Smoking and cause-specific mortality in large prospective studies from Cuba, Mexico, and the United States



Blake Thomson, MPhil

Balliol College, University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Clinical Trial Service Unit and Epidemiological Studies Unit,

Nuffield Department of Population Health

Michaelmas 2019

“Death in old age is inevitable, but death before old age is not.” – Sir Richard Doll

For mom,

Whose compassion inspired my passion for global health

And whose death at age 59 years, well before old age, has given my life a resolute and single-minded purpose: to prevent death and disability before old age

Acknowledgements

I am grateful to the participants of the Cuba Prospective Study, the Mexico City Prospective Study, and the US National Health Interview Survey, who contributed their time and personal information for the advancement of science. I would also like to thank the teams that directed these studies and gathered the data, with special thanks to Dr Alfredo Dueñas Herrera and Dr Nurys Armas Rojas of the Cuba Prospective Study, and Dr Roberto Tapia-Conyer, Dr Jesus Alegre-Díaz, and Dr Pablo Kuri-Morales of the Mexico City Prospective Study. The complementary efforts of Prof Sir Richard Peto, Prof Rory Collins, Prof Jonathan Emberson, Prof Sarah Lewington, and Dr Gary Whitlock from Oxford have set the standard for equitable and rigorous scientific collaborations that transcend borders. Your work has inspired me, and surely countless others.

I give my profound thanks to my supervisors, who have conducted active follow-up over this 3-year prospective study towards the endpoint of DPhil (crossing their fingers that follow-up would not be censored due to migration). Thank you for everything, Jon (Emberson), Sarah (Lewington), Ben (Lacey), and (Sir) Richard (Peto). I feel compelled to point out that the prevalence of Professorship among my supervisory team was 75% at the end of the study period, compared to 25% at the beginning of the study period. Having spent three years thinking about causal relationships, however, I must conclude that my role in their meteoric rises in academia was moderate at best. That being said, at the time of writing, only one-fourth of my supervisory team had been recognised in the Queen's Birthday Honours List. Jon, Sarah, and Ben: no pressure, take your time.

I am indebted (figuratively!) to those who have funded the education culminating in this qualification: the Nuffield Department of Population Health (DPhil), the Gates Cambridge

Trust (MPhil), the Flinn Foundation (BA), and the personal endowment of David A. Thomson and Mary S. Finnerty (ABC's, counting toes, etc.). Thank you also to Prof Emmanuela Gakidou, who hired me for my first quantitative global health position.

Thank you to those in the department who gave critical input on my DPhil research, with particular thanks to Prof Sarah Parish, Prof Robert Clarke, and Prof Tim Key for their invaluable comments as the thesis took shape. Thank you to the broader Mexico City Prospective Study team – Diego, Louisa, Rachel, Raúl, Will – for supporting my research and for creating this incredible scientific resource. Thank you for the memories in Mexico City as well, with the notable exception of the trip to Applebee's. Thank you, also, to the broader Cuba Prospective Study team – Julie and Paul in particular – for your statistical support (and the occasional brownie). I must also recognise the students who supported me throughout this process, with particular thanks to Debbie, Diego, Filipa, Jess, Lauren, Peter, Shawon, and of course, the inimitable Katherine, who at many points made bad days good and good days hilarious. And thank you, Richard, for inspiring me far more than you could possibly imagine. I said at your retirement symposium that I planned on spending my 110th birthday, in the year 2100, showing how many tobacco deaths were avoided in the 21st century thanks to your foundational work. I wasn't kidding.

Finally, my deepest thanks goes to my family of lawyers – Mom, Dad, Hunter, and Brenda – who taught me the importance of being able to argue either side of a case objectively (and to use those powers for good). Thank you to Aunt Jane, Uncle Tom, and my extended family for your love and support. Roxanna, you are included here as family despite being, by my father's rough calculation, approximately one-half Thomson as of August 2019 (which may already exceed the recommended amount). In an era of countless distractions, thank you for keeping me focused on the things that matter.

