this approach, researchers hoped to use metabonomics technologies to correlate an individual's phenotype with an individualized probability of having a certain outcome to treatment or disease. They hoped, in other words, to be able to detect the presence of conditions like diabetes before patients displayed harmful symptoms.

As an example of this type of approach, a researcher in the BMM was overseeing a study on personalized nutrition, in which a diverse group of people from the United Kingdom were placed on five diets, characterized by unsaturated fat, low fat, high fat, low glycemic index, high glycemic index. The study aimed to compare the effect of the diet on each participant by measuring their biology before and after the dietary intervention. Because each individual participant had entered the study with a different set of bodily conditions and dietary habits, each individual had a different starting metabolism. The goal of the study, therefore, was not to determine if participants were healthy or sick, but rather to determine how each individual responded differently to the controlled diets of the study. Therefore, the study sought to determine the range of normal metabolisms that existed across different age groups, genders, and ethnicities in the United Kingdom, by using "every person [as] their own control."

Overall, pharmacometabonomic practices entail individual biological measurements that are used to produce individualized approaches to health and disease. Such practices place value on the determination of *health as an ideal or optimal state of being*. By conceptualizing health and disease in relation to an individual's unique metabolism, pharmacometabonomics attempts to determine—with the aid of statistics—the best possible outcome for each individual patient. In doing so, pharmacometabonomic practices enact health as something prescriptive, as something to which individuals should strive. Consequently, they enact personalized medicine as that which can provide an optimal state of being for each individual patient, according to individual metabolic measurements.

Though the Imperial Clinical Phenome Centre and the patient journey were explicitly oriented at the provision of personalized diagnosis and treatment, establishing the point at which individuals moved from a normal to an abnormal phenotype was not a trivial thing (see Chapter 4). Monitoring metabolic phenotypes required more than the measurement of a single patient as he or she moved through the hospital. It also required the comparison of groups of similar patients—who would undoubtedly display similar metabolic phenotypes—to establish a range of normal and abnormal metabolism. As one researcher said:

> So if you have a good idea for what is normal, and you've got somebody who's abnormal, then you can do an intervention on them to see whether or not whatever parameters you're measuring get them close to normal…On the other hand, understanding what is normal is a non-trivial question. Because what is normal for you as an American woman and me as a British middle aged man, right, is actually quite different in terms of a whole series of parameters. So, one of the things that we've been doing in the background with the epidemiologists is trying to establish [the normal] in populations. Chinese, Japanese, English, Americans, they're all very different!

To this end, following the 2012 London Olympics, the BMM, in a joint venture with King's College London, commissioned the use of the Olympic doping testing facilities to create the "MRC-NIHR Phenome Centre" (see Chapter 2). Funded by a combination of private and public entities, the Centre was scheduled to open in 2013 at Hammersmith Hospital, another of Imperial College's medical campuses.

The overall goal of the MRC-NIHR Phenome Centre was to carry out population phenotyping of human beings via the analysis of large numbers of biological samples[75]. The Centre's activities were directed at longitudinal patient modelling based on the concept of metabolome-wide associate studies (MWAS), the use of metabonomics techniques to identify metabolic variables that were correlated with epidemiological risk factors. The concept of MWAS emerged in the late 2000s through the BMM's work on the International Collaborative Study of Macronutrients, Micronutrients and Blood Pressure (INTERMAP). INTERMAP examined the urine of 4630 people from four different populations in the United States, United Kingdom, Japan, and China, in order to find correlations between lifestyle factors and blood pressure (Holmes et al. 2008a; Imperial College London 2012b) (see Figure 34). From the study, metabonomics researchers not only demonstrated that each population was characterized by a distinct metabolic phenotype, but also that the metabolites which

[75] The Centre aimed to foster collaboration among researchers conducting population-level research at four other institutions—Oxford University, Cambridge University, University College London, and King's College London—all of which are NIHR-BRCs (National Institute for Health Research 2012). Imperial College is the second highest recipient of NIHR-BRC funding in the United Kingdom (University College London 2013).

made each phenotype unique could be linked to differences in diet, gut microbial population, and drug intake. Thus, in a field traditionally dominated by genetic profiling and genome wide association studies (GWAS) (Nicholson 2006:1), the INTERMAP study provided a "proof of principle" that metabonomics technologies could be used to examine metabolic phenotypes across large populations, and consequently to correlate such metabolic phenotypes to environmental exposures and disease factors.

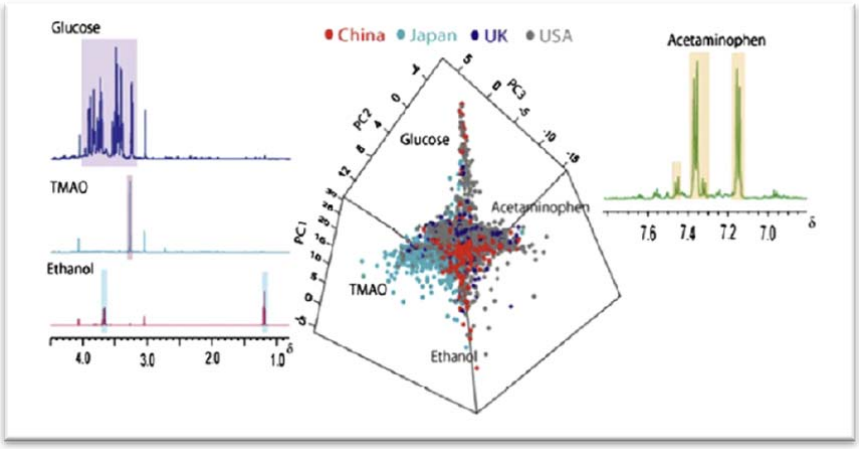


Figure 34: Molecular epidemiology and population phenotyping.
Diagram depicting the concept of "molecular epidemiology" as it was carried out in the INTERMAP Project (Bictash et al. 2010). The middle image shows a multivariate diagram of the metabolic characteristics of the four populations—people from China, Japan, the UK, the USA—while the right and left images depict specific chemicals that are responsible for the differences in these populations.

Researchers in the BMM emphasized that MWAS approaches to the patient journey would reveal how both populations and diseases could be characterized as molecularly diverse entities. Patients exposed to different environmental or lifestyle factors were likely to respond differently to diseases and interventions, and to have molecularly diverse phenotypes. For example, echoing Margaret Lock's (Lock 1995) work on aging, women from the United States were likely to experience menopause and hormone replacement therapy (HRT) differently from women from Japan. Moreover, patients diagnosed according to standardized symptoms and physiological clinical signs often experienced disease in fundamentally different ways, requiring different hospitalization lengths, responding differently to drugs, and experiencing different side-effects from treatment. Thus, researchers expected metabonomics technologies to provide information about the range of phenotypes displayed by groups of people, which could be measured and quantified with statistics.

In contrast to pharmacometabonomic practices, molecular epidemiology practices entail measurements of the biology of populations, which are used to tailor medical interventions to different groups of people. However, molecular epidemiology practices also place value on the determination of *health as an average or typical state of being*. By conceptualizing health and disease in relation to the metabolism of groups of individuals, molecular epidemiology attempts to determine—with the aid of statistics—the range of possible outcomes for different populations. As one researcher said: "Because you are getting information from a huge amount of people, it's more [correct], it's more true." Thus, molecular epidemiology practices enact health as something descriptive, as a guideline for the "typical" or "standard" health of individuals. In doing so, they enact personalized medicine as that which can provide the best overall care to a population of patients, according to population-wide metabolic measurements.

Co-producing Individuals and Populations

Taken together, the Clinical Phenome Centre and the MRC-NIHR Phenome Centre represent efforts to put "personalized" medicine into practice, but in distinctly different ways. Practices of individual phenotyping enact personalized medicine as something that is concerned with optimal states of individual health, while practices of population phenotyping enact personalized medicine as something that is concerned with the average or range of possible health outcomes for groups of individuals. Consequently, these two different approaches to personalized medicine engender not only different practices for investigating health and disease, but also different concepts of what health and disease entail.

The differences between individual and population phenotyping, however, belie the ways in which these two practices are inter-related. Both individual and population phenotyping are united in their use of multivariate statistical practices to investigate processes of health and disease. They are also united in the value they place on the collection and interpretation of metabolic information to glean understandings of disease diagnosis and treatment. Consequently, these two approaches to personalized medicine, though they entail different practices and concepts of health, are entangled with one another. They do not represent two different approaches to personalized medicine, but rather *two different poles of the overall practice of personalized medicine, one focused on individuals and the other focused on populations*. As metabonomics researchers themselves write: "[Individualized] healthcare and molecular epidemiology are thus effectively two sides of the same 'systems

biology coin'; the essential differences are with respect to the type of medical end points or outcomes that are to be modeled" (Nicholson 2006).

To this end, in their discussions of the practice of personalized medicine, researchers in the BMM frequently emphasized that populations and individuals could not be considered independently from one another. For example, at the end of a seminar, a group of researchers discussed the logistical challenges of using metabonomics technologies to provide personalized medicine. One researcher who worked with human studies questioned how it would be possible to provide medicine to individuals in a truly personalized way. Logistically, how would researchers be able to conduct validated statistical analyses on a single person? What would serve as a proper control, and how would researchers generate robust statistical correlations with a sample size of one? In response, the seminar speaker—Ian Wilson, the scientist from AstraZeneca whose work I discuss in Chapter 4—responded:

> I firmly believe…that based on some of the work that's done here is the only way to understand individuals is to understand populations. So I think that you need to do these larger epidemiological studies so that you can say, well a normal human being spans this amount of variability. And also to anchor some of that back into the current clinical chemistry tests.

According to the Ian Wilson, individuals could not be understood without populations. Population information was necessary to provide the context for individual information: it allowed scientists to understand how individuals fit into the larger group of which they were a part (see Figure 35). Similarly, Jeremy Nicholson emphasized in an interview that populations could not be understood without individuals. He said: "It's only when you've got a lot of similar patient journeys, where the biology and the interventions are very similar…that you can try and produce a more generalized model." Individual information was necessary to provide in-depth case studies of the metabolic effects of health and disease across different environments and interventions.

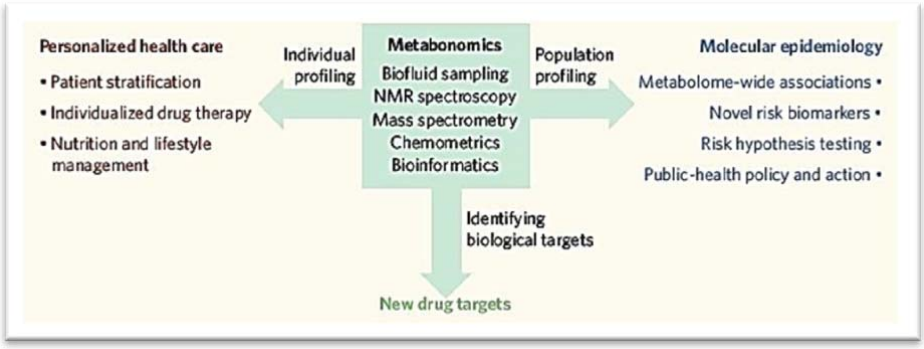


Figure 35: Co-production of individuals and populations.
Diagram depicting the entanglement of individuals and populations in "personalized health care" and "molecular epidemiology" (Nicholson et al. 2008). Note how metabonomics is seen as carrying out both "individual profiling" and "population profiling."

Similarly, Jacob—the researcher whose work I discuss in Chapter 5—emphasized that populations and individual studies "were complimentary," with populations acting as "as a kind of calibration" to establish how common or widespread the biology of an individual person might be. But for Jacob, population data was also a way to forecast the likelihood that a disease might occur. Population data allowed researchers to calculate the probability that an individual might become sick or might respond poorly to a medical intervention. Because metabonomics research was involved in predicting future disease, experiments aimed at the development of personalized medicine involved, as another surgeon commented "an element of chance" and "of trying to guess what's going to happen."

In many ways, discussions of personalized medicine within the BMM speak to Tiago Moreira's (2011) historical discussion of the Baltimore Longitudinal Study of Aging. Moreira describes how the study entailed contrasting but inter-related approaches to the study of populations, which involved serial observations of individuals and the calculation of averages of populations. According to Moreira, the decision to pursue either approach entailed politically motivated questions about the overarching aims and uses of the study. Examining "population laboratories" or "laboratory populations" entailed, respectively, investigations of either the parameters of the normal aging process, or the specific instances of individual variation from the normal. Thus, metabonomics researchers' concerns with examining both individuals and populations speak to a long-standing tension within the biomedical research community about the use of biological information for the study of health disease.

By combining individual and population data, metabonomics researchers carry out what philosopher Ian Hacking (1990) refers to as the "taming of chance," in which descriptive observations are transformed into prescriptive laws and statements about causality. Consequently, the taming of chance in personalized medicine entails a shifting and increasing role for informational practices, with which observations about individuals are transformed into future statements about individuals. In this way, informational practices in personalized medicine are used to manage contingency by calculating and making the future knowable. Such informational practices are implicated not only in describing the boundaries between the normal and abnormal, but also in making prescriptive statements about the point at which individuals and populations will transition into diseased states of being. Therefore, the use of statistical and biochemical technologies to make statements about future health and disease affords new realms of power and control over individuals and populations.

Ultimately, "personalized medicine" within the BMM entails a discursive focus on the individual, which belies the ways in which metabonomics conscripts individuals and populations into the study of disease. This is in part a reflection of the ways in which metabonomics research is embedded within popular efforts and funding calls, as I have discussed in Chapter 2, to develop translational technologies for personalized medicine. But as researchers seek to tailor medical practice to individuals, they invariably draw on population-derived ranges and averages. On the other hand, as researchers seek to learn about the population-wide distribution of disease, they invariably examine case studies of smaller groups of individuals and phenotypes. Individuals and populations do not exist separately or on their own. Rather, each concept gives meaning to the other. Thus, the vision of personalized medicine promoted within the BMM provides an ideal case study for examining how the rhetoric and practice of personalized medicine are not always aligned.

Personalized Medicine in Practice

While the following section examined how individuals and populations are co-produced in personalized medicine, this section examines how personalized medicine is carried out in practice by metabonomics researchers in the BMM. To do so, this section details attempts to improve upon the statistical algorithms that are used to assess the severity, prognosis, and treatment for liver disease through the development of multivariate statistical models of clinical data. In doing so, it draws upon my discussion of the application of multivariate statistics to clinical data in Chapter 5, in order to emphasize the ways in which

metabonomics transforms the form, meaning, and value of information in the diagnosis of liver disease. Consequently, this section argues that attempts to develop "personalized" multivariate statistical models struggle to improve the medical care given to individual patients because of a reliance on population statistics.

Towards the end of my fieldwork, I closely followed the research of Andrew, a clinician who specialized in liver disease and transplantation in the UK health system, but who was also doing a 6-month research project in the BMM. Prior to his work as a clinician, Andrew had completed his PhD in quantum physics, but after working briefly as a phlebotomist, had decided to pursue a post-graduate degree in medicine in London. Combining an interest in quantum mathematics with a decade of medical training, Andrew had a deep interest in the statistical methods that were used in the clinic to assess liver disease. When I asked Andrew if he struggled—like many of the other clinician-researchers I interviewed—to understand the multivariate statistics common to metabonomics practices, he replied that after working in a computationally-intense field like quantum physics, he was comfortable with the complexities of data analysis.

Diagnosing and treating liver failure, according to Andrew, was a challenging field of work. Liver disease is the fifth largest cause of death in the UK, and occurs when the liver is no longer able to carry out its key functions in metabolism, immunity, and detoxification. Liver failure can be acute—due to drug overdose or trauma—or chronic—due to longer-term conditions like alcoholism and hepatitis C. Though people can sometimes recover or live with liver failure, patients are often faced with the need for liver transplantation due to declines in their quality of life. The allocation of livers, however, is a tricky thing, as they are in high demand but have a low availability. To this end, clinical tools have been developed to aid doctors in their decision making about how to allocate livers to individuals within populations.