Publications and presentations

Work towards the completion of this thesis has resulted in the following publications and presentations:

Published

Armas Rojas N*, Lacey B*, Lewington S*, Varona Pérez P, Burrett JA, Morales Rigau JM, Sherliker P, Boreham J, Hernández López OJ, **Thomson B**, Achiong Estupiñan F, Díaz González M, Rosquete Muñoz N, Cendra Asencio M, Emberson J, Peto R, Dueñas Herrera A. Cohort Profile: the Cuba Prospective Study. *Int J Epidemiol.* 2019; published online ahead of print (doi: 10.1093/ije/dyy297).
*Joint first authors

Submitted or in preparation

Thomson B*, Armas-Rojas N*, Lacey B*, Burrett J, Varona-Pérez P, Calderón Martínez M, Lorenzo-Vázquez E, Bess Constantén S, Manuel Morales Rigau J, Jesús Hernández López O, Ángel Martínez Morales M, Alonso Alomá I, Achiong Estupiñan F, Díaz González M, Rosquete Muñoz N, Cendra Asencio M, Emberson J, Peto R†, Lewington S†, Dueñas Herrera A†. Childhood smoking and adult mortality in Cuba: prospective study of 150,000 men and women. (Submitted)
*Joint first authors; †Joint senior authors

Thomson B*, Alegre-Díaz J*, Wade R, Ramirez R, Lacey B, Herrington W, López Cervantes M, Gnatiuc L, Solano M, Peto R, Collins R, Tapia-Conyer R, Kuri-Morales P, Lewington S, Emberson JR. Smoking and cause-specific mortality among 150,000 Mexican adults followed for 17 years. (In preparation)
*Joint first authors

Presentations (selected)

Smoking and premature mortality: 17-year prospective study of 120,000 Cuban adults. Presented at the Nuffield Department of Population Health (NDPH) Symposium 2019. March 2019.

Awarded 2nd prize at the annual NDPH poster presentation competition

Smoking kills, but quitting works: Continuing the tale of two Richards. Presented at the "Richard Peto: Studying the Bleeding Obvious" Symposium, Oxford, UK. June 2018.

Smoking and cause-specific mortality in the Cuba Prospective Study of 140,000 adults. Presented at The 30th Latin American Congress of Cardiology and 9th Cuban Congress of Cardiology, Havana, Cuba. June 2018.

Large-scale studies of Latin-American populations. Presented at The 20th Anniversary Symposium at the Institute for Health Promotion, Graduate School of Public Health, Yonsei University, Seoul, South Korea. April 2018.

Smoking and mortality in Cuba. Presented at the National Institute of Cardiology, Havana, Cuba. March 2018.

Smoking and mortality: 14-year prospective study of 150,000 Mexican adults. Presented at the British Heart Foundation (BHF) Centre for Research Excellence (CRE) Symposium 2017 (November 2017); and the NDPH Symposium 2018 (January 2018).

Awarded Judge's Choice at the BHF CRE poster presentation competition

Awarded 1st prize at the annual NDPH poster presentation competition

Abstract

Smoking and cause-specific mortality in large prospective studies from Cuba, Mexico, and the United States

Blake Thomson, Balliol College, University of Oxford

A thesis submitted for the degree of Doctor of Philosophy, Michaelmas 2019

Background: Tobacco smoking caused an estimated 100 million deaths in the 20th century, and unless widespread cessation and prevention occur, it could cause one billion deaths in the 21st century. An estimated 50 million smokers live in Latin America, but there is little direct evidence of the effects of smoking in Latin American populations.

Methods: Information from three large, prospective studies of Latin American populations was used to determine the relevance of smoking and smoking cessation to mortality. In 1996-2002, 146 556 Cuban adults aged ≥30 years were recruited into the Cuba Prospective Study, of whom 8691 died at ages 30-69 during follow-up. In 1998-2004, 159 755 Mexican adults aged ≥35 years were recruited into the Mexico City Prospective Study, of whom 19 091 died at ages 35-89 during follow-up (10 773 without major chronic disease at recruitment, and 8318 with major chronic disease). Finally, in 1997-2014, 446 531 adults aged ≥30 years were recruited into the US National Health Interview Survey, of whom 48 440 died at ages 30-89 during follow-up (3242 Mexican Americans, 2177 other Hispanics, 34 482 non-Hispanic Whites, 6978 non-Hispanic Blacks, and 1561 participants of other race/ethnicity). In each prospective study, Cox regression (adjusting for age-at-risk, sex, education, area, and alcohol consumption) was used to relate smoking habits at recruitment to mortality rate ratios (RRs). Resurveys of Cuban and Mexican participants informed 'usual' smoking behaviours.

Results: In the Cuba Prospective Study, half of men and one-quarter of women were smokers at recruitment. One-third of current cigarette smokers reported having started smoking regularly in childhood (age <15 years). Compared with never smokers, all-cause

mortality RRs at ages 30-69 for Cuban cigarette smokers (mean 15 cigarettes/day) who had started at ages 5-9, 10-14, 15-19, and ≥20 years were 2.51 (95%CI 2.21-2.85), 1.83 (1.72-1.95), 1.56 (1.46-1.65) and 1.50 (1.39-1.62), respectively (overall RR for current smokers: 1.66 [1.59-1.74]). In the Mexico City Prospective study, half of men and one-quarter of women were smokers at recruitment, and one-fifth of smokers began smoking before age 15, but consumption was low (8 cigarettes/day among daily smokers). Among Mexican adults without chronic disease at recruitment, compared with never smokers, the all-cause mortality RR at ages 35-69 for smoking ≥10 cigarettes/day was 1.52 (1.36-1.70), and was similar at ages 70-89 and among men and women (overall RR for daily smokers: 1.27 [1.19-1.35]). In the US National Health Interview Survey prospective study, one-fifth of smokers began smoking before age 15, which was associated with three times the mortality rate at ages 30-69 of never smoking, with even greater risk among those starting to smoke before age 10. Compared with never smokers, the all-cause mortality RR for daily smokers was 1.39 (1.24-1.54) among Mexican Americans, 2.45 (2.20-2.72) among other Hispanics, 2.74 (2.61-2.89) among non-Hispanic Whites, 2.29 (2.19-2.41) among non-Hispanic Blacks, and 2.08 (1.84-2.35) among participants of other races. In each of these populations, smoking cessation by age ~40 avoided most of the excess risk associated with continued smoking.

Conclusions: In Cubans, Mexicans, and Mexican Americans, the excess mortality risk associated with smoking was substantial, but was lower than the risk associated with smoking among Western populations. This may be due to a combination of intermittent smoking patterns in each of these populations, lower levels of consumption among Mexicans and Mexican Americans, and variability in products smoked over time. This thesis provides the first reliable, direct evidence of the relationship between smoking and mortality in Cuba and Mexico, which led to estimates of smoking-attributable mortality similar to those from the Global Burden of Disease (GBD) Study. Therefore, the GBD Study may provide the best available estimates of smoking-attributable mortality in other Latin American populations in which there is not yet direct evidence.

Statement of originality

The work presented in this thesis is my own, with research questions and statistical analyses shaped by the supervision of Prof Jonathan Emberson, Prof Sarah Lewington, Dr Ben Lacey, and Prof Sir Richard Peto. I had no personal involvement in the conduct of the baseline surveys or resurveys for the Cuba Prospective Study, the Mexico City Prospective Study, or the US National Health Interview Survey. The former data sets were provided to me by the respective Oxford co-PI's, Prof Sarah Lewington and Prof Jonathan Emberson, and the latter data set is in the public domain. I was not involved in the data collection for the US National Health Interview Survey, though I sought, acquired, and cleaned these data myself. I wrote all statistical programmes myself, but several SAS macros and plots were informed by the prior work of Ms Rachel Wade, Mr Paul Sherliker, Dr Julie Burrett, and Prof Jonathan Emberson. I conducted all analyses and prepared the draft of this thesis myself, which was further refined with the guidance of my supervisors. Any errors are entirely my own.