From a biomedical perspective, Andrew described, it was difficult to diagnose the severity of liver failure—and consequently the need for a liver transplant—because the condition was marked by a diffuse and diverse set of symptoms. As many of the organs to which the liver was linked failed simultaneously, patients displayed common symptoms such as jaundice, encephalopathy, kidney failure, and coagulopathy, but in varying orders and severities. Furthermore, from an ethical and public health perspective, it was difficult to make decisions about liver allocation because of high demand and short supply. Dealing

with livers as a scarce resource, clinicians had to strike a balance between doing what was right for the individual patient, and also allocating organs for the population. Giving a liver to someone who was not judged to be in immediate need meant denying a liver—and possibly causing death—for another person. Andrew said: "They might die, or you [might] misclassify them and give them an organ [that was meant for] someone else. And you will never know if you were right or wrong."

To deal with these challenges, clinicians used several statistical algorithms, which differed according to the type of liver failure, to determine which patients would qualify for a liver transplant. Like individual phenotyping practices, this was an attempt to "prognosticate" patients: to establish a point at which liver failure was severe enough to warrant a transplant, but not so severe that the patient would die before a transplant could be done. These statistical models were based on certain "criteria," biological parameters that were determined by blood measurements and clinical examination. In the end, these statistical models produced a numerical score that reflected the severity of liver disease. The higher the score, the higher the chance of death, and consequently the more urgent the need for a transplant.

Andrew's work focused on liver transplantation for a condition known as "acute-on-chronic" liver failure, in which patients with chronic liver disease developed rapid-onset, acute liver failure. His research was an attempt to develop new biochemical and statistical techniques to evaluate and diagnose acute-on-chronic liver failure. This was largely in an effort to improve on the current clinical tool, an algorithm called Model of End-Stage Liver Disease (MELD)[76]. Using linear regression on a combination of clinical measures of serum bilirubin, creatinine, and prothrombin time, MELD calculated a patient's probability of dying over a certain period of time without a liver transplant (see Figure 36). As Andrew explained, "Basically everyone just does it on a website now, you just type in your results and you get a score. And the higher the score, with a max of forty, the more likely a patient is to die." Moreover, the diagnosis of acute-on-chronic liver failure was not determined solely according to MELD. In order to qualify for a transplant, patients were also required to

[76] MELD—like other clinical algorithms—has a distinct history of use in the UK health system. MELD was implemented in 2002 to replace the older system for assessing need for transplantation, which entailed the combination of the Child-Turcotte-Pugh (CTP) score and overall waiting time for a liver transplant. The implementation of MELD, therefore, represented a change to the "sickest first" policy (Cholongitas et al. 2010), in which people were ranked on the list for a liver transplant based on the severity of their disease, and irrespective of the waiting time of other patients (Wiesner et al. 2003).

display signs that another organ—such as the kidney or the brain—was failing in addition to the liver.

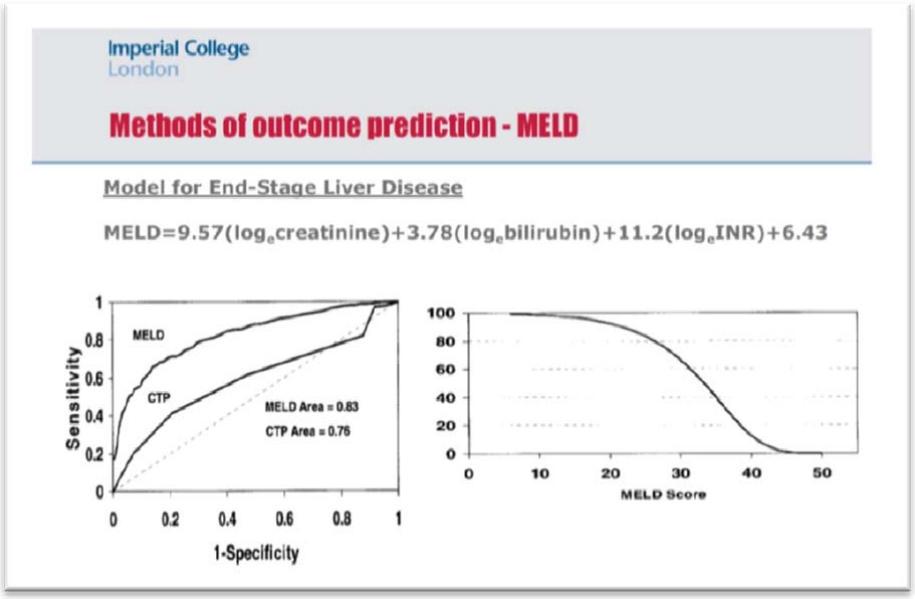


Figure 36: Predicting liver failure with MELD.
Powerpoint slide from Andrew's presentation showing the formula for the MELD algorithm (top), as well as its ROC Curve (left) and association between score and chance of dying (right).

As I spoke to Andrew about the use of MELD to assess acute-on-chronic liver failure, he expressed concern that the clinical algorithm prevented clearly sick but "borderline" patients—those who were on the threshold of a statistical cut-off for treatment—from receiving liver transplantation. The reliance on MELD was problematic because it created inflexible boundaries between those who received treatment and those who did not. While a MELD score of 18 was associated with a 40% chance of dying, a score of 22 was associated with an 80% chance, which was a severe estimate for death. Because their condition could change rapidly and for the worse, borderline patients with a score of 18 or 19 could progress to full-blown liver failure before their MELD score was next measured by a clinician.

Thus, Andrew's comments expressed concern about clinicians' reliance on statistical models for determining liver failure, as he said "none of [the models] are perfect, they're a summary under a…curve." His comments highlight how statistical models are institutionalized technologies, which might appear to be reproducible and objective, but are associated with distinct histories, values, and practices. As statistical models like MELD

create neutral numerical thresholds for liver transplantation, they generate biomedical boundaries that have moral implications for the kinds of patients who are deemed to have life-threatening diseases, and consequently to qualify for life-saving treatments. Ultimately, as Andrew questioned the utility of such numerical boundaries to evaluate the severity and prognosis of disease, he also raised questions about the role and compatibility of human judgement—as I have discussed in Chapter 5—with statistical technologies.

To illustrate how statistical scoring systems proved problematic for borderline cases of disease, Andrew gave an example of one of his patients, a woman who suffered from liver failure due to a drug overdose. Andrew and his clinical team sensed that the woman's prognosis was poor, but they were unable to list her for a liver transplant because she was missing symptoms of encephalopathy, a state of confusion that indicates that the brain is beginning to fail. Convinced that the woman was in need of a liver transplant, Andrew and his team kept the woman in the hospital, expecting her condition to decline. In doing so, they relied on what Andrew referred to as a "gut feeling," a combination of clinical judgment and experiences, which included sensing the woman's pulse, color of her skin, quality of her respiration, and general change in state over time. After a few weeks, as predicted, the woman went into a coma as a result of encephalopathy due to liver failure. Because she had finally displayed the final criterion of encephalopathy, Andrew and his team were able to list her for a transplant, and in the end, she survived. Andrew described the experience:

> What we all known as doctors and intensivists, is basically you go up to a patient and they're unconscious. They're yellow…they're cold, and it's just bits of information that fit into that pattern that you don't like. But even if you go incrementally up to, oh, their pulse is 90 three minutes ago and now its 110, it's not a big change. But it still adds to the pattern of, they're doing badly.

According to Andrew, in such borderline cases of disease, clinicians often knew that patients were in need of a transplant before the statistical models indicated so. They carried out, as another liver specialist described, "their own pattern recognition," in which they applied statistical models repeatedly to patients until they displayed a result that would qualify them for a liver transplant. Andrew commented:

> You [are] clearly having somewhat of a problem that doctors with 30 years of experience in liver intensive care will say that a patient…is not going to survive. And they'll go back to them every six hours and examine them again to see if they get another criterion, so they can list them…You always come back to the patient and decide, are they getting better or worse?

Overall, Andrew's case highlights how the treatment of liver failure is an inherently personalized practice. As Andrew described how the clinical team had been able to sense,

through interacting with the woman repeatedly, that her liver would eventually fail, he emphasized that the management of disease was interpretive and individualized. Clinicians did much more than apply diagnostic technologies and note numerical MELD scores. Rather, they drew on their training and judgement to respond to liver failure as a dynamic and individual condition (see Chapter 5). Importantly, clinicians relied on non-statistical information—on information that was not easily quantified with laboratory technologies—to provide adequate medical care and treatment. In doing so, they were attuned to the fact that not all patients displayed the same symptoms, and also to the fact that individual patients changed dynamically over time.

Ultimately, Andrew used this moment of ambiguity, as I discuss in the section that follows, to suggest that more personalized approaches to the diagnosis of liver disease were necessary, and could be achieved by using multivariate approaches to biology. At stake in this assertion, however, was the extent to which certain aspects of health and disease—or of individuals and populations—could be quantified and examined in statistical models. Consequently, Andrew's story of the young woman typifies an alternative form of "personalized medicine," in which clinicians draw on trained judgement and non-statistical information to provide unique and tailored approaches to liver failure patients. It echoes Ian Whitmarsh's (2011) observations of personalized medicine in a genetics clinic in Barbados, in which care is "personalized" not because of genetic technologies, but because of the affective care patients receive through their interactions with medical professionals. Thus, as patients are configured into a series of measurable and seemingly objective variables, such variables do not necessarily capture those elements of health that enable clinicians to carry out the diagnosis and treatment of disease.

Diagnosing Individual Patients with Multivariate Statistics

Several weeks after my interview with Andrew, I heard him present one aspect of his research on liver disease at a seminar in St. Mary's Hospital. Standing in the front of a small room filled with a combination of clinical practitioners and metabonomics researchers, Andrew described his work to develop advanced statistical models—using the multivariate statistics that metabonomics researchers applied to their biochemical data—that could provide a more personalized alternative to clinical models like MELD. After providing a brief overview of the status of liver disease and treatment in the UK, Andrew described his

work to carry out "multivariate modelling of the continuous non-spectroscopic data that you capture as part of the normal, daily care in the liver intensive care unit."

Working with a clinical database collected between 2002 and 2007 from liver disease patients in the intensive care unit (ITU) at King's College London, Andrew had applied common multivariate statistical techniques to a variety of clinical data, which included heart rate, respiratory rate, platelet count, and blood pH. While MELD was based on three clinical variables, Andrew's multivariate statistical analysis was based on thirty six variables in total, all of which were derived from bedside or blood tests (see Figure 37). Consequently, Andrew had compared the results from the MELD algorithm, which was based on logistic regression, with the results of his own multivariate statistical analysis, which was based on a technique called orthogonal partial least squares discriminant analysis (OPLS-DA). By incorporating a larger number of biological measurements into his algorithm, Andrew hoped to be able to better tailor the diagnosis and treatment of liver failure to individual patients.

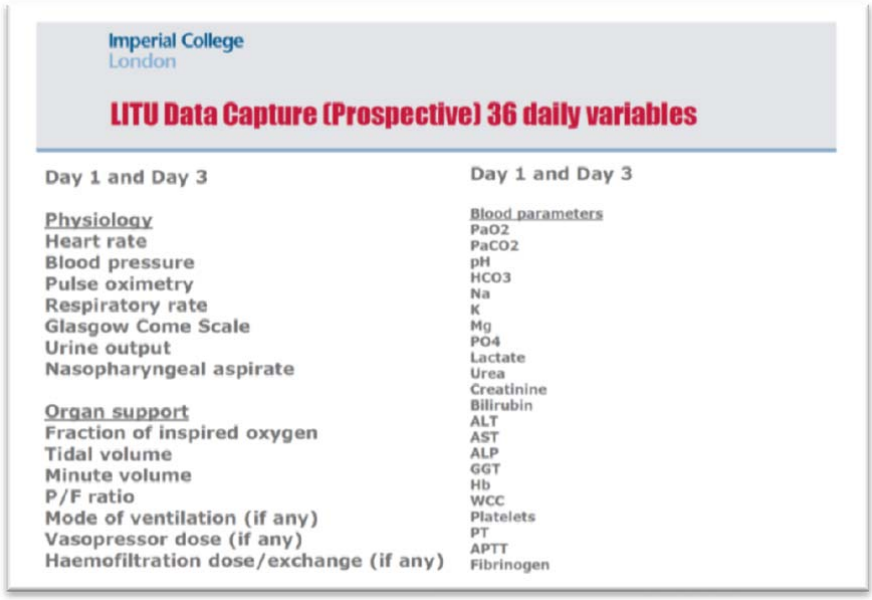
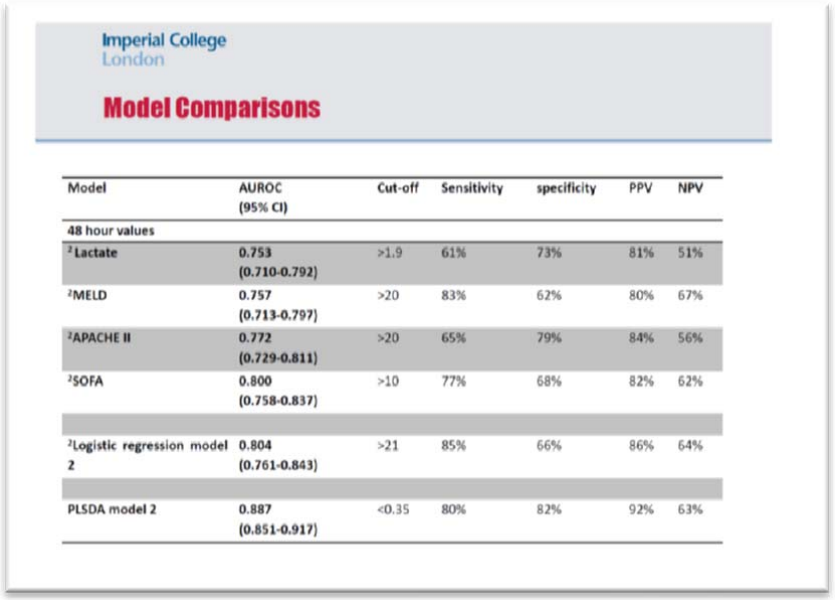


Figure 37: Clinical information collected in the liver ITU.
Powerpoint slide from Andrew's presentation showing the clinical information collected in the liver ITU, which he used for his OPLS-DA algorithm for calculating liver failure.

Presenting the results of the comparison, Andrew asserted that his multivariate approach could better predict the threshold for liver failure in individual patients with acute-on-chronic liver disease. He based his assertion on the relative ability of the two methods to

correctly classify patients as diseased or not diseased, using the widely accepted measures of sensitivity, specificity, and area under a receiver operating curve (AUROC). Sensitivity and specificity are statistical measures of the performance of a classification test: sensitivity measures the "true positive rate" or the ability of a test to correctly identify people who are positive for a disease, while specificity measures the "true negative rate" or the ability of test to correctly identify people who are negative for a disease. Consequently, the AUROC (which has a maximum of 1) combines measures of sensitivity and specificity to indicate the overall accuracy of a classification test, or in other words to indicate how well the test discriminates between people who do and do not have the disease (Tape 2013).

Comparing his OPLS-DA analysis with MELD, Andrew showed that his multivariate approach, which he had applied to data taken 48 hours after patient admission, was better able to classify patients with acute-on-chronic liver disease. While Andrew's method provided a sensitivity of 80%, a specificity of 82%, and an AUROC of 0.887, MELD in comparison provided a sensitivity of 83%, a specificity of 62%, and an AUROC of 0.757. The lower numbers for MELD, according to Andrew, were a quantitative indication that his OPLS-DA model was better able to assess liver failure in acute-on-chronic patients, and provided evidence that his method should be tested in other patient cohorts (see Figure 38).


Imperial College London

Model Comparisons

Model	AUROC (95% CI)	Cut-off	Sensitivity	specificity	PPV	NPV
48 hour values						
[2]Lactate	0.753 (0.710-0.792)	>1.9	61%	73%	81%	51%
[2]MELD	0.757 (0.713-0.797)	>20	83%	62%	80%	67%
[2]APACHE II	0.772 (0.729-0.811)	>20	65%	79%	84%	56%
[2]SOFA	0.800 (0.758-0.837)	>10	77%	68%	82%	62%
[2]Logistic regression model 2	0.804 (0.761-0.843)	>21	85%	66%	86%	64%
PLSDA model 2	0.887 (0.851-0.917)	<0.35	80%	82%	92%	63%

Figure 38: Comparison between univariate and multivariate ways for predicting liver failure.
Powerpoint slide from Andrew's presentation showing a comparison between the sensitivity and specificity of various models, including MELD and his OPLS-DA model.