Contents

Acknowledgements....iii

Abstract....vii

Statement of originality....ix

List of Tables....xvi

List of Figures....xix

Chapter 1: Introduction....1

1.1 History of the evidence: Before and after the British Doctors Study....2

1.2 Cohort effects (and their relevance to smoking among women)....4

1.3 The relevance of different smoking behaviours to mortality....6

1.4 The current state of the evidence on smoking and resultant policies....9

1.5 The need for continued research on smoking in diverse populations globally....11

1.6 Outline of the thesis....12

Chapter 2: Review of the literature....14

2.1 Chapter aims....14

2.2 Methodology....14

2.3 Case-control evidence....15

2.3.1 Heart disease....16

2.3.2 Cancer....17

2.4 Prospective evidence....19

2.5 Summary of findings....23

Chapter 3: Smoking and mortality in the Cuba Prospective Study....29

3.1 Chapter aims....29

3.2 Baseline survey....29

3.2.1 Study design....29

3.2.2 Study population 30
3.2.3 Data collection 31
3.3 Resurvey 31
3.4 Follow-up 32
3.5 Methods 32
3.5.1 Assessment of smoking 32
3.5.2 Selection of covariates 33
3.5.3 Participant exclusions 34
3.5.4 Statistical analysis 35
3.6 Results 38
3.6.1 Participant characteristics 38
3.6.2 Trends in smoking behaviour across birth cohorts 38
3.6.3 Characteristics of participants at baseline and resurvey by smoking status 39
3.6.4 Association between current smoking and mortality 40
3.6.4.1 All-cause mortality 40
3.6.4.2 Cause-specific mortality 42
3.6.5 Association between smoking cessation and all-cause mortality 44
3.6.6 Population attributable fraction of mortality associated with ever smoking 44
3.7 Discussion 45
3.7.1 Summary of findings 45
3.7.2 Strengths and limitations 46
3.7.3 Findings in the context of prior evidence 49
3.8 Conclusion 53
Chapter 4: Smoking and mortality in the Mexico City Prospective Study 74
4.1 Chapter aims 74
4.2 Baseline survey 74

4.2.1 Study design 74
4.2.2 Study population 75
4.2.3 Data collection 75
4.3 Resurvey 76
4.4 Follow-up 77
4.5 Methods 78
4.5.1 Assessment of smoking 78
4.5.2 Selection of covariates 79
4.5.3 Participant exclusions 80
4.5.4 Statistical analysis 80
4.6 Results 83
4.6.1 Participant characteristics 83
4.6.2 Trends in smoking behaviour across birth cohorts 84
4.6.3 Characteristics of participants at baseline and resurvey by smoking status 85
4.6.4 Association between smoking and mortality among those with and without chronic disease at recruitment 87
4.6.5 Association between daily smoking and mortality among participants without prior chronic disease at recruitment 88
4.6.5.1 All-cause mortality 88
4.6.5.2 Cause-specific mortality 90
4.6.6 Association between smoking cessation and all-cause and cause-specific mortality among those without chronic disease at recruitment 92
4.6.7 Population attributable fraction of mortality associated with ever smoking 93
4.7 Discussion 94
4.7.1 Summary of findings 94
4.7.2 Strengths and limitations 96

4.7.3 Findings in the context of prior evidence 99
4.8 Conclusion 104
Chapter 5: Smoking and mortality in the US National Health Interview Survey 123
5.1 Chapter aims 123
5.2 Baseline survey 124
5.2.1 Study design 124
5.2.2 Study population 124
5.2.3 Data collection 125
5.3 Resurvey 126
5.4 Follow-up 126
5.5 Methods 128
5.5.1 Assessment of smoking 128
5.5.2 Assessment of race/ethnicity 129
5.5.3 Selection of covariates 129
5.5.4 Participant exclusions 130
5.5.5 Statistical analysis 131
5.6 Results 133
5.6.1 Participant characteristics by race 133
5.6.2 Trends in smoking behaviour across birth cohorts by race 134
5.6.3 Characteristics of participants by race and smoking status at recruitment . 136
5.6.4 Pooled results: Association between smoking (and quitting) and mortality. 137
5.6.4.1 Participant characteristics by age started smoking 137
5.6.4.2 Association between age started smoking and mortality 137
5.6.4.3 Association between amount smoked and mortality 138
5.6.4.4 Association between smoking cessation and mortality 139
5.6.4.5 National estimates of smoking among US adults (NHIS 2018) 140

5.6.5 Association between smoking and mortality by race 140
5.6.5.1 Overall 140
5.6.5.2 By age started and amount smoked 141
5.6.6 Association between smoking cessation and mortality by race 142
5.6.7 Population attributable fraction of mortality associated with ever smoking by race 143
5.7 Discussion 144
5.7.1 Summary of findings 144
5.7.2 Strengths and limitations 145
5.7.3 Findings in the context of prior evidence 148
5.8 Conclusion 151
Chapter 6: Discussion 178
6.1 Thesis overview 178
6.2 Summary of main findings across studies 179
6.2.1 Trends in smoking across birth cohorts and sex 179
6.2.2 Prospective evidence on the hazards of smoking 180
6.2.3 Prospective evidence on the benefits of quitting 182
6.3 Consistency with published literature 183
6.4 Findings in global context: Evidence from the cross-sectional Global Adult Tobacco Surveys in 30 low- and middle-income countries 185
6.5 Strengths and limitations of the thesis 189
6.5.1 Strengths 189
6.5.2 Limitations 190
6.6 Implications for future research 192
6.6.1 Assessment of exposure 192
6.6.2 Assessment of outcomes 194

6.6.3 Prospective study of lifestyle risk factors in diverse populations globally 195
6.7 Conclusions 197
References 215
Appendix A: Additional results for the Cuba Prospective Study 223
Appendix B: Additional results for the Mexico City Prospective Study 232
Appendix C: Additional results for the US National Health Interview Survey prospective study 245
Appendix D: Additional information and methods for the Global Adult Tobacco Surveys 256
Appendix E: Questionnaires for the Cuba Prospective Study and the Mexico City Prospective Study 260

List of Tables

List of Abbreviations xxv

Table 2.1: Literature review search terms 25

Table 2.2: Case-control studies of smoking in Latin American populations (≥1000 cases), in the order described in text 26

Table 2.3: Prospective studies of smoking in Latin American populations (≥1000 participants), in order of publication date 27

Table 3.1: ICD-9 and ICD-10 codes for cause-specific mortality endpoints 55

Table 3.2: Baseline characteristics of the 143 920 participants in the Cuba Prospective Study, by age and sex 56

Table 3.3: Baseline characteristics of the 120 929 participants aged 30-69 at recruitment, by smoking status 57

Table 3.4: Baseline characteristics of the included male participants aged 30-69 at recruitment, by smoking status 58

Table 3.5: Baseline characteristics of the included female participants aged 30-69 at recruitment, by smoking status 59

Table 3.6: Smoking characteristics at resurvey (mean 6 years later) according to baseline smoking status, ages 30-69 at recruitment 60

Table 3.7: Cause-specific mortality rate ratios (RRs) and population attributable fractions (PAFs) for ever versus never smoking at ages 30-69, by sex 61

Table 4.1: ICD-10 codes for cause-specific mortality endpoints 105

Table 4.2: Baseline characteristics of the 158 688 participants aged 35-89 at recruitment, by age and sex 106

Table 4.3: Baseline characteristics of the 130 791 included participants aged 35-89 at recruitment, by sex and smoking status 107

Table 4.4: Smoking characteristics at resurvey according to baseline smoking status, among those without chronic disease at baseline, ages 35-89 108
Table 4.5: Cause-specific mortality rate ratios (RRs) and population attributable fractions (PAFs) for ever versus never smoking at ages 35-89, by sex 109
Table 4.6: All-cause mortality rate ratios (RRs) and population attributable fractions (PAFs) for ever versus never smoking at ages 35-89, by chronic disease status at recruitment 110
Table 5.1: ICD-10 codes for 10 cause-specific mortality endpoints (as provided by the National Center for Health Statistics, Maryland, USA) 152
Table 5.2: Baseline characteristics of the US (NHIS) cohort participants at ages 30-84, by race/ethnicity 153
Table 5.3: Baseline characteristics of the US (NHIS) cohort participants at ages 30-84, by sex and race/ethnicity 154
Table 5.4 (Part 1 of 2): Baseline characteristics of the US (NHIS) cohort participants at ages 30-84, by race/ethnicity and smoking status 155
Table 5.4 (Part 2 of 2): Baseline characteristics of the US (NHIS) cohort participants at ages 30-84, by race/ethnicity and smoking status 156
Table 5.5: Pooled US (NHIS) baseline characteristics by smoking habits, ages 30-69 157
Table 5.6: Estimated number of adult smokers (age≥18) in the United States in 2018 by age started smoking 158
Table 5.7: Population attributable fraction (PAF) for ever versus never smoking at ages 30-89, by race and sex 159
Table 6.1: Population attributable fractions for mortality from medical causes associated with ever-regular smoking in Cuba, Mexico, and the United States 199
Table 6.2: Global Adult Tobacco Survey respondent characteristics by country 200