After Andrew's hour-long seminar, the clinicians and metabonomics researchers in attendance began a heated debate about the meaning of his research. This spoke broadly to debates about the use of models—what they meant, how they were generated, how they might be used—within clinical practice. Though Andrew was hesitant to claim that his research could revolutionize the diagnosis of liver failure, one metabonomics researcher was convinced that Andrew's OPLS-DA model provided a significant improvement. "It's almost 12%, so 12% is considerable," she said. Envisioning the impact Andrew's model would have on a cohort of 100 to 1000 patients, she emphasized: "Even though it's not that many, from a mathematical point of view it's an improvement."

However, a practicing gastroenterologist based in St. Mary's Hospital felt otherwise. He questioned whether the difference in accuracy—76% for MELD and 88% for Andrew's OPLS-DA model—was large enough to make a case for the implementation of the multivariate statistical analysis in clinical practice. The gastroenterologist said:

> It's just that the values didn't look that different, for all of your different approaches that you used…The area under the ROC curves, they were all fairly similar…And I'm a great advocate of multivariate pattern recognition analysis, but I think if you were trying…to tell people that this is the…wonderful thing, then…the values have got to be very different to …what has been used so far.

Such comments were emblematic of the different approaches to the study of disease that existed in laboratory and clinical spheres of practices (see Chapter 5). Coming to Andrew's defense, the surgeon named William—whose work I detail in Chapters 4 and 5—commented that he was encouraged by Andrew's results, precisely because the corroborated the current clinical assessments for liver failure. He said: "Frankly, if your…data said that creatinine, bilirubin, and [prothrombin time] were not that important, I would be more worried…Because you can't just say…hundreds of years of clinical experience are wrong." Andrew's research was not expected to provide a "quantum leap" in the assessment of liver failure, but rather to show small and incremental improvements in accuracy.

Embedded in the debate about Andrew's research was the assertion—as discussed in Chapter 5—that the use of sophisticated molecular technologies and the modelling of complex biological information could result in improved disease diagnosis and treatment. When I asked Andrew about his research several weeks after the seminar, he asserted that the key to overcoming issues with the diagnosis of borderline cases of liver disease was to

develop better ways of measuring peoples' biology. He asserted, in other words, that a technological solution would improve the diagnosis of and care for liver failure. What was absent from the discussion surrounding Andrew's research, however, were questions about the increased value placed on biological information—and on what aspects of individual health such biological information could capture—in the diagnosis and treatment of disease. This was ironic in some ways, given Andrew's insistence in the previous section on the role of human capacities such as "gut feeling" in the diagnosis of disease.

To this end, the gastroenterologist continued to question the value of Andrew's research, saying that he was not convinced that the difference between the two algorithms would be useful for clinicians faced with the challenge of treating individual patients. As Andrew responded:

> So I completely agree if it's a patient in front of you. ROC curves for the one patient is, you know, useless. If you're deciding on what we're going to do as a group of gastroenterologists and intensivists, which is many of the kind of reasons for using [statistical algorithms] and deciding what's the burden of disease and how we're going to assign resources, maybe it could be different.

Within Andrew's reply was the recognition that statistical models for the assessment of disease were based on measurements and assessments of populations rather than individuals. The results of MELD and Andrew's OPLS-DA model were useful for understanding how to allocate livers and manage the burden of disease in populations. In contrast, such statistical models were not as useful for determining the care of individual patients. As another clinician commented: "So there's a lot of tools and there's lots of graphs and tables…[that] give someone an estimate of risk, but they're very crude. And they're based on populations, and those populations are not necessarily representative of the person in front of you." Thus, the difference between 76% and 88%, which was hailed as a significant improvement of Andrew's OPLS-DA model over MELD, did not affect the care of individual patients. It did not, in other words, make it easier for clinicians to make difficult decisions about how to diagnose and treat liver disease on a case-by-case basis. Echoing this, Andrew said to me on a separate occasion that the hardest part of his job as a liver specialist was in using clinical statistical tools like MELD to do the best for the individual patient rather than the general population.

In conclusion, my discussion of Andrew's research on models of liver disease shows how researchers rely on and value metabolic information and statistical practices to create diagnostic measures that are seen as "tailored" to individual patients. Multivariate statistical

models are seen to provide more accurate diagnostic indicators than models based on linear regression and univariate logic. However, during the daily practice of medical care, in which patients display individual cases of and trajectories for disease, statistical models in general can prove unhelpful for the management of individual patients. This is because, I argue, the efficacy and value of statistical models like MELD and OPLS-DA are evaluated based on their success in allocating organs to and saving the lives of populations rather than individuals. In clinical settings, the importance placed on the collection and interpretation of biochemical and statistical information does not always hold true, as diseases cannot always be reduced to measurable variables, and as the trained judgement of clinicians proves helpful for responding to dynamically changing diseases.

Consequently, this section argues that the practices which rhetorically focus on individualized approaches to health and disease, in reality, enact medical interventions that place value on the health of populations over individuals. Because of a reliance on multivariate statistics and biochemical information, metabonomics practices directed at the development of personalized medicine succeed only in providing *individually-tailored probabilities* for populations. Such practices do not succeed in providing individualized approaches to patient care, which are typified in this chapter by the interpretive capacities of clinicians. Thus, as metabonomics researchers carry out experiments in the name of personalized medicine, their reliance on statistical practices produces diagnoses and treatments for health and disease that benefit populations rather than individuals. From this emerges a fundamental tension contained within personalized medicine, as both individuals and populations are implicated in medical interventions, but in unequal ways.

Discussion

Personalized medicine is enacted through metabonomics practices in unique ways, and with important challenges to the provision of medical care in modern clinical settings. "Personalized medicine" implies a medical practice that is tailored to individuals, and that provides a uniquely new approach to disease diagnosis and treatment through the use of molecular technologies. What I show throughout this chapter, however, is that personalized medicine's discursive emphasis on the individual belies the ways in which personalized medicine investigates and co-produces individuals and populations. Individuals cannot exist without populations, and the practices to investigate individuals inherently draw on populations to provide their meaning and context.

By emphasizing the co-production of individuals and populations, this chapter provides an alternative vision to Nikolas Rose's (2007) argument that the modern life sciences have become "molecularized," and consequently that biopolitics is being carried out not from above on populations but from below on individuals (Braun 2007). Interpreting Foucault's work on biopolitics in light of the rise of the post-genomic sciences, Rose argues that life and vitality—and their accompanying techniques for exercising power—are envisioned at the molecular level. Viewed within this framework, personalized medicine articulates biomedicine's increased focus on individual bodies, biologies, and health.

Drawing on my discussion of the BMM's efforts to develop biomedical Centres for individual and population phenotyping, I emphasize that the health of "individuals" only gains meaning in relation to the statistical measurements of "populations." I argue that personalized medicine differs from other forms of medical practice not in its ability to provide individually-tailored care, but in its reliance on molecular and statistical technologies for the investigation of health and disease. Such technologies entail new ways of making life and populations calculable, particularly through an increased focus on future disease risk. They focus on the suggestion or possibility of disease, rather than on its current manifestation. This shift away from therapies and towards prevention and diagnostics involves, as Kaushik Sunder Rajan (2006) comments, a type of forecasting or divination about the future. Individual health is reduced to a "probability calculus as a potential target for therapeutic intervention" (Rajan 2006:167), such that the target of medical intervention is not the manifestation but rather the suggestion or risk of disease.

The centrality of informational practices to the BMM's vision of personalized medicine, moreover, raises the overarching question of what type of "person" is being enacted through metabonomics research practices. Individual persons and populations are arbitrary categories, such that individual persons have been articulated and investigated in a variety of historically-situated ways. As Richard Tutton writes, "it is not the case that 'personalized medicine' has always existed: rather there have been different historical forms of 'personalization' over time" (Tutton 2012:1726). For example, Donna Haraway (1997) has suggested that in the age of genomics, the individual has increasingly been conceived of in terms of "the gene." Consequently, in the age of personalized medicine, how is the individual being conceived of and constructed by metabonomics?

As metabonomics engages in the rhetoric of personalization and complexity (see Chapter 3), it attempts to capture and model the living organism as a dynamic entity. But in doing so, it enacts the "person" as a molecular and statistical entity, intervening into their biology to characterize them according to biochemicals and probabilities. At stake, then, is the extent to which various aspects of individual health and disease can be captured and quantified by such informational practices. As Andrew's discussion of the young woman suffering from liver failure highlights, much of the "information" that clinicians use to diagnose and treat disease comes from the trained judgment of professionals (see Chapter 5) and their embodied interactions with patients. Thus, the claim that biology can be increasingly understood with the aggregation of statistical information (see Chapter 5) eclipses the fact that information about rapidly changing conditions, as well as their unique manifestation in individuals, is not easily quantified.

However, it is not only the individual and population that are being re-configured by metabonomics, but also the environment. As metabonomics enacts persons and populations in informational ways, it accounts for metabolic variables generated from a combination of endogenous metabolism, bacterial metabolism, and environment. Metabonomics, it might seem, suggests a more holistic and integrated understanding of individuals and populations. It suggests a move away from seeing the organism within its environment, and towards a view that sees the *environment as part of—and embedded within—the organism*. This speaks to Hannah Landecker's (2011) recent work on the ways in which nutritional epigenetics renders food as a form of environmental exposure, and in doing so makes environments come to matter in particular, notably molecular, ways. But as the environment comes into view, it—like persons and populations—is rendered into statistical and molecular forms. Embedded within this holistic view of the organism and the environment, therefore, is an environment that, like the complex metabolism I have discussed in Chapter 3, is portrayed in calculated and controlled ways.

In the end, enactments of individuals as measurable biological systems facilitate the calculation and prediction of health and disease in populations, but impede the diagnosis and treatment of dynamic conditions and individual bodies. The co-production of individuals and populations through metabonomics practices, therefore, is carried out in unequal ways. Despite a discursive focus on the health of the individual, metabonomics practices are aimed at the health of populations. As they are applied to clinical issues, their reliance on statistical and biochemical technologies demands certain medical interventions, and certain values on

information rather than the trained judgements of clinicians, to make health and disease knowable. With an informational enactment of persons and populations, perhaps the “social” aspects of personhood—the relational and dynamic capacities long described by anthropologists—are beginning to disappear.

Conclusion

I this dissertation, I have traced the contours, practices, and challenges that characterize research in the field of "metabonomics," the post-genomic study of metabolism. Linking metabonomics to the Biomolecular Medicine Laboratory (BMM) at Imperial College London, one of the field's primary sites of practice, I have provided a critical lens onto the behind-the-scenes, everyday work that goes into transforming biological materials into biological conclusions. Efforts to conduct experiments, to produce and interpret data, and to translate findings into medical practice rely on human and technological capabilities. Consequently, I have demonstrated how metabonomics researchers are engaged in a constant effort to understand, produce, and stabilize the biological relevance of their work. Researchers are engaged in various types of rhetorical work to transform contingent observations into authoritative practices for diagnosing and treating disease.

This dissertation makes three main contributions to anthropological studies of science. Firstly, it demonstrates *the centrality and materiality of data practices to metabonomics*—and to the life sciences more generally—by portraying the challenges inherent in the production and negotiation of large volumes and specific types of biological data. Data, which exist as measurements, statistical calculations, and computerized visualizations, cannot be reduced solely to the realm of cognitive representations. Instead, data must be examined as material things that entail distinct constellations of practices, technologies, and people. Within the metabonomics laboratory, data practices entail the selection and coding of statistical algorithms, the organization of biochemical data into tables and matrices, the interpretation of multi-dimensional statistical results, and the production of computer visualizations. These are material things that have historically developed at the intersection of distinct people, technologies, and ideologies, and which subsequently come to have specific and situated

uses. But data practices also entail the constant negotiation of ideas and values in order to produce knowledge and facts about the world. Data is never pre-existing or "raw": rather it entails choices about how to carry out experiments, as well as statements about the "correct" ways of doing research. Consequently, as metabonomics researchers carry out data practices in varied ways, they attempt to understand, test, and redefine the limits of quantitative, statistical reasoning in the biomedical sciences.

This dissertation, however, it not just about data and statistics. It is also about how *data practices are used to engage with the biological concepts of metabolism, the normal, health, and "life itself"* (Foucault 1990). Secondly, this dissertation demonstrates how data practices are entangled with contemporary ways of thinking about, intervening into, and enacting biology. Certain forms of data practices, most notably multivariate statistics, enable researchers to grapple with biology as a complex informational system. As researchers create large volumes of quantitative data, they both produce and necessitate methods for interpreting multifaceted biological systems. The tools and practices for investigating biology are intertwined with—shape and are shaped by—biomedical concepts and ideologies. Thus, multivariate statistics entail normative visions not only about the form that data should take, but also about the value data should have. Objective knowledge of biology can only be achieved, researchers assert, by engaging with multiple variables that must be analyzed in combination.

Thirdly, this dissertation explores how—and to what consequences—*metabonomics researchers understand, define, and translate health and disease into increasingly statistical and biochemical terms*. As they use biochemical and statistical technologies, metabonomics researchers struggle to evaluate and capture the dynamic and relational concepts of health and disease. The reduction of health and disease to numbers and measurements—though it involves constantly evolving technologies—has long been a feature of biomedical research. As such, the quantitative language and practices of the "normal" and "abnormal" allow health and disease to be rendered as statistical and information objects, but this rendering is not without fundamental problems. Despite claims to the molecular objectivity of metabonomics practices, health and disease cannot easily be reduced to numerical values and thresholds. Claims to molecular objectivity, therefore, sidestep fundamental questions about the ability of metabonomics practices to engage with biological complexity and normativity, human interpretation, and individual vitality. The treatment of the human body as a complex and

dynamic organism necessitates not only data, but also the trained judgments and individual attention of medical professionals.

On a broad level, this dissertation builds on anthropology's commitment to portray scientific research as a form of cultural practice. Building on Kaushik Sunder Rajan's (2005) call to explore "emergence"—processes of "reorganization and recalibration" rather than novelty—it examines *how, by what processes, and to what extent* persistent cultural themes and ideas are articulated in metabonomics practices. Consequently, in this dissertation I have argued that the practices of metabonomics are tied to and shaped by the historical, cultural, political, and economic context of research in the BMM. Metabonomics is surrounded by particular historical narratives, which emphasize the field's ties to multivariate statistics and nuclear magnetic resonance, as well as to funding and commercial opportunities. Moreover, contemporary projects in the BMM are linked to widespread efforts to develop translational research and personalized medicine. Technologies and practices do not operate within a vacuum, and instead are linked to 21st century themes such as the emergence of "big science" following the Human Genome Project (HGP), and the informaticization of the life sciences.

In detailing scientific practice as a form of culture, this dissertation also depicts the ways in which cultural idioms and biological knowledge interact and mutually inform one another. Biological concepts are imbued with and reflective of cultural ideas and values. Taking cues from feminist studies of technoscience, I have argued that the "natural" objects of metabonomics research are entangled with quantitative and statistical ways of thinking and doing research. Metabolism, tissues, diseases, populations are enacted in particular ways, such that they shape and are shaped by technologies and practices. The links between the laboratory and culture are not always overt, but the fundamental ideologies and values of metabonomics research relate to modern society's preoccupation with numbers, patterns, and probabilities. The understandings of biology and life produced by the 21st century life sciences are not the result of technological determinism, but rather entail the dynamic interaction and feedback loops between culture and science. Thus, I have built upon anthropology's imperative to move beyond simplistic accounts of the implications of the life sciences on social life and personhood, and have instead shown the ways in which social practices and meanings, and biological technologies and knowledge, are co-produced.

Central Themes and Problematics

In drawing this dissertation to a close, this section reviews the three central themes—data as material practices, technological enactments of biology, and the statisticalization of health and disease—to show how they flow across the various chapters. Before doing so, I turn to a story about Ryan, a researcher whose work I describe in Chapter 4, because it encapsulates the main challenges and tensions that my study of metabonomics has raised. It demonstrates, in other words, what is at stake in a changing landscape of laboratory research for notions of experimentation, understandings of biological life, and capacities to affect medical practices.

I spoke to Ryan over lunch one day about his research, which involved a joint project between Imperial College London and the National Institute of Health (NIH) in the United States. His research explored the metabolic pathways involved in the development of inborn errors of metabolism, a group of genetic diseases that are largely caused by defects in single genes coding for enzymes or proteins involved in various types of metabolism. To carry out this research, Ryan had arranged for the NIH to provide him with access to molecular biology facilities and collect biological samples. Working with a laboratory at the NIH that specialized in research on undiagnosed metabolic diseases, Ryan travelled to the United States several times a year to gain access to molecular biology facilities, and most importantly, to biological samples from patients. Back at Imperial, he used these samples to carry out metabonomics experiments, with which he examined the metabolic pathways involved in the development of inborn errors of metabolism, and attempted to correlate this information with genetic information.

As Ryan explained his research project, he lamented that he was having a difficult time fitting in with his laboratory at the NIH. He enjoyed the people and the facilities, but struggled to get along with the head of the laboratory. A renowned scientist who specialized in research on a number of metabolic diseases like cystinosis, alkaptonuria, and polycystic kidney disease, the head of the laboratory drew on a variety of genetic techniques to determine the molecular mechanisms underlying the development of disease. Research on inborn errors of metabolism has a long history that dates back to the work of British physician Archibald Garrod, who identified that alkaptonuria—a disorder of phenylalanine and tyrosine metabolism—was caused by a Mendelian recessive trait. Importantly, Garrod's work was one of the building blocks for the "one gene-one enzyme" hypothesis, a central

tenet of molecular biology which, although now out-dated, stated that single genes lead to the production of single enzymes, and affect single steps in a metabolic pathway.

To this end, the techniques that the head of the NIH laboratory used to investigate inborn errors of metabolism did not include metabonomics, with which the head of the laboratory had little experience or interest. The head of the laboratory did not believe that Ryan's work was relevant to the study of inborn errors of metabolism: he was unconvinced that Ryan's work with nuclear magnetic resonance (NMR) and mass spectrometry (MS)—much less his work with multivariate statistics and chemometrics—would be useful for conducting research that would ultimately help diagnose patients with metabolic diseases. Though Ryan had published a number of papers in high-impact journals and had carried out a range of metabonomics experiments, the head of the laboratory had remarked with dissatisfaction: "Does it bother you that you haven't done any experiments yet for your PhD?"

The head of the laboratory's dissatisfaction with Ryan's work crystallizes the central themes that flow throughout this work. Most notably, Ryan's metabonomics research entailed a distinct set of practices for generating, manipulating, and interpreting data. At the BMM, Ryan spent most of his time at his computer, generating multivariate statistical analyses rather than doing the "wet" experiments of molecular biology. This entailed a different definition of "experimentation," as Ryan's experimental practices were unfamiliar to scientists engaging in the study of inborn errors of metabolism. However, Ryan's research entailed not only different practices, but also different understandings and enactments of what biology was in the first place. Researchers in the field of inborn errors of metabolism engaged with a more linear concept of biology, which followed the notion that single gene defects led to abnormalities in protein synthesis or pathways. In contrast, metabonomics researchers like Ryan engaged with a model of biological causality that was multifactorial and contingent, in which metabolic pathways and interactions occurred across multiple places and scales.

Consequently, it was not surprising that the head of the laboratory distrusted the ability of metabonomics—which as I have argued in Chapter 2 is an emerging and as yet unknown field—to address clinical issues such as the diagnosis of disease. Though he was caught up in boundary work over the meaning of metabonomics in relation to genomics, he also flagged the inherent challenges faced by metabonomics researchers in making clinical

sense of their data. Consequently, this episode reflects the ongoing challenges faced by the field of metabonomics, as it positions data as central to investigations of biology, as its creates visions of life that are caught up with its technologies, and as it attempts to implement increasingly statistical conceptions of disease within medical practice.

Data as Material Practices

Data and informational practices are becoming increasingly central to the biological sciences, and are changing how researchers both practice and understand biology. As part of the broad pattern of informaticization of the life sciences, data practices increasingly define the epistemic machinery and status of metabonomics as a field. They are intrinsic, both conceptually and historically, to the way metabonomics produces and values knowledge. Consequently, computation, quantification, and statistics are central not only to the generation and management of information, but also to its analysis and interpretation in relation to biology. Specific ways of manipulating data are central to claims about "good" research practices, but are also central to the generation of insightful biological claims.

To this end, metabonomics places value not only on data and statistics, but also *on certain forms of data and statistics*. The choices that researchers make to carry out data analysis are constrained by the types of analysis that are considered "correct" according to the norms of metabonomics as a field, and according to the institutions of the BMM as a laboratory. Researchers are taught to view and approach data in a multivariate way—as a combination of multiple variables whose meaning lies in a pattern as a whole—and to resort to certain types of algorithms for data analysis. Consequently, multivariate statistics are normalized and naturalized, such that they become intrinsic both to the ways that knowledge is produced and to the ways the field defines itself. Multivariate statistics entail the "right" way of doing data analysis. They provide a more insightful and objective view of biological processes and systems than other types of analysis, in particular univariate statistics.

Multivariate statistics, in turn, are not pre-existing or neutral entities, but rather are tied to distinct historical and material developments. Though researchers draw distinctions between univariate and multivariate ways of analysing and interpreting data, the distinction between the two approaches is not so clear cut. Multivariate statistics consist of many different mathematical approaches and computational techniques, which involve different and overlapping historical lineages. On a broader level, the development of multivariate statistics is deeply embedded within post-WWII developments in computing, which enabled

researchers to both produce increased volumes of data and also draw on increased computational power.

The practices of multivariate statistics, however, are also characterized by an inherent tension between institutionalized ways of producing knowledge and the craftwork that knowledge production entails. On the one hand, as I have argued throughout this dissertation, the use of statistics is interpretive and creative. Because researchers in the BMM do not typically rely on standardized, black-boxed computer software to carry out data analysis, they make active decisions about which types and styles of analysis to conduct. As one researcher commented in a seminar, the answers researchers achieved from their data depended on the kinds of algorithms and models they used. He said: "With PCA all you're really doing is creating a different description of the data, and you're really focused on what the model is telling you…And the whole subjectivity of it…is kind of inherent." Viewed in this way, the practice of statistics involves inherent choices and values, but also involves interpretive aspects of art and craft.

Technological Enactments of Biology

Particular conceptions of biology, metabolism, health, and life are *made and made real* through the data practices of metabonomics. Technologies for investigating biology are intertwined with views and ideologies of biology: just as biology and culture co-produce one another, so do multivariate statistics and complex, informational views of metabolism. As metabonomics researchers invest in a multivariate statistical view of the world as something new and exciting, they move away from a view of biology as something that can be understood through linear causes and effects. Studies of metabolism cannot operate with the metaphors of code eschewed in early forms of genomics, but must make recourse instead to notions of systems and networks.

To this end, the entanglement of technological practices and biology allows metabonomics researchers to enact complex views of metabolism. As biological experiments produce large volumes of data, researchers assert the need to use multivariate statistics in order to understand and engage with the complexity of biology. But as metabonomics embraces notions of biological complexity—promoting understandings of metabolism as a system of interlocking biochemicals, organs, organisms, and environments—it does so in controlled and limited ways. By encountering complexity through those things that can be quantified in laboratory practice, metabonomics struggles to engage with those things that are

not easily measured or quantified. To this end, metabonomics entails *normative visions of biological complexity*, which discount the kinds of dynamic complexity that clinicians routinely encounter with individual patients in medical practice.

But the entanglement of technological practices and biology also poses key challenges to metabonomics' conceptions of biology and disease. As metabonomics researchers enact statistical and biochemical concepts of the normal and abnormal, they transform the dynamic states of health and disease into concepts that can be measured and quantified within the laboratory environment. Though organisms change over time and in relation to the environment, their state of being is transformed into a binary and decontextualized fact. But as researchers grapple with the inherent contingency of normal and abnormal, they encounter difficulties in directly relating these concepts to health and disease. Statistical and biochemical enactments of disease struggle to engage with the variability and dynamism of life, which are operating concepts outside of the controlled context of the laboratory.

The Statisticalization of Health and Disease

Data practices are becoming increasingly central not only to the understanding of, but also to the diagnosis and treatment of health and disease. Metabonomics is characterized fundamentally by the notion that quantitative measurements and statistical probabilities can supplant, or at the very least improve upon, clinical practices in a variety of settings. Researchers place value on the generation and interpretation of multivariate forms of data, proclaiming that they are more precise and "objective" ways of assessing health and disease. To this end, as researchers attempt to develop alternative forms of tissue classifications or disease diagnostics, they place value on the collection and analysis of molecular data. This is particularly salient in attempts to replace histopathological practices with magic angle spinning nuclear magnetic resonance (MAS-NMR) and matrix-assisted laser desorption/ionization mass spectroscopy imaging (MALDI-MSI). As such, researchers conceal the contingency involved in metabonomics practices, and place value on the practices of statistics and biochemistry over the practices of human judgement.

Amidst the value placed on data practices in the diagnosis and treatment of disease, the biological meaning and relevance of data practices remains a central concern to metabonomics researchers. Because the meanings contained within multivariate statistical analyses are never inherent or pre-existing, the interpretation of metabonomics data is a central, ongoing, and challenging endeavor. Metabonomics researchers continually struggle

to negotiate the interpretation and use of complex biological information. As researchers enact tissues as a series of biochemical signals and mathematical patterns, they are faced with questions about how to link quantitative measures to biological processes. Moreover, as researchers attempt to implement such quantitative measures in clinical practice, they are faced with further questions about how—if at all—complex information can be used in clinical and surgical scenarios that rely on rapid, straightforward decision-making.

The translation of metabonomics knowledge into clinical practices also raises fundamental questions about the ways in which statistical practices can be used to produce knowledge of and interventions for both individuals and populations. Individuals and populations are not pre-existing categories, and are instead made real through laboratory practices aimed at quantifying health and disease. Within metabonomics, individuals and populations are recast as a sum total of metabolic and statistical variables, such that bodily conditions are measured in relation to environmental variables. As researchers attempt to develop "personalized medicine," they place discursive focus on the investigation and management of individuals. In practice, however, researchers do not focus solely on individuals, and instead carry out practices that co-produce information about individuals and populations. As researchers draw on informational personalized medicine practices rather than interpretive clinical practices to tailor therapies to individuals, they enact organisms as enclosed information systems rather than dynamic entities. In doing so, they struggle to produce knowledge of and clinical applications aimed at the care of individual persons, and instead produce informational and probabilistic understandings of populations.

Future Implications

Amidst my concern with tracking the practices, challenges, and implications of using data practices to comment on health and disease, this dissertation has also explored the extent to which observations about metabonomics—as a field, and as a site of practice at Imperial College London—can be generalized to talk about the contemporary sciences more broadly. Since the completion of my fieldwork, both the BMM and the field of metabonomics have undergone processes of rearrangement and restructuring. To this end, in the summer of 2012, less than one year after I had completed my fieldwork, the BMM gained access to the Olympic drug testing facilities and launched a major project to carry out the large-scale, longitudinal phenotyping research. At the same time, the BMM recruited and hired a well-known researcher whose work with mass spectrometry was one of the major sources of

competition during my fieldwork. The laboratory began to take on a large number of clinical researchers to do short-term clinical projects, in a move to expand its translational research capacities. Placing a permanent stamp on these changes, at the beginning of 2013 the laboratory renamed itself from Biomolecular Medicine (BMM) to Computation and Systems Medicine (CSM). The shift in names itself was indicative of the laboratory's increasing focus on translational and personalized medicine, and highlighted the ways in which metabonomics was emerging and evolving over time.

In addition, many of the researchers whose lives and work form the core of my writing have since moved on from the laboratory to have families, to conclude their own doctoral theses and relocate to other research positions, to return to their home countries, to go back to their medical training, or to switch to other research fields or occupations. When I chatted with one researcher who had left the BMM to pursue a research fellowship at another University, she commented that the BMM had completely changed. The laboratory, she claimed, was full of unfamiliar faces, along with many recently acquired pieces of equipment. As she described the influx of people and equipment, I recalled seeing a stern announcement on the laboratory's mailing list about the tightening up of the equipment booking system, due to the increased demands on the laboratory's facilities and resources.

My dissertation, therefore, provides a snapshot of a rapidly changing area of research, in which biological ideas and technologies are evolving concurrently. On a broad scale, the past few years has seen a surge of developments in metabonomics throughout the United Kingdom, Europe, and the United States. The number of laboratories conducting research in metabonomics has grown considerably, supported by the funding of national bodies, investments by commercial entities, and increased efforts to mobilize metabonomics technologies in clinical settings. As a key example, the National Institute of Health (NIH) has committed more than $50 million to metabolomics research over five years between 2012 and 2017. Investments in metabonomics are increasing, placing the field's research capacities alongside more well-known fields like proteomics and epigenetics.

The past few years have also seen the fields of metabonomics and metabolomics become increasingly intertwined and integrated, though they entail different sets of histories and practices. Citation data indicates that usage of the term metabonomics is decreasing in relation to the term metabolomics, signalling perhaps that the fields are being centralized and subsumed under one heading (see Figure 39). More and more, the use of MS rather than

NMR, which continues to form a key component of the BMM's activities, is seen as central to the investigation of metabolism[77]. To this end, metabonomics will continue to grow and evolve, as it becomes increasingly integrated with not only metabolomics, but also the broad attempts of the life sciences to characterize and influence the treatment of human disease.

Figure 39: Google Trends information comparing web searches for metabolomics and metabonomics.
Note how (1) metabolomics is more widely known and cited than metabonomics, and (2) metabolomics is, according to Google Trends data, a more well-known field of research.

Given the centrality of data practices to the life sciences as a whole, therefore, the claims made in this dissertation are not limited to my ethnography of the BMM, or to the field of metabonomics. The data practices that characterize and are central to metabonomics are common to many other fields in both the social and natural sciences, which are concerned with the generation and interpretation of large volumes of information. Multivariate statistics and complex thinking are widespread throughout post-genomic fields like bioinformatics, epigenetics, proteomics, and systems biology, which have similar methods and ask related questions to metabonomics.

But data practices are also becoming increasingly central to fields of research that emerged before or have existed independently of the post-genomic sciences. It is not only the hard sciences, but also social sciences that are attempting to harness the "power" of data to reveal complex trends and patterns about human behavior and culture. As social scientists

[77] MS, which characterizes the field of metabolomics rather than metabonomics, is hailed as being able to resolve—although not necessarily interpret—a greater number and variety of metabolites.

attempt not only to access large datasets, but also to transform qualitative information into quantitative data, they raise similar questions about the value of data in-and-of-itself, the role of human interpretation, and the ability to understand persons and populations. Consequently, the issues raised in this thesis extend beyond metabonomics to touch upon fields as diverse as anthropology, epidemiology, economics, and climate science.

Metabonomics, therefore, provides a key site for identifying and engaging with the data practices, and the ensuing negotiations and challenges, that make up modern cultures of biomedical research. For example, as I travelled on the train to London during my fieldwork, I spoke to a psychology researcher who used functional magnetic resonance imaging (fMRI), a technology that measures brain activity via blood flow, to examine the biological processes underlying consumer choices. I was struck by the overarching similarities in the ways he described the challenges inherent in his everyday work. These included coping with the volume of data generated by fMRI experiments, and figuring out how best—and objectively—to make sense of such data. Though psychology and metabonomics were fields and institutions apart, they espoused similar values regarding the objectivity of molecular data, but also similar problems regarding data generation and interpretation.

Such invocations of objectivity relate in many ways to recent attempts throughout the biological sciences to generate more accurate and precise molecular diagnostics for health and disease. On a broad scale, researchers throughout Europe and the United States have embraced the notion of "precision medicine" (Committee on a Framework for Development a New Taxonomy of Disease 2011; Mirnezami et al. 2012; Katsnelson 2013) as a means to develop a molecular classification or taxonomy for disease. Precision medicine, they claim, should be based on a combination of "multiparametric molecular data with clinical data…environmental data, and health outcomes" (Desmond-Hellmann 2012:1). On a small scale, researchers at Cambridge University recently announced the application of deep space astronomy imaging algorithms to the diagnosis of cancer pathology. They claimed that "the computer was just as accurate as the manual system" (Paddock 22 February 2013), with manual system referring to human histopathologists. Overall, such claims typify movements in the biological sciences to eclipse the interpretive capacities of human beings with the accumulation and automation of data through patterns, systems, and multivariate models.