Table A1: Baseline characteristics of the participants aged 70-89 at recruitment, by sex and smoking status 224
Table B1: Baseline characteristics of the participants at ages 35-89, by sex and smoking status at recruitment, among those with chronic disease at recruitment 233
Table B2: Baseline characteristics of the participants at ages 35-69, by sex and smoking status at recruitment, among those without chronic disease at recruitment) 234
Table B3: Smoking characteristics at resurvey according to baseline smoking status, ages 35-69 235
Table B4: Smoking characteristics at resurvey according to baseline smoking status, among those with chronic disease at recruitment, ages 35-89 236
Table B5: Population attributable fraction (PAF) at ages 35-69, by disease and sex ... 237
Table C1: National estimates of never, ex, non-daily, and daily smokers by race in the United States in 2018 (in millions) 246
Table C2: Unweighted sample size of never, ex, non-daily, and daily smokers by race in the United States in 2018 (NHIS) 247

List of Figures

Figure 2.1: Flow chart of selection process for reports included in the systematic review 28
Figure 3.1: Smoking prevalence in the Cuba Prospective Study 62
Figure 3.2: Mean cigarettes/day, and age started smoking, among 44 042 current daily smokers 63
Figure 3.3: All-cause mortality rate ratio at ages 30-69, cigarette smoker vs never smoker, by age started smoking and usual amount smoked 64
Figure 3.4: All-cause mortality rate ratio at ages 30-69, cigarette smoker vs never smoker, by age started smoking and usual amount smoked 65
Figure 3.5: All-cause mortality rate ratio at ages 30-69, cigarette smoker vs never smoker, by age started smoking and sex 66
Figure 3.6: All-cause mortality rate ratio at ages 30-69, cigarette smoker vs never smoker, by usual amount smoked and sex 67
Figure 3.7: Risks associated with current cigarette vs never smoking by levels of factors adjusted for in main analyses 68
Figure 3.8: Cause-specific mortality rate ratio at ages 30-69, cigarette vs never smoker, by age started smoking 69
Figure 3.9: Cause-specific mortality rate ratio at ages 30-69, cigarette vs never smoker, by amount smoked 70
Figure 3.10: Cause-specific mortality rate ratio at ages 30-69, cigarette or cigar smoker vs never smoker, by product smoked 71
Figure 3.11: All-cause mortality rate ratio at ages 30-69, cigarette smoker vs never smoker, by age at cessation and years since quitting 72
Figure 3.12: Cigarettes sold per adult per day in Cuba, 1980-2015 73

Figure 4.1: Baseline smoking prevalence in the Mexico City Prospective Study..... 111
Figure 4.2: Mean cigarettes/day at recruitment and age started smoking among the 28 122 daily smokers..... 112
Figure 4.3: Association between smoking habits and all-cause mortality among those with vs without chronic disease at recruitment by smoking status, age 35-89..... 113
Figure 4.4: All-cause mortality rate ratio at ages 35-69 and 70-89, daily smoker vs never smoker, by usual amount smoked..... 114
Figure 4.5: All-cause mortality rate ratio at ages 35-69 and 70-89, daily smoker vs never smoker, by age started smoking..... 115
Figure 4.6: All-cause mortality rate ratio at ages 35-69 and 70-89, daily smoker vs never smoker, by usual amount smoked and sex..... 116
Figure 4.7: All-cause mortality rate ratio at ages 35-69 and 70-89, daily smoker vs never smoker, by age started smoking and sex..... 117
Figure 4.8: All-cause mortality rate ratio at ages 35-89, daily smoker vs never smoker, by levels of factors adjusted for in main analyses..... 118
Figure 4.9: Cause-specific mortality at ages 35-89, daily smoker vs never smoker, by amount smoked..... 119
Figure 4.10: Cause-specific mortality at ages 35-89, daily smoker vs never smoker, by age started smoking..... 120
Figure 4.11: Cause-specific mortality at ages 35-89, never vs ‘long-term’ ex-daily and current daily smokers..... 121
Figure 4.12: Cause-specific mortality at ages 35-89, never vs ex-daily smokers, by quitting duration..... 122
Figure 5.1: Baseline smoking prevalence among Mexican Americans in the US NHIS 160
Figure 5.2: Baseline smoking prevalence among other Hispanics in the US NHIS..... 161