On a broader level, issues with data also typify broad discussions surrounding the use of "big data," large data sets that cut across multiple types and realms of information, in

society[78]. Big data, which largely draws on existing data sets, is increasingly used in finance, business, and research to make observations, claims, and decisions about money, human behaviour, and the allocation of resources[79]. It is used by prominent internet companies such as Google and Facebook, and is increasingly being mobilized to tailor and predict individual choices. As philosopher Sabina Leonelli writes, big data underlies the notion that "pooling together results…maximizes the changes of spotting significant patterns in data that are collected, and thus of transforming data into knowledge" (Leonelli forthcoming). Big data, consequently, raises many of the fundamental questions presented throughout this dissertation about the appropriate uses and values of data, and in particular about the suitability of data practices to address dynamic objects or to eclipse human intuition and judgment.

To this end, the *New York Times* recently reported on a newly formed company called "Factual," whose sole aim is to use statistics to identify every fact that exists in the digital world, in an effort to find the answers to the world's most pressing problems. Factual typifies big data's approach to the world, in which data is believed to be the be-all-and-end-all to the world's problems, but also has much in common with metabonomics. Factual's existence is made possible by modern society's obsession with data: it conceives of the world as "one big data problem," such that "data and algorithms [can be used] to find patterns in nature and society, for scientists to observe and businesses to exploit" (Hardy 24 March 2012). In doing so, Factual raises important questions. Does big data need human judgment to figure out which correlations are meaningful and how? Does big data consider whether some types of information, in particular the moral and ethical dimensions of human health and behaviour, are more suited to quantification than others?

Such questions—although they are being considered in relation to the widely publicized failure of big data in the Financial Crisis in 2008, or in relation to the

---

[78] For more information on the intersection between big data and society, see a series of recent media articles on the subject, which include columns by *The New York Times* Op-Ed journalist David Brooks (15 April 2013; 18 February 2013), a special focus in *The New York Times* "Bits Blog" (Clifford 19 June 2013), and a special report by *The Economist* (27 February 2010).

[79] As a recent *New York Times* article describes:

> Big Data is a vague term, used loosely, if often, these days. But put simply, the catchall phrase means three things. First, it is a bundle of technologies. Second, it is a potential revolution in measurement. And third, it is a point of view, or philosophy, about how decisions will be — and perhaps should be — made in the future. (Lohr 19 June 2013)

controversies surrounding the United States National Security Agency's (NSA) use of private and personal data for intelligence uses[80]—form one of the central concerns of this dissertation. Asked in the context of the application of data practices to biomedical practice, they highlight philosophical and moral worries about what limits exist—if any—for the movement of statistics into realms of practice traditionally occupied by human interpretation and judgment. Can computer-assisted diagnostics replace the skill of trained and experienced doctors, whose interpretive abilities are considered by some to be "as barbaric as blood-letting" (Davidi 12 April 2013) in the face of data-driven research? Can metabonomics and other fields of post-genomic research overcome the very real bottleneck of data interpretation by enlisting the combined aid of "algorithms, machines, and people" (Patterson 5 December 2011)? Such questions remained unanswered and unexamined, and require the urgent attention of social scientists.

Consequently, this dissertation raises key concerns about the ability and limits of statistics in the biological sciences to capture the vitality and normativity of life. Though metabonomics researchers assert that challenges with disease diagnosis and treatment can be solved by the aggregation of large volumes of information, I am sceptical that statistics and informational practices can engage with health and disease as individual, dynamic, and complex entities. To this end, the trained judgment and interpretive abilities of medical practitioners still remain central to the diagnosis and treatment of disease. Visions of the technological futures of clinical and public health interventions insist that technologies can replace human judgment with precise molecular measurements. But on a fundamental level, such technologies succeed only in rearranging—and not removing—human judgment from biomedical practice.

Furthermore, as the post-genomic sciences produce increasingly statistical enactments of complexity and disease, this dissertation explores whether we are truly moving beyond—or rather finding ourselves wading among new permutations of—the reductionism of molecular biology and genetics. Scholars such as Margaret Lock (2012), Nikolas Rose (2013), Jorge Niewöhner (2011), Stefan Helmreich (2009), and Hannah Landecker (2011) have suggested that with the rise of post-genomic fields like epigenetics, the biological sciences are providing alternative frameworks for engaging with the dynamic relationships

[80] A quote from this commentary, which explores the similarities between Silicon Valley technology businesses and the NSA, exclaims: "We are all in these Big Data business models."

between nature and nurture, or between the organism and its environment. It is my belief, however, that although biological sciences are engaging with increasingly relational and dynamic conceptions of life, we are being faced with newer and more convoluted notions of reductionism, which are enmeshed within rhetorics of complexity and holism. Though metabolism is rendered as dynamic, or as Hannah Landecker writes, "a set of molecules that both interacts with environmental molecules and is iteratively conditioned by exposure to the environment" (Landecker 2011:180), metabolism is nonetheless a site of calculation and regulation of both the body and the environment at the molecular level.

In the end, anthropologists will play an important role in examining the challenges and consequences that lie ahead, as the life sciences produce new understandings of life, as biomedical practitioners experience changes in the practices for diagnosis disease, and as biomedicine is rearranged to accommodate the technologies, ideologies, and values of the post-genomic sciences. We must engage—as many scholars have proclaimed following the criticisms of the Science Wars and social constructionist approaches to science—with these emerging discourses and practices as things that have the potential to inform our own social-scientific ways of seeing the world. But we must also be critical of the ideas and values that they espouse. Ultimately, our ability to see sites of scientific practice as things that are both embedded within larger networks of knowledge production and also connected to broader patterns of informaticization and commodification, gives us the opportunity to challenge existing and emerging notions of biological life, and to find new links between science and culture.

Bibliography

Abram, S. and M. E. Lien (2011). "Performing Nature at World's Ends." Ethnos **76**(1): 3-18.

Ackerknecht, E. H. (1982). *A Short History of Medicine*. Baltimore: Johns Hopkins University Press.

Adelswärd, V. and L. Sachs (1996). "The Meaning of 6.8: Numeracy and Normality in Health Information Talks." Social Science & Medicine **43**(8): 1179-87.

Agrocos. (2012). "Bruker BioSpin GmbH (Germany)." from http://www.agrocos.eu/index.php?option=com_content&view=article&id=65&Itemid=80.

Alac, M. (2008). "Working with Brain Scans." Social Studies of Science **38**(4): 483-508.

Alonso, W. and P. Starr (1987). *The Politics of Numbers*. New York: Russell Sage Foundation Publications.

Amit, V. (2000). *Constructing the Field: Ethnographic Fieldwork in the Contemporary World*. Abingdon: Routledge.

Ankeny, R. A. and S. Leonelli (2011). "What's So Special about Model Organisms?" Studies in History and Philosophy of Science Part A **42**(2): 313-23.

Armstrong, D. (1995). "The Rise of Surveillance Medicine." Sociology of Health & Illness **17**(3): 393-404.

Baker, M. (27 February 2013). "Big Biology: The 'Omes Puzzle." Nature News, from http://www.nature.com/news/big-biology-the-omes-puzzle-1.12484.

Baker, M. (2011). "Companies Ponder How Truly 'Personal' Medicines Can Get." Nature Medicine **17**(5): 519-19.

Barad, K. (2003). "Posthumanist Performativity: Toward an Understanding of How Matter Comes to Matter." Signs **28**(3): 801-31.

Barton, R. H. (2011). "A Decade of Advances in Metabonomics." Expert Opinion on Drug Metabolism & Toxicology **7**(2): 129-36.

Bates, S. (2010). "Progress Towards Personalized Medicine." Drug Discovery Today **15**(3-4): 115-20.

Beaulieu, A. (2001). "Voxels in the Brain: Neuroscience, Informatics and Changing Notions of Objectivity." Social Studies of Science **31**(5): 635-80.

Beaulieu, A. (2004). "From Brainbank to Database: The Informational Turn in the Study of the Brain." Studies in History and Philosophy of Science Part C **35**(2): 367-90.

Beck, U. (1992). *Risk Society: Towards a New Modernity*. London: Sage.

Beger, R. and T. Colatsky (2012). "Metabolomics Data and the Biomarker Qualification Process." Metabolomics **8**(1): 2-7.

Bettelheim, F. A., W. H. Brown, et al. (2009). *Introduction to General, Organic and Biochemistry*. Belmont: Brooks Cole.

Bhattacharya, A. (3 February 2010). "Chemistry: Breaking the Billion-Hertz Barrier." Nature, from http://www.nature.com/news/2010/100203/full/463605a.html.

Bhattacharya, A. (14 December 2009). "Surgeons Get Real-Time Tissue Profiling." Nature, from http://www.nature.com/news/2009/091214/full/news.2009.1128.html.

Bictash, M., T. M. Ebbels, et al. (2010). "Opening up the "Black Box": Metabolic Phenotyping and Metabolome-Wide Association Studies in Epidemiology." Journal of Clinical Epidemiology **63**(9): 970-79.

Blow, N. (2008). "Metabolomics: Biochemistry's New Look." Nature **455**(7213): 697-700.

Braun, B. (2007). "Biopolitics and the Molecularization of Life." Cultural Geographies **14**(1): 6-28.

Brooks, D. (15 April 2013). "What You'll Do Next." New York Times, from http://www.nytimes.com/2013/04/16/opinion/brooks-what-youll-do-next.html.

Brooks, D. (18 February 2013). "What Data Can't Do." New York Times, from http://www.nytimes.com/2013/02/19/opinion/brooks-what-data-cant-do.html.

Brown, N. and M. Michael (2003). "A Sociology of Expectations: Retrospecting Prospects and Prospecting Retrospects." Technology Analysis & Strategic Management **15**(1): 3-18.

Bruker Biospin. (2013). "JuiceScreener - Technical Details." from http://www.bruker.com/products/mr/nmr/food-screener/juicescreener/technical-details.html.

Burri, R. V. (2008). "Doing Distinctions Boundary Work and Symbolic Capital in Radiology." Social Studies of Science **38**(1): 35-62.

Buscher, M., D. Goodwin, et al. (2010). *Ethnographies of Diagnostic Work: Dimensions of Transformative Practice*. Basingstoke: Palgrave Macmillan.

Butler, J. (1993). *Bodies That Matter: On the Discursive Limits of "Sex"*. London: Routledge.

Cajka, T., K. Riddellova, et al. (2011). "Ambient Mass Spectrometry Employing a DART Ion Source for Metabolomic Fingerprinting/profiling: A Powerful Tool for Beer Origin Recognition." Metabolomics **7**(4): 500-08.

Calvert, J. (2008). "The Commodification of Emergence: Systems Biology, Synthetic Biology and Intellectual Property." BioSocieties **3**(4): 383-98.

Calvert, J. and J. H. Fujimura (2011). "Calculating Life? Duelling Discourses in Interdisciplinary Systems Biology." Studies in History and Philosophy of Science Part C **42**(2): 155-63.

Canguilhem, G. (1989). *The Normal and the Pathological*. Cambridge: MIT Press.

Canguilhem, G. and J. Savage (2001). "The Living and Its Milieu." Grey Room: 7-31.

Carmel, S. (2012). "The Craft of Intensive Care Medicine." Sociology of Health & Illness **35**(5): 731-45.

Carusi, A. (2012). "Making the Visual Visible in Philosophy of Science." Spontaneous Generations **6**(1): 106-14.

Chan, E. C. Y., K. K. Pasikanti, et al. (2011). "Global Urinary Metabolic Profiling Procedures Using Gas Chromatography-Mass Spectrometry." Nature Protocols **6**(10): 1483-99.

Chin, L., J. N. Andersen, et al. (2011). "Cancer Genomics: From Discovery Science to Personalized Medicine." Nature Medicine **17**(3): 297-303.

Cholongitas, E., G. Germani, et al. (2010). "Prioritization for Liver Transplantation." Nature Reviews Gastroenterology and Hepatology **7**(12): 659-68.

Chow-White, P. a. and M. Garcia-Sancho (2011). "Bidirectional Shaping and Spaces of Convergence: Interactions between Biology and Computing from the First DNA Sequencers to Global Genome Databases." Science, Technology & Human Values **37**(1): 124-64.

Clayton, T. A., D. Baker, et al. (2009). "Pharmacometabonomic Identification of a Significant Host-Microbiome Metabolic Interaction Affecting Human Drug Metabolism." Proceedings of the National Academy of Sciences **106**(34): 14728-33.

Clayton, T. A., J. C. Lindon, et al. (2006). "Pharmaco-Metabonomic Phenotyping and Personalized Drug Treatment." Nature **440**(7087): 1073-77.

Clifford, S. (19 June 2013). "Using Data to Stage-Manage Paths to the Prescription Counter." New York Times, from http://bits.blogs.nytimes.com/2013/06/19/using-data-to-stage-manage-paths-to-the-prescription-counter/.

Code, L. (2000). *Encyclopedia of Feminist Theories*. London: Taylor & Francis.

Cohn, S. (2004). "Increasing Resolution, Intensifying Ambiguity: An Ethnographic Account of Seeing Life in Brain Scans." Economy & Society **33**(1): 52-76.

Cohn, S. (2007). "Seeing and Drawing: the Role of Play in Medical Imaging." In: *Skilled Visions: Between Apprenticeship and Standards* by C. Grasseni. Oxford: Berghahn Books, 91-105.

Collier, S. J. and A. Ong (2005). *Global Assemblages Anthropological Problems*. Malden: Blackwell.

Committee on a Framework for Development a New Taxonomy of Disease (2011). Toward Precision Medicine: Building a Knowledge Network for Biomedical Research and a New Taxonomy of Disease. Washington DC, National Academy of Sciences.

Coopmans, C. (2011). "'Face Value': New Medical Imaging Software in Commercial View." Social Studies of Science **41**(2): 155-76.

Coughlan, S. (26 September 2011). "Harvard Endowment Rises to £21bn." BBC News, from http://www.bbc.co.uk/news/education-15061374.

Csordas, T. J. (1994). *Embodiment and Experience: The Existential Ground of Culture and Self*. Cambridge: Cambridge University Press.

Daston, L. and P. Galison (2007). *Objectivity*. New York: Zone Books.

Davidi, A. (12 April 2013). "Big Data and Society - Interview with Kenneth Cukier, The Economist." The Guardian, from http://www.guardian.co.uk/media-network/media-network-blog/2013/apr/12/big-data-privacy-economist.

Desmond-Hellmann, S. (2012). "Toward Precision Medicine: A New Social Contract?" Science Translational Medicine **4**(127): 127ed3.

Desrosières, A. and C. Naish (2002). *The Politics of Large Numbers: A History of Statistical Reasoning*. Cambridge: Harvard University Press.

Doing, P. (2009). *Velvet Revolution at the Synchrotron: Biology, Physics, and Change in Science*. Cambridge: MIT Press.

Douglas, M. (2002). *Purity and Danger: An Analysis of Concepts of Pollution and Taboo*. London: Routledge.

Draghia-Akli, R. (2012). "Enabling Personalized Medicine in Europe: A Look at the European Commission's Funding Activities in the Field of Personalized Medicine Research." Personalized Medicine **9**(2): 151-55.

Dumit, J. (2003). *Picturing Personhood: Brain Scans and Biomedical Identity*. Princeton: Princeton University Press.

Edwards, P. N. (2010). *A Vast Machine: Computer Models, Climate Data, and the Politics of Global Warming*. Cambridge: MIT Press.

Edwards, P. N., M. S. Mayernik, et al. (2011). "Science Friction: Data, Metadata, and Collaboration." Social Studies of Science **41**(5): 667-90.

Eisenberg, A. (2 February 2013). "Genomic Analysis, the Office Edition." New York Times, from http://www.nytimes.com/2013/02/03/business/knomes-new-machine-to-aid-labs-in-genomic-analysis.html?smid=tw-share.

Ensmenger, N. (2012). "Is Chess the Drosophila of Artificial Intelligence? A Social History of an Algorithm." Social Studies of Science **42**(1): 5-30.

Ernst, W. (2006). *Histories of the Normal and the Abnormal: Social and Cultural Histories of Norms and Normativity*. Abingdon: Routledge.

Esbensen, K. and P. Geladi (1990). "The Start and Early History of Chemometrics: Selected Interviews. Part 2." Journal of Chemometrics **4**(6): 389-412.

EuroBioForum. (2013). "United Kingdom." from http://www.eurobioforum.eu/2105/observatory/united-kingdom/.

European Commission. (2013a). "The EU Framework Programme for Research and Innovation." from http://ec.europa.eu/research/horizon2020/index_en.cfm?pg=h2020.

European Commission. (2013b). "Seventh Framework Programme (Fp7)." from http://cordis.europa.eu/fp7/home_en.html.

European Science Foundation (2012). Personalised Medicine for the European Citizen: Towards More Precise Medicine for the Diagnosis, Treatment and Prevention of Disease (iPM).

Evans, J. (1995). *Biomolecular NMR Spectroscopy*. Oxford: Oxford University Press.

Evans, J. P., E. M. Meslin, et al. (2011). "Deflating the Genomic Bubble." Science **331**(6019): 861.

Farmer, P. (2001). *Infections and Inequalities: The Modern Plagues*. Berkeley: University of California Press.

Farrant, R. D., J. C. Lindon, et al. (1992). "An Automatic Data Reduction and Transfer Method to Aid Pattern Recognition Analysis and Classification of NMR Spectra." Journal of Pharmaceutical and Biomedical Analysis **10**(2–3): 141-44.

Fassin, D. (2007). *When Bodies Remember: Experiences and Politics of AIDS in South Africa*. Berkeley: University of California Press.

Filler, A. G. (2009). The History, Development and Impact of Computed Imaging in Neurological Diagnosis and Neurosurgery: CT, MRI, and DTI. Santa Monica, Institute for Nerve Medicine.

Fleck, L. (1986). "Some Specific Features of the Medical Way of Thinking [1927]." In: *Cognition and Fact - Materials on Ludwik Fleck* by R. S. Cohen and T. Schnelle. Dordrecht: D. Reidel Publishing Company, 39-46.

Fonville, J. M., C. Carter, et al. (2012a). "Robust Data Processing and Normalization Strategy for MALDI Mass Spectrometric Imaging." Analytical Chemistry **84**(3): 1310-19.

Fonville, J. M., C. L. Carter, et al. (2012b). "Hyperspectral Visualization of Mass Spectrometry Imaging Data." Analytical Chemistry **85**(3): 1415-23.

Fonville, J. M., S. E. Richards, et al. (2010). "The Evolution of Partial Least Squares Models and Related Chemometric Approaches in Metabonomics and Metabolic Phenotyping." Journal of Chemometrics **24**: 636-49.

Forsythe, D. E. (1999). "Ethics and Politics of Studying up in Technoscience." Anthropology of Work Review **20**(1): 6-11.

Fortun, K. and M. Fortun (2005). "Scientific Imaginaries and Ethical Plateaus in Contemporary US Toxicology." American Anthropologist **107**(1): 43-54.

Fortun, M. (2008). *Promising Genomics: Iceland and deCODE Genetics in a World of Speculation*. Berkeley: University of California Press.

Foucault, M. (1977). *Discipline & Punish: The Birth of the Prison*. Sheridan: Vintage.

Foucault, M. (1990). *The History of Sexuality*. New York Vintage.

Foucault, M. (2003). *The Birth of the Clinic*. London: Routledge.

Foucault, M., L. H. Martin, et al. (1988). *Technologies of the Self: A Seminar with Michel Foucault*. Amherst: University of Massachusetts Press.

Fox Keller, E. (2002). *The Century of the Gene*. Cambridge: Harvard University Press.

Frank, R. and R. Hargreaves (2003). "Clinical Biomarkers in Drug Discovery and Development." Nature Reviews Drug Discovery **2**(7): 566-80.

Franklin, S. (1995). "Science as Culture, Cultures of Science." Annual Review of Anthropology(24): 163-84.

Franklin, S. (2002). "The Anthropology of Science." In: *Exotic No More: Anthropology on the Front Lines* by J. MacClancy. Chicago: University of Chicago Press, 351-58.

Franklin, S. and C. Roberts (2006). *Born and Made: An Ethnography of Preimplantation Genetic Diagnosis*. Princeton: Princeton University Press.

Fujimura, J. H. and R. Rajagopalan (2011). "Different Differences: The Use of 'Genetic Ancestry'versus Race in Biomedical Human Genetic Research." Social Studies of Science **41**(1): 5-30.

Fullwiley, D. (2011). *The Enculturated Gene: Sickle Cell Health Politics and Biological Difference in West Africa*. Princeton: Princeton University Press.

Gartland, K. P. R., S. M. Sanins, et al. (1990). "Pattern Recognition Analysis of High Resolution 1H NMR Spectra of Urine. A Nonlinear Mapping Approach to the Classification of Toxicological Data." NMR in Biomedicine **3**(4): 166-72.

Gasteiger, J. (2006). "Chemoinformatics: A New Field with a Long Tradition." Analytical and Bioanalytical Chemistry **384**(1): 57-64.

Geladi, P. (1988). "Notes on the History and Nature of Partial Least Squares (PLS) Modelling." Journal of Chemometrics **2**(4): 231-46.

Geladi, P. and K. Esbensen (1990). "The Start and Early History of Chemometrics: Selected Interviews. Part 1." Journal of Chemometrics **4**(5): 337-54.

Genome Canada. (2007). "The Human Metabolome Project." from http://www.metabolomics.ca/.

Genser, B., P. J. Cooper, et al. (2007). "A Guide to Modern Statistical Analysis of Immunological Data." BMC immunology **8**(1): 27.

Gibbon, S. (2007). *Breast Cancer Genes and the Gendering of Knowledge: Science and Citizenship in the Cultural Context of the" New" Genetics*. Basingstoke: Palgrave Macmillan.

Gieryn, T. F. (1983). "Boundary-Work and the Demarcation of Science from Non-Science: Strains and Interests in Professional Ideologies of Scientists." American Sociological Review **48**(6): 781-95.

Gitelman, L. (2013). *Raw Data Is an Oxymoron*. Cambridge: MIT Press.

Glasner, P. (2002). "Beyond the Genome: Reconstituting the New Genetics." New Genetics and Society **21**(3): 267-77.

Good, B. (1994). *Medicine, Rationality, and Experience: An Anthropological Perspective*. Cambridge: Cambridge University Press.

Goodacre, R. (2005). "Metabolomics–The Way Forward." Metabolomics **1**(1): 1-2.

Gougeon, R. D., M. Lucio, et al. (2009). "The Chemodiversity of Wines Can Reveal a Metabologeography Expression of Cooperage Oak Wood." Proceedings of the National Academy of Sciences **106**(23): 9174–79.

Grant, D. M. and R. K. Harris, Eds. (1997). Encyclopedia of Nuclear Magnetic Resonance. London, John Wiley & Sons.

Grasseni, C. (2009). *Skilled Visions: Between Apprenticeship and Standards*. Oxford: Berghahn Books.

Greiffenhagen, C. and W. Sharrock (2011). "Does Mathematics Look Certain in the Front, but Fallible in the Back." Social Studies of Science **41**(6): 839-66.

Gwynne, P. and G. Heebner. (2007). "Drug Discovery and Biotechnology Trends." from http://www.sciencemag.org/site/products/ddbt_0207_Final.xhtml.

Hacking, I. (1990). *The Taming of Chance*. Cambridge: Cambridge University Press.

Hacking, I. (1992). "The Self-Vindication of the Laboratory Sciences." In: *Science as Practice and Culture* by A. Pickering. Chicago: University of Chicago Press, 29-64.

Hacking, I. (1999). *The Social Construction of What?* Cambridge: Harvard University Press.

Hacking, I. (2007). Kinds of People: Moving Targets. Proceedings of the British Academy, London.

Hadolt, B., V. Hörbst, et al. (2012). "Biomedical Techniques in Context: On the Appropriation of Biomedical Procedures and Artifacts." Medical Anthropology **31**(3): 179-95.

Hagen, J. B. (1998). "The Origins of Bioinformatics." Nature Medicine **2**: 231-36.

Hamburg, M. A. and F. S. Collins (2010). "The Path to Personalized Medicine." New England Journal of Medicine **363**(4): 301-04.

Haraway, D. (1988). "Situated Knowledges: The Science Question in Feminism and the Privilege of Partial Perspective." Feminist Studies **14**(3): 575-99.

Haraway, D. (1991a). "A Cyborg Manifesto: Science, Technology, and Socialist-Feminism in the Late 20th Century." In: *Simians, Cyborgs, and Women: The Reinvention of Nature* by D. Haraway. New York: Routledge, 117-58.

Haraway, D. (1991b). *Simians, Cyborgs and Women: The Reinvention of Nature*. New York: Routledge.

Haraway, D. (1997). *Modest−Witness@ Second− Millennium. FemaleMan− Meets− OncoMouse: Feminism and Technoscience*. New York: Routledge.

Haraway, D. J. (1995). "Universal Donors in a Vampire Culture: It's All in the Family: Biological Kinship Categories in the Twentieth-Century United States." In: *Uncommon Ground: Toward Reinventing Nature* by W. Cronon. New York: Norton, 321-22.

Hardy, Q. (24 March 2012). "Just the Facts. Yes, All of Them." New York Times, from http://www.nytimes.com/2012/03/25/business/factuals-gil-elbaz-wants-to-gather-the-data-universe.html?pagewanted=all.

Hedgecoe, A. (2004). *The Politics of Personalised Medicine: Pharmacogenetics in the Clinic*. Cambridge: Cambridge University Press.

Helmreich, S. (2000). *Silicon Second Nature: Culturing Ariticial Life in a Digital World*. Berkeley: University of California Press.

Helmreich, S. (2008). "Species of Biocapital." Science as Culture **17**(4): 463-78.

Helmreich, S. (2009). *Alien Ocean: Anthropological Voyages in Microbial Seas*. Berkeley: University of California Press.

Hinterberger, A. (2012). "Investing in Life, Investing in Difference: Nations, Populations and Genomes." Theory, culture & society **29**(3): 72-93.

Hirschauer, S. (1991). "The Manufacture of Bodies in Surgery." Social Studies of Science **21**(2): 279-319.

Holmberg, C., C. Bischof, et al. (2012). "Making Predictions: Computing Populations." Science, Technology & Human Values **38**(3): 398-420.

Holmes, E., R. L. Loo, et al. (2008a). "Human Metabolic Phenotype Diversity and Its Association with Diet and Blood Pressure." Nature **453**: 396-400.

Holmes, E., J. K. Nicholson, et al. (1992). "Mapping the Biochemical Trajectory of Nephrotoxicity by Pattern Recognition of NMR Urinanalysis." NMR in Biomedicine **5**(6): 368-72.

Holmes, E., I. D. Wilson, et al. (2008b). "Metabolic Phenotyping in Health and Disease." Cell **134**(5): 714-17.

Homer-Vanniasinkam, S. and J. Tsui (2012). "The Continuing Challenges of Translational Research: Clinician-Scientists' Perspective." Cardiology Research and Practice: 1-5.

Hopkins, M. M., P. A. Martin, et al. (2007). "The Myth of the Biotech Revolution: An Assessment of Technological, Clinical and Organisational Change." Research policy **36**(4): 566-89.

Hotz, R. L. (13 August 2012). "Here's an Omical Tale: Scientists Discover Spreading Suffix." The Wall Street Journal, from http://online.wsj.com/article/SB10000872396390444840104577551433143153716.html.

Hunter, P. (2009). "Reading the Metabolic Fine Print." EMBO Reports **10**(1): 20-23.

Imperial College London. (2011a). "Annual Report and Accounts 2010-2011." from https://workspace.imperial.ac.uk/finance/Public/annual_report/annual_report_10_11.pdf.

Imperial College London. (2011b). "Biomolecular Medicine." from http://www1.imperial.ac.uk/surgeryandcancer/divisionofsurgery/biomol_med/.

Imperial College London. (2012a). "Financial Support." from http://www1.imperial.ac.uk/surgeryandcancer/divisionofsurgery/biomol_med/financial/.

Imperial College London. (2012b). "INTERMAP." from http://www1.imperial.ac.uk/publichealth/departments/ebs/projects/cdel/intermap/.

Imperial College London. (2012c). "Nestle-ICL Research Alliance." from http://www1.imperial.ac.uk/surgeryandcancer/divisionofsurgery/biomol_med/key_projects/nestle/.

Imperial College London. (2012d). "NIHR Imperial Biomedical Research Centre,." from https://workspace.imperial.ac.uk/medicine/Public/FoM/A4%20Landscape%2016pp.pdf.

Imperial College London. (2013). "Why Choose Imperial?", from http://www3.imperial.ac.uk/parents/whyimperial.

Imperial NIHR Biomedical Research Centre. (2011). "An 'Intelligent Knife' That Tells the Surgeon Where to Cut." from http://imperialbrc.org/our-impact/case-studies/intelligent-knife-surgery.

Jasanoff, S. (2002). "Science and the Statistical Victim." Social Studies of Science **32**(1): 37-69.

Jasanoff, S. (2004). *States of Knowledge: The Co-Production of Science and Social Order*. Abingdon: Routledge.

Jha, A. (7 October 2010). "Science Funding Cuts: We Won't Fill the Gaps, Say Firms and Charities." The Guardian, from http://m.guardian.co.uk/science/2010/oct/07/science-funding-cuts-firms-charities?cat=science&type=article.

Jha, A. and I. Sample. (15 August 2011). "Chemistry Cuts Will Do 'Irreparable' Damage, Top Scientists Warn." The Guardian, from http://www.guardian.co.uk/science/2011/aug/15/chemistry-funding-cuts-scientists-warn.

Joyce, K. A. (2008). *Magnetic Appeal: MRI and the Myth of Transparency*. Ithaca: Cornell University Press.

Katsnelson, A. (2013). "Momentum Grows to Make 'Personalized' Medicine More 'Precise'." Nature Medicine **19**(243): 249.

Kay, L. E. (2000). *Who Wrote the Book of Life?: A History of the Genetic Code*. Palo Alto: Stanford University Press.

Keating, P. and A. Cambrosio (2003). *Biomedical Platforms: Realigning the Normal and the Pathological in Late-Twentieth-Century Medicine*. Cambridge: MIT Press.

Keating, P. and A. Cambrosio (2012). "Too Many Numbers: Microarrays in Clinical Cancer Research." Studies in History and Philosophy of Biological and Biomedical Sciences **43**(1): 37-51.

Keller, E. F. (2000). "Models of and Models For: Theory and Practice in Contemporary Biology." Philosophy of Science **67**: 72-86.

Keller, E. F. (2002). *The Century of the Gene*. Cambridge: Harvard University Press.

Kinross, J. M., E. Holmes, et al. (2011). "Metabolic Phenotyping for Monitoring Surgical Patients." Lancet **377**(9780): 1817-9.

Kleinman, A. (1982). "Neurasthenia and Depression: A Study of Somatization and Culture in China." Culture, Medicine and Psychiatry **6**(2): 117-90.

Knorr-Cetina, K. (1999). *Epistemic Cultures: How the Sciences Make Knowledge*. Cambridge: Harvard University Press.

Kohli-Laven, N., P. Bourret, et al. (2011). "Cancer Clinical Trials in the Era of Genomic Signatures: Biomedical Innovation, Clinical Utility, and Regulatory-Scientific Hybrids." Social Studies of Science **41**(4): 487-513.

Konrad, M. (2005). *Narrating the New Predictive Genetics: Ethics, Ethnography and Science*. Cambridge University Press Cambridge.

Kowalski, B. and C. Bender (1972). "Pattern Recognition. Powerful Approach to Interpreting Chemical Data." Journal of the American Chemical Society **94**(16): 5632-39.

Landecker, H. (2007). *Culturing Life: How Cells Became Technologies*. Cambridge: Harvard University Press.

Landecker, H. (2011). "Food as Exposure: Nutritional Epigenetics and the New Metabolism." BioSocieties **6**(2): 167-94.

Latour, B. (1987). *Science in Action: How to Follow Scientists and Engineers Through Society*. Cambridge: Harvard University Press.

Latour, B. and S. Woolgar (1986). *Laboratory Life: The Construction of Scientific Facts*. Princeton: Princeton University Press.