Figure 5.3: Baseline smoking prevalence among non-Hispanic Whites in the US NHIS 162
Figure 5.4: Baseline smoking prevalence among non-Hispanic Blacks in the US NHIS 163
Figure 5.5: Baseline smoking prevalence among other race/ethnicity in the US NHIS 164
Figure 5.6: Mean cigarettes/day and age started smoking by race and sex among 68 853 daily smokers in the US NHIS prospective study (1997-2014)..... 165
Figure 5.7: All-cause mortality rate ratio at ages 30-69, daily vs never smokers, by age started smoking 166
Figure 5.8: All-cause mortality rate ratio at ages 30-69, daily vs never smokers, by amount smoked 167
Figure 5.9: All-cause mortality rate ratio by age quit smoking among those quitting at ages 15-54, compared to never and current smokers at baseline, age at risk 55-89 ... 168
Figure 5.10: All-cause mortality rate ratio by years since quitting smoking, compared to never and current smokers at baseline, age at risk 30-89 169
Figure 5.11: All-cause mortality rate ratio at ages 30-89, current daily smoker vs never smoker, by race 170
Figure 5.12: All-cause mortality rate ratio at ages 30-89, current daily smoker vs never smoker, by race 171
Figure 5.13: All-cause mortality rate ratio at ages 30-89, current daily smoker vs never smoker, by age started smoking and race 172
Figure 5.14: All-cause mortality rate ratio at ages 30-89, current daily smoker vs never smoker, by amount smoked and race 173
Figure 5.15: All-cause mortality rate ratio by years since quitting smoking and race, compared to never and current smokers at baseline, age at risk 30-89..... 174

Figure 5.16: All-cause mortality rate ratio by age quit smoking and race among those quitting at ages 15-54, compared to never and current smokers at baseline, age at risk 55-89 175
Figure 5.17: All-cause mortality rate ratio by years since quitting smoking and race, compared to never and current smokers at baseline, age at risk 30-89 176
Figure 5.18: All-cause mortality rate ratio by age quit smoking and race among those quitting at ages 15-54, compared to never and current smokers at baseline, age at risk 55-89 177
Figure 6.1: Prevalence of smoking frequency among men in GATS surveys, by country 201
Figure 6.2: Prevalence of smoking frequency among women in GATS surveys, by country 202
Figure 6.3: Proportion of male current smokers who smoke daily in GATS surveys, by country 203
Figure 6.4: Proportion of female current smokers who smoke daily in GATS surveys, by country 204
Figure 6.5: Cigarettes smoked per day among male cigarette smokers in GATS surveys, by country 205
Figure 6.6: Cigarettes smoked per day among female cigarette smokers in GATS surveys, by country 206
Figure 6.7: Age at smoking initiation among male cigarette smokers in GATS surveys, by country 207
Figure 6.8: Age at smoking initiation among female cigarette smokers in GATS surveys, by country 208
Figure 6.9: Adult smokers who began smoking daily at ages 10-14 in GATS surveys, by country 209

Figure 6.10: Adult smokers who began smoking daily at ages 5-9 in GATS surveys, by country 210
Figure 6.11: Quit ratio among ever-smoking men in GATS surveys, by country 211
Figure 6.12: Quit ratio among ever-smoking women in GATS surveys, by country 212
Figure 6.13: Time to first smoking after waking among men in GATS surveys, by country 213
Figure 6.14: Time to first smoking after waking among women in GATS surveys, by country 214
Figure A1: Cigarettes smoked per day at recruitment, by product smoked 225
Figure A2: Cigars smoked per day at recruitment, by product smoked 226
Figure A3: All-cause mortality rate ratio at ages 70-89, cigarette smoker versus never smoker, by age started smoking and usual amount smoked 227
Figure A4: All-cause mortality rate ratio at ages 70-89, cigarette smoker versus never smoker, by age started smoking and usual amount smoked 228
Figure A5: Cause-specific mortality rate ratio at ages 70-89, cigarette versus never smoker, by age started smoking 229
Figure A6: Cause-specific mortality rate ratio at ages 70-89, cigarette versus never smoker, by amount smoked 230
Figure A7: All-cause mortality rate ratio at ages 30-69, cigarette smoker versus never smoker, by age at cessation and years since quitting 231
Figure B1: Cigarettes smoked per occasion by smoking frequency among current smokers 238
Figure B2: Risk of cause-specific mortality among daily vs never smokers by amount smoked, ages 35-69 239
Figure B3: Risk of cause-specific mortality among daily vs never smokers by age started, ages 35-69 241

Figure B4: Risk of cause-specific mortality among daily vs never smokers by amount smoked, ages 35-89 242
Figure B5: Risk of cause-specific mortality among daily vs never smokers by age started, ages 35-89 243
Figure B6: Risk of cause-specific mortality, never vs ex-daily smokers by quitting duration, ages 35-89 244
Figure C1: Mortality rate ratio at ages 30-69 for major smoking-related causes, daily vs never smokers at recruitment, by age started smoking 248
Figure C2: All-cause mortality rate ratio at ages 70-89, daily vs never smokers, by age started smoking 249
Figure C3: All-cause mortality rate ratio at ages 70-89, daily vs never smokers, by amount smoked 250
Figure C4: Mortality rate ratio at ages 30-69 for major smoking-related causes, daily vs never smokers at recruitment, by amount smoked 251
Figure C5: Mortality rate ratio at ages 30-89 for major smoking-related causes, by age quit smoking among those quitting at ages 15-54, compared to never and current smokers at recruitment 252
Figure C6: Mortality rate ratio at ages 30-89 for major smoking-related causes, by years since quitting smoking, compared to never and current smokers at recruitment 253
Figure C7: All-cause mortality rate ratio by age quit smoking among those quitting at ages 15-54, compared to never and current smokers at baseline, age at risk 55-89 ... 254
Figure C8: All-cause mortality rate ratio by years since quitting smoking, compared to never and current smokers at baseline, age at risk 30-89 255

List of Abbreviations

BMI	Body mass index
CI	Confidence interval
COPD	Chronic obstructive pulmonary disease
CTSU	Clinical Trial Service Unit and Epidemiological Studies Unit
CVD	Cardiovascular disease
FCTC	Framework Convention on Tobacco Control
GATS	Global Adult Tobacco Survey
ICD	International Classification of Diseases
IHD	Ischaemic heart disease
LMICs	Low- and middle-income countries
MI	Myocardial infarction
NCD	Non-communicable disease
NCHS	National Center for Health Statistics
NDI	National Death Index
NDPH	Nuffield Department of Population Health
NHIS	National Health Interview Survey
OR	Odds ratio
PAF	Population attributable fraction
RR	Rate ratio
SD	Standard deviation
SDG	Sustainable Development Goal
WHO	World Health Organisation

Chapter 1: Introduction

Smoking is a leading preventable cause of death globally, with an estimated 6.4 million deaths attributable to smoking in 2015.[1] Tobacco smoking caused an estimated 100 million deaths in the 20th century, and unless widespread cessation and prevention occur, it is projected that it could cause one billion deaths in the 21st century.[2]

The causal association between smoking and premature mortality is well established. Studies in the UK, US, and Australia have shown that half of lifelong smokers in these populations were killed by their habit, with a quarter of these deaths occurring before the age of 70 (here termed 'premature death').[3-5] However, those who quit smoking before age 40 avoided most of the excess risk of premature death associated with continued smoking in each of these populations. Nevertheless, despite this clear evidence on the hazards of smoking and benefits of quitting in these populations, continued research into the health effects of tobacco worldwide remains an urgent global health priority.

Although most tobacco deaths in the last century occurred in high-income countries, this century most deaths caused by tobacco will occur in low- and middle-income countries (LMICs).[6] To date, however, little direct evidence from these countries has been produced. Despite consistent contemporary evidence on the association between smoking and premature death from populations in several high-income countries, the health effects of smoking are likely to differ in populations in which the cigarette epidemic is not yet mature.[7] In addition, the relevance of smoking to mortality may change over time due to changes in several factors: smoking patterns (especially younger age at initiation and prolonged smoking without cessation), the composition of products smoked, other major causes of death, and underlying death rates in the population being studied.[8,9] Therefore, further research is needed.

This chapter provides an overview of the major findings on tobacco smoking and its health effects over the past century, provides a rationale for the continued study of smoking in diverse populations globally