Law, J. (2009). "Actor Network Theory and Material Semiotics." In: *The new Blackwell companion to social theory* by B. S. Turner. Chichester: Blackwell, 141-58.

Lee, J. K., P. D. Williams, et al. (2008). "Data Mining in Genomics." Clinics in Laboratory Medicine **28**(1): 145-66.

Leonelli, S. (2012). "When Humans Are the Exception: Cross-Species Databases at the Interface of Biological and Clinical Research." Social Studies of Science **42**(2): 214-36.

Leonelli, S. (forthcoming). "Why the Current Insistence on Open Access to Scientific Data? Big Data, Knowledge Production and the Political Economy of Contemporary Biology." Bulletin of Science and Technology Studies.

Leonelli, S. (in press). "Integrating Data to Acquire New Knowledge: Three Modes of Integration in Plant Science." Studies in History and Philosophy of Science Part C.

Lévi-Strauss, C. (1962). *The Savage Mind*. Chicago: University of Chicago Press.

Lindon, J. C. (2010). "Metabonomics History." from http://www1.imperial.ac.uk/surgeryandcancer/divisionofsurgery/biomol_med/education/metabonomics_history/.

Lindon, J. C., E. Holmes, et al. (2003). "Peer Reviewed: So What's the Deal with Metabonomics?" Analytical Chemistry **75**(17): 384-91.

Lindon, J. C., H. C. Keun, et al. (2005). "The Consortium for Metabonomic Toxicology (COMET): Aims, Activities and Achievements." Pharmacogenomics **6**(7): 691-99.

Lindon, J. C. and J. K. Nicholson (2008). "Spectroscopic and Statistical Techniques for Information Recovery in Metabonomics and Metabolomics." Annual Review of Analytical Chemistry **1**: 45-69.

Lindon, J. C., J. K. Nicholson, et al. (2007). *The Handbook of Metabonomics and Metabolomics*. Radarweg: Elsevier.

Lippman, A. (1992). "Led (Astray) by Genetic Maps: The Cartography of the Human Genome and Health Care." Social Science & Medicine **35**(12): 1469-76.

Lock, M. (2011). "Dementia Entanglements in a Postgenomic Era." Science, Technology & Human Values **36**(5): 685-703.

Lock, M. (2012). "The Epigenome and Nature/Nurture Reunification: A Challenge for Anthropology." Medical Anthropology **32**(4): 291-308.

Lock, M., J. Freeman, et al. (2007). "Susceptibility Genes and the Question of Embodied Identity." Medical Anthropology Quarterly **21**(3): 256-76.

Lock, M. and P. A. Kaufert (1998). *Pragmatic Women and Body Politics*. Cambridge: Cambridge University Press.

Lock, M. and S. Lindenbaum (1993). *Knowledge, Power and Practice: The Anthropology of Medicine and Everyday Life*. Berkeley: University of California Press.

Lock, M. and V. K. Nguyen (2010a). "The Normal Body." In: *An Anthropology of Biomedicine* by M. Lock and V. K. Nguyen. Chichester: Blackwell Publishing, 32-56.

Lock, M. and N. Vinh-Kim (2010b). *An Anthropology of Biomedicine*. Chichester: Wiley-Blackwell.

Lock, M. M. (1995). *Encounters with Aging: Mythologies of Menopause in Japan and North America*. London: University of California Press.

Lock, M. M., A. Young, et al. (2000). *Living and Working with the New Medical Technologies: Intersections of Inquiry*. Cambridge: Cambridge University Press.

Lohr, S. (19 June 2013). "Sizing Up Big Data, Broadening Beyond the Internet." New York Times, from http://bits.blogs.nytimes.com/2013/06/19/sizing-up-big-data-broadening-beyond-the-internet/.

Löwy, I. (1990). "The Strength of Loose Concepts: Boundary Concepts, Federative Experimental Strategies and Disciplinary Growth: The Case of Immunology." History of Science **30**(90): 371-96.

Löwy, I. (1996). *Between Bench and Bedside: Science, Healing, and Interleukin-2 in a Cancer Ward*. Cambridge: Harvard University Press.

Löwy, I. (2009). *Preventive Strikes: Women, Precancer, and Prophylactic Surgery*. Baltimore: Johns Hopkins University Press.

Lynch, M. and S. Woolgar (1990). *Representation in Scientific Practice*. Cambridge: MIT Press.

M'charek, A. (2005). *The Human Genome Diversity Project: An Ethnography of Scientific Practice*. Cambridge: Cambridge University Press.

Mackenzie, A. (2003). "Bringing Sequences to Life: How Bioinformatics Corporealizes Sequence Data." New Genetics and Society **22**: 315-32.

Malerba, F., R. Nelson, et al. (1999). "'History-Friendly' Models of Industry Evolution: The Computer Industry." Industrial and Corporate Change **8**(1): 3-40.

Mansell, P. (14 October 2010). "UK Launches Stratified Medicines Innovation Platform." Pharma Times, from http://www.pharmatimes.com/article/10-10-14/UK_launches_Stratified_Medicines_Innovation_Platform.aspx.

Marcus, G. E. (1995). "Ethnography in/of the World System: The Emergence of Multi-Sited Ethnography." Annual Review of Anthropology **24**: 95-117.

Martin, E. (1995). *Flexible Bodies: Tracking Immunity in American Culture from the Days of Polio to the Age of AIDS*. Boston: Beacon Press.

McGoey, L. (2009). "Pharmaceutical Controversies and the Performative Value of Uncertainty." Science as Culture **18**(2): 151-64.

Metabolomics Society. (2013). "Databases." from http://www.metabolomicssociety.org/databases.

Metzler, I. (2010). "Biomarkers and Their Consequences for the Biomedical Profession: A Social Science Perspective." Personalized Medicine: Future Medicine **7**(4): 407-20.

Mirnezami, R., J. Nicholson, et al. (2012). "Preparing for Precision Medicine." New England Journal of Medicine **366**(6): 489-91.

Mitchell, R. and C. Waldby (2010). "National Biobanks: Clinical Labor, Risk Production, and the Creation of Biovalue." Science, Technology & Human Values **35**(3): 330-55.

Mol, A. (1998). "Lived Reality and the Multiplicity of Norms: A Critical Tribute to George Canguilhem." Economy and Society **27**(2-3): 274-84.

Mol, A. (2002). *The Body Multiple: Artherosclerosis in Practice*. Durham: Duke University Press.

Mol, A. and J. Law (2002). *Complexities*. Durham: Duke University Press.

Moody, G. (2004). *Digital Code of Life: How Bioinformatics Is Revolutionizing Science, Medicine, and Business*. Hoboken: John Wiley & Sons.

Moreira, T. (2006). "Heterogeneity and Coordination of Blood Pressure in Neurosurgery." Social Studies of Science **36**(1): 69-97.

Moreira, T. and P. Palladino (2011). "'Population Laboratories' or 'Laboratory Populations? Making Sense of the Baltimore Longitudinal Study of Aging, 1965–1987." Studies in History and Philosophy of Science Part C **42**(3): 317-27.

Morgan, M., C. A. Barry, et al. (2011). "Implementing 'Translational' Biomedical Research: Convergence and Divergence among Clinical and Basic Scientists." Social Science & Medicine **73**(7): 945-52.

Moser, I. (2011). "Dementia and the Limits to Life: Anthropological Sensibilities, STS Interferences, and Possibilities for Action in Care." Science, Technology & Human Values **36**(5): 704-22.

Mukhopadhyay, R. (2013). "Q&A with Jeremy Nicholson." American Society for Biochemistry and Molecular Biology Today, from http://www.asbmb.org/asbmbtoday/asbmbtoday_article.aspx?id=24036.

Myers, N. (2006). "Animating Mechanism: Animations and the Propagation of Affect in the Lively Arts of Protein Modelling." Science Studies **19**(2): 6-30.

Myers, N. (2008). "Molecular Embodiments and the Body-Work of Modeling in Protein Crystallography." Social Studies of Science **38**(2): 163-99.

Nader, L. (1972). "Up the Anthropologist: Perspectives Gained from Studying Up." Reinventing Anthropology **1972**: 284-311.

National Institute for Health Research. (2012). "Biomedical Research Centres." from http://www.nihr.ac.uk/infrastructure/Pages/infrastructure_biomedical_research_centres.aspx.

National Institute of Health. (19 September 2012). "NIH Announces New Program in Metabolomics." from http://www.nih.gov/news/health/sep2012/od-19.htm.

Natural Environment Research Council. (2006). "What Is Genomics and Post Genomics?", from http://www.nerc.ac.uk/research/programmes/proteomics/background/whatis.asp?cookieConsent=A.

Navon, D. (2011). "Genomic Designation: How Genetics Can Delineate New, Phenotypically Diffuse Medical Categories." Social Studies of Science **41**(2): 203-26.

Nelkin, D. and M. Lindee (2004). *The DNA Mystique: The Gene As a Cultural Icon*. Ann Arbor: University of Michigan Press.

Nestle. (2013). "Metabonomics - What Metabolic Profile are You?", from http://www.nestle.com/randd/innovations/allinnovations/metabonomics.

Nicholson, J. K. (2006). "Global Systems Biology, Personalized Medicine and Molecular Epidemiology." Molecular Systems Biology **2**(1): 1-6.

Nicholson, J. K., E. Holmes, et al. (2012a). "Host-Gut Microbiota Metabolic Interactions." Science **336**(6086): 1262-67.

Nicholson, J. K., E. Holmes, et al. (2012b). "Metabolic Phenotyping in Clinical and Surgical Environments." Nature **491**(7424): 384-92.

Nicholson, J. K., E. Holmes, et al. (2005). "Gut Microorganisms, Mammalian Metabolism and Personalized Health Care." Nature Reviews Microbiology **3**(5): 431-38.

Nicholson, J. K. and J. C. Lindon (2008). "Systems Biology: Metabonomics." Nature **455**(7216): 1054-56.

Nicholson, J. K., J. C. Lindon, et al. (1999). "Metabonomics: Understanding the Metabolic Responses of Living Systems to Pathophysiological Stimuli Via Multivariate Statistical Analysis of Biological NMR Spectroscopic Data." Xenobiotica **29**(11): 1181-89.

Nicholson, J. K. and I. D. Wilson (2003). "Understanding 'Global' Systems Biology: Metabonomics and the Continuum of Metabolism." Nature Reviews Drug Discovery **2**(8): 668-76.

Niewöhner, J. (2011). "Epigenetics: Embedded Bodies and the Molecularisation of Biography and Milieu." BioSocieties **6**(3): 279-98.

Nightingale, P. (2000). "Economies of Scale in Experimentation: Knowledge and Technology in Pharmaceutical R&D." Industrial and Corporate Change **9**(2): 315-59.

November, J. A. (2012). *Biomedical Computing: Digitizing Life in the United States*. Baltimore: Johns Hopkins University Press.

O'Connell, T. and P. Watkins (2010). "The Application of Metabonomics to Predict Drug-Induced Liver Injury." Clinical Pharmacology & Therapeutics **88**(3): 394-99.

O'Malley, M., A. Powell, et al. (2007a). "Knowledge-Making Distinctions in Synthetic Biology." BioEssays **30**(1): 57-65.

O'Malley, M. A., J. Calvert, et al. (2007b). "The Study of Socioethical Issues in Systems Biology." American Journal of Bioethics **7**(4): 67-78.

O'Connell, D. and D. Roblin (2006). "Translational Research in the Pharmaceutical Industry: From Bench to Bedside." Drug Discovery Today **11**(17–18): 833-38.

Oliver, S., M. Winson, et al. (1998). Systematic Functional Analysis of the Yeast Genome. European Symposium of Life Sciences Research in Space (Oser).

Olson, V. A. (2010). "The Ecobiopolitics of Space Biomedicine." Medical Anthropology **29**(2): 170-93.

Oudshoorn, N. (1990). "On the Making of Sex Hormones: Research Materials and the Production of Knowledge." Social Studies of Science **20**(1): 5-33.

Paddock, C. (22 February 2013). "Astronomy Algorithms Help Diagnose Aggressive Tumors." from http://www.medicalnewstoday.com/articles/256739.php.

Parkin, D. J. and S. J. Ulijaszek (2007). *Holistic Anthropology: Emergence and Convergence*. Oxford: Berghahn Books.

Patterson, D. (5 December 2011). "Computer Scientists May Have What It Takes to Help Cure Cancer." New York Times, from http://www.nytimes.com/2011/12/06/science/david-patterson-enlist-computer-scientists-in-cancer-fight.html.

Patti, G. J., O. Yanes, et al. (2012). "Metabolomics: The Apogee of the Omics Trilogy." Nature Reviews Molecular Cell Biology **13**: 263-69.

Pearson, H. (2007). "Meet the Human Metabolome." Nature **446**(7131): 8-8.

Penders, B., K. Horstman, et al. (2008). "Walking the Line Between Lab and Computation: The "Moist" Zone." BioScience **58**(8): 747-55.

Petryna, A. (2002). *Life Exposed: Biological Citizens After Chernobyl*. Princeton: Princeton University Press.

Pfaffenberger, B. (1988). "The Social Meaning of the Personal Computer: Or, Why the Personal Computer Revolution Was No Revolution." Anthropological Quarterly: 39-47.

Pickering, A. (1992). *Science as Practice and Culture*. Chicago: University of Chicago Press.

Pollack, A. (1 December 2011). "DNA Sequencing Caught in Deluge of Data." New York Times, from http://www.nytimes.com/2011/12/01/business/dna-sequencing-caught-in-deluge-of-data.html?ref=health.

Pollack, A. (14 June 2010). "The Genome at 10." New York Times, from http://www.nytimes.com/2010/06/15/business/15genome.html.

Porter, T. M. (1988). *The Rise of Statistical Thinking, 1820-1900*. Princeton: Princeton University Press.

Prentice, R. (2005). "The Anatomy of a Surgical Simulation: The Mutual Articulation of Bodies in and through the Machine." Social Studies of Science **35**(6): 837-66.

Rabinow, P. (1996). *Making PCR: A Story of Biotechnology*. Chicago: University of Chicago Press.

Rajan, K. S. (2005). "Subjects of Speculation: Emergent Life Sciences and Market Logics in the United States and India." American Anthropologist **107**(1): 19-30.

Rajan, K. S. (2006). *Biocapital: The Constitution of Postgenomic Life*. Durham: Duke University Press Books.

Rajan, K. S. and S. Leonelli (forthcoming). "Biomedical Trans-actions " Public Culture.

Raman, S. and R. Tutton (2010). "Life, Science, and Biopower." Science, Technology & Human Values **35**(5): 711-34.

Rapp, R. (1999). *Testing Women, Testing the Fetus: The Social Impact of Amniocentesis in America*. New York: Routledge.

Räsänen, M. and J. M. Nyce (2013). "The Raw is Cooked: Data in Intelligence Practice." Science, Technology & Human Values.

Reardon, J. (2011). "The 'Persons' and 'Genomics' of Personal Genomics." Personalized Medicine **8**(1): 95-107.

Rheinberger, H. J. (1997). *Toward a History of Epistemic Things: Synthesizing Proteins in the Test Tube*. Palo Alto: Stanford University Press.

Rose, N. (2001). "The Politics of Life Itself." Theory, culture & society **18**(6): 1-30.

Rose, N. (2007). *The Politics of Life Itself. Biomedicine, Power and Subjectivity in the 21st Century*. Princeton: Princeton University Press.

Rose, N. (2009). "Normality and Pathology in a Biomedical Age." The Sociological Review **57**: 66-83.

Rose, N. (2013). "The Human Sciences in a Biological Age." Theory, culture & society **30**(1): 3-34.

Saini, A. (2012). "London's Olympic Drug Testing Lab to Become National Phenome Center." Science **3**(337): 6094.

Scheper-Hughes, N. and M. M. Lock (1987). "The Mindful Body: A Prolegomenon to Future Work in Medical Anthropology." Medical Anthropology Quarterly **1**(1): 6-41.

Science and Technology Committee (2009). Science and Technology Committee - Second Report on Genomic Medicine, House of Lords.

Scowcroft, H. (2011). "Our Stratified Medicine Programme – What Is It and How Will It Work?", from http://scienceblog.cancerresearchuk.org/2011/11/21/our-stratified-medicine-programme-what-is-it-and-how-will-it-work/.

Seeley, E. H. and R. M. Caprioli (2008). "Molecular Imaging of Proteins in Tissues by Mass Spectrometry." Proceedings of the National Academy of Sciences **105**(47): 18126-31.

Seising, R. (2008). "On the Absence of Strict Boundaries—Vagueness, Haziness, and Fuzziness in Philosophy, Science, and Medicine." Applied soft computing **8**(3): 1232-42.

Shapin, S. and S. Schaffer (1985). *Leviathan and the Air-Pump: Hobbes, Boyle, and the Experimental Life*. Princeton: Princeton University Press.

Silverstein, A. M. (1989). *A History of Immunology*. London: Academic Press.

Singer, M. (1990). "Reinventing Medical Anthropology: Toward a Critical Realignment." Social Science & Medicine **30**(2): 179-87.

Skogerson, K., R. Runnebaum, et al. (2009). "Comparison of Gas Chromatography-Coupled Time-of-Flight Mass Spectrometry and 1H Nuclear Magnetic Resonance Spectroscopy Metabolite Identification in White Wines from a Sensory Study Investigating Wine Body." Journal of Agricultural and Food Chemistry **57**(15): 6899-907.

Small Things Considered: The Microbe Blog. (2009). "Of Terms in Biology: Metabolomics and Metabonomics." from http://schaechter.asmblog.org/.a/6a00d8341c5e1453ef0115709c8e27970b-300wi.

Sommerlund, J. (2006). "Classifying Microorganisms: The Multiplicity of Classifications and Research Practices in Molecular Microbial Ecology." Social Studies of Science **36**(6): 909-28.

Star, S. L. and J. R. Griesemer (1989). "Institutional Ecology, Translations and Boundary Objects: Amateurs and Professionals in Berkeley's Museum of Vertebrate Zoology, 1907-39." Social Studies of Science **19**(3): -420.

Stevens, H. (2011). "On the Means of Bio-Production: Bioinformatics and How to Make Knowledge in a High-Throughput Genomics Laboratory." BioSocieties **6**(2): 217-42.

Stoeckli, M., P. Chaurand, et al. (2001). "Imaging Mass Spectrometry: A New Technology for the Analysis of Protein Expression in Mammalian Tissues." Nature Medicine **7**(4): 493-96.

Strathern, M. (1992a). *After Nature: English Kinship in the Late Twentieth Century*. Cambridge: Cambridge University Press.

Strathern, M. (1992b). *Reproducing the Future: Essays on Anthropology, Kinship and the New Reproductive Technologies*. Manchester: Manchester University Press.

Strathern, M. (1996). "Cutting the Network." The Journal of the Royal Anthropological Institute **2**(3): 517-35.

Street, A. (2011). "Artefacts of Not-Knowing: The Medical Record, the Diagnosis and the Production of Uncertainty in Papua New Guinean Biomedicine." Social Studies of Science **41**(6): 815-34.

Swedlow, J. R., G. Zanetti, et al. (2011). "Channeling the Data Deluge." Nature Methods **8**(6): 463-65.

Tape, T. G. (2013). "The Area Under an ROC Curve." from http://gim.unmc.edu/dxtests/Default.htm.

Thacker, E. (2005). *The Global Genome: Biotechnology, Politics, and Culture*. Cambridge: MIT Press.

The Economist (27 February 2010). Data, Data Everywhere: A Special Report on Managing Information.

The Wellcome Trust. (2013). "History of Henry Wellcome." from http://www.wellcome.ac.uk/About-us/History/index.htm.

Throsby, K. and C. Roberts (2010). "Getting Bigger: Children's Bodies, Genes and Environments." The Sociological Review **58**(s1): 73-92.

Times Higher Education. (2012). "World University Rankings 2012-2013." from http://www.timeshighereducation.co.uk/world-university-rankings/2012-13/world-ranking.

Timmermans, S. and M. Berg (1997). "Standardization in Action: Achieving Local Universality through Medical Protocols." Social Studies of Science **27**(2): 273-305.

Titz, A. (2006). "The Borderline between Medicinal Products and Food Supplements." In: *Responsibilities in the Efficient Use of Medicinal Products* by J. Valverde. Amsterdam: IOS Press, 37-50.

Traweek, S. (1988). *Life Times and Beamtimes. The World of High Energy Physicists*. Cambridge: Harvard University Press.

Tripathy, B. (2012). *RSSDI: Textbook of Diabetes Mellitus*. New Delhi: Jaypee Brothers Medical Publishers.

Trusheim, M. R., E. R. Berndt, et al. (2007). "Stratified Medicine: Strategic and Economic Implications of Combining Drugs and Clinical Biomarkers." Nature Reviews Drug Discovery **6**(4): 287-93.

Tursz, T., F. Andre, et al. (2011). "Implications of Personalized Medicine—Perspective from a Cancer Center." Nature Reviews Clinical Oncology **8**(3): 177-83.

Tutton, R. (2012). "Personalizing Medicine: Futures Present and Past." Social Science & Medicine **75**: 1721-28.

Tutton, R. and B. Prainsack (2011). "Enterprising or Altruistic Selves? Making up Research Subjects in Genetics Research." Sociology of Health & Illness **33**(7): 1081-95.

U.S. Food and Drug Administration. (2013). "Table of Pharmacogenomic Biomarkers in Drug Labels." from http://www.fda.gov/drugs/scienceresearch/researchareas/pharmacogenetics/ucm083378.htm.

UK Department of Health. (1 August 2012). "A Phenomenal Legacy for London 2012." from http://mediacentre.dh.gov.uk/2012/08/01/a-phenomenal-legacy-for-london-2012/.

Umetrics. (2012). "Methods: History." from http://www.umetrics.com/default.aspx?id=7943.

University College London. (2013). "Facts and Figures." from http://www.ucl.ac.uk/slms/about-us/facts-figures.

Varmuza, K. and P. Filzmoser (2009). *Introduction to Multivariate Statistical Analysis in Chemometrics*. Boca Raton: CRC Press.

Venture Navigator. (2007). "A Recent History of the Pharmaceutical Industry." from http://www.venturenavigator.co.uk/content/154.

Wainwright, S. P., C. Williams, et al. (2006). "From Bench to Bedside? Biomedical Scientists' Expectations of Stem Cell Science as a Future Therapy for Diabetes." Social Science & Medicine **63**(8): 2052-64.

Waters Corporation. (26 September 2006). "Waters Corporation and Imperial College London Expand Systems Biology Collaboration." from http://www.waters.com/waters/newsDetail.htm?id=10001312&locale=en_GB.

Wenner, M. (17 June 2008). "Jeremy Nicholson's Gut Instincts: Researching Intestinal Bacteria." Scientific American, from http://www.scientificamerican.com/article.cfm?id=jeremy-nicholsons-gut-instincts.

Whitmarsh, I. (2011). "American Genomics in Barbados: Race, Illness, and Pleasure in the Science of Personalized Medicine." Body & Society **17**(2-3): 159-81.

Whyte, S. R. (1998). *Questioning Misfortune: The Pragmatics of Uncertainty in Eastern Uganda*. Cambridge: Cambridge University Press.

Wiesner, R., E. Edwards, et al. (2003). "Model for End-Stage Liver Disease (MELD) and Allocation of Donor Livers." Gastroenterology **124**(1): 91-96.

Wikipedia. (2013). "List of UK Universities by Endowment." from http://en.wikipedia.org/wiki/List_of_UK_universities_by_endowment.

Will, C. M. (2007). "The Alchemy of Clinical Trials." BioSocieties **2**(1): 85-99.

Wilson-Kovacs, D. M. and C. Hauskeller (2011). "The Clinician-Scientist: Professional Dynamics in Clinical Stem Cell Research." Sociology of Health & Illness **34**(4): 497-512.

Wistuba, I. I., J. G. Gelovani, et al. (2011). "Methodological and Practical Challenges for Personalized Cancer Therapies." Nature Reviews Clinical Oncology **8**(3): 135-41.

Wong, S. (14 November 2012). "New Centre Heralds Age of Precision Medicine." Imperial College London, from http://www3.imperial.ac.uk/newsandeventspggrp/imperialcollege/newssummary/news_14-11-2012-17-37-6.

Wong, S. (18 May 2011). "“Molecular Fingerprinting” Will Improve Monitoring of Surgical Patients, Experts Say." Imperial College London, from http://www3.imperial.ac.uk/newsandeventspggrp/imperialcollege/newssummary/news_18-5-2011-10-41-30.

Woolgar, S. and J. Lezaun (2013). "The Wrong Bin Bag: A Turn to Ontology in Science and Technology Studies?" Social Studies of Science **43**(3): 321-40.

Wynne, B. (2005). "Reflexing Complexity." Theory, culture & society **22**(5): 67-94.

Xu, J. and A. Hagler (2002). "Chemoinformatics and Drug Discovery." Molecules **7**(8): 566-600.

Appendix: List of Technical Terminology

Bioinformatics

A set of practices for handling, manipulating, and storing biological data, and ultimately for transforming data into knowledge. Also referred to as **computational biology**, such practices rely on the use of computers and statistics, and have become central to biology following the Human Genome Project (HGP) of the 1990s.

Biomarker

Measurable and quantifiable biological entities that can be statistically correlated with health and disease. Biomarkers can technically correspond to any physiological or anatomical measurement—for example blood cholesterol or body mass index (BMI)—but in the post-genomic sciences they largely refer to proteins, genes, or other small molecules.

Chemometrics

A field that entails the application of mathematical and statistical methods to complex chemical data. Though the field dates back to the 1970s, chemometric techniques provide the overarching framework for statistical data analysis, and are central to metabonomics as a field.

MALDI-MSI

The abbreviation for matrix-assisted laser desorption/ionization mass spectroscopy imaging, a type of mass spectroscopy (MS) that is used to carry out the molecular imaging of solid tissues. Originally developed within the field of proteomics, MALDI-MSI uses a laser to irradiate molecular locations—termed “pixels”—within prepared tissue slices. Each pixel

that is irradiated by laser gives rise to a mass spectrum, which details the molecular composition of a particular molecular location. In combination, the information generated by MALDI-MSI can be used to generate a two dimensional map of the chemical patterns within a tissue slice, or can be used to analyze the characteristics of specific pixels or metabolites.

MAS-NMR

The acronym for magic angle spinning nuclear magnetic resonance, a type of nuclear magnetic resonance (NMR) that is used to determine the biochemical composition of solid tissues. "Magic angle spinning" refers to the fact that the solid samples are spun via an air turbine mechanism within the NMR spectrometer, such that an average set of chemical shift values are obtained for each solid sample regardless of its molecular homogeneity.

MELD

The abbreviation for Model of End-Stage Liver Disease, a clinical algorithm that is used to assess patients with "acute-on-chronic" liver failure. MELD uses linear regression on a combination of clinical measures—serum bilirubin, creatinine, and prothrombin time—to calculate a patient's probability of dying over a certain period of time without a liver transplant.

Metabolome

The sum of the metabolites within an organism, which are reflective of the metabolism of and biochemical reactions occurring within an organism.

Metabonomics

The post-genomic field of research on metabolism. The field dates back to the 1999, though it entails a combination of practices, technologies, and ideas that emerged long before that in the fields of biochemistry, drug discovery, statistics, and physics. The official definition of metabonomics, as provided by the 1999 *Xenobiotica* article that inaugurated the field, reads: "The quantitative measurement of the dynamic multiparametric metabolic response of living systems to pathophysiological stimuli or genetic modification." In practice, metabonomics is synonymous with the field of **metabolomics**, though metabolomics originated from a different set of practices, people, and historical circumstances.

Multivariate statistics

A type of statistical analysis that involves the observation and analysis of many variables simultaneously, often within large data sets. Common forms of multivariate statistics that are used within metabonomics include principal components analysis (PCA) and partial least squares regression (PLS)—techniques for dealing with data that date back to the early 20th century, though other more recent forms include self-organizing maps (SOMs) and manifold learning.

Importantly, within metabonomics multivariate statistics are seen as both superior to and also ontologically different from **univariate statistics**, which measure single variables at a time. The distinction between the two types of statistical analysis, however, is not so clear cut, as the history of both is deeply interwoven with post World War II developments in mathematical techniques and computing.

MS

The abbreviation for mass spectrometry, a technology that enables the determination of the molecular composition of biological samples through ionization. MS measures the mass-to-charge ratios of the charged molecules or molecular fragments that result from the ionization of chemical compounds. The scientific principles underlying MS were developed in the late 19th century by scientists observing the properties of gas discharges, which separated into cathode (negatively charged) and anode (positively charged) rays. Soon thereafter it was discovered that the application of electric and magnetic fields to anode rays could separate the molecules within the gas discharge according to their mass-to-charge ratio.

MWAS

The abbreviation for metabolome wide association studies, which entail the application of metabonomics techniques to large-scale epidemiological studies. In MWAS studies, researchers examine the correlations between the metabolic characteristics of populations and epidemiological risk factors. The concept of MWAS emerged in the late 2000s through work on the International Collaborative Study of Macronutrient, Micronutrients and Blood Pressure (INTERMAP), which examined the urine of 4630 people from four different populations—in the United States, United Kingdom, Japan, and China—to find correlations between lifestyle factors and blood pressure.

NIHR-BRC

The abbreviation for the National Institute for Health Research (NIHR) Biomedical Research Council (BRC) in the United Kingdom. BRCs are based within outstanding NHS and University partnerships, and are established to drive progress on innovation and translational research in biomedicine into NHS practice. The first round of BRCs were established by the NIHR in 2007, followed by a second round in 2011, which established BRCs at the University of Cambridge, University College London, King's College London, Imperial College London, Newcastle University, University of Oxford, Institute of Cancer Research, and University of Southampton.

NMR

The abbreviation for nuclear magnetic resonance spectroscopy, a technology that exploits the magnetic properties of atomic nuclei to determine the chemical and physical properties of molecules. The term "nuclear magnetic resonance" refers to the fact that the nuclei of atoms absorb and re-emit energy at various frequencies in response to magnetic fields of various strengths. By detecting the unique magnetic signatures of atoms and compounds, NMR allows researchers to identify the chemical structure—and also confirm the identity—of unknown biochemicals. Though NMR was developed from experimentation on radars and radio frequency during World War II, it was first applied to the study of biological samples in the 1950s. Beyond its use in metabonomics, NMR forms the basis for the technology of magnetic resonance imaging (MRI), which is widely used for clinical imaging.

PCA and PLS

The abbreviations for principal components analysis and partial least squares regression, respectively. Both PCA and PLS are common types of multivariate statistical analysis. PCA, which was invented in 1901 by Karl Pearson, is an unsupervised statistical technique that converts a set of observations (data) into a series of "principal components," a set of linearly uncorrelated variables, in which the first principal component represents the direction of largest possible variance in the data. PLS, which was invented in the mid-20th century by the econometrician Herman Wold, is a supervised statistical technique. While it is similar to PCA, PLS finds the greatest co-variance within data, by dividing the data into two matrices, X and Y, made of independent and dependent variables.

Personalized Medicine

A paradigm stating that medical diagnoses and treatments can and should be tailored to the unique, individual biology of each patient. The implementation of personalized medicine relies on the use of molecular technologies to find the right treatment for the right patient at the right time. Similar to personalized medicine, **stratified medicine** is seen as the practical implementation of personalized medicine, in which diagnoses and treatments are tailored to molecularly-similar sub-populations of diseases or patients.

Pharmacometabonomics

The use of metabonomics techniques to predict drug toxicity and response in individuals. In pharmacometabonomics experiments, measurements of initial metabolic phenotypes are used to mathematically model responses to subsequent drug therapies and interventions. The concept was first developed through studies of liver toxicity in rats in the mid-2000s in collaboration with researchers from the pharmaceutical company Pfizer. It was later applied to humans to study the adverse effects of the drug acetaminophen on the liver.

Post-Genomic

The study of the effects and mechanisms of gene regulation and expression. Post-genomic fields seek not to identify the structure or sequences of genes, but rather to study how they have higher biological meanings and functions. The best known post-genomic research include epigen(etics/omics), transcriptomics, metabo(lomics/nomics), and lipidomics, but also encompass broader fields such as systems biology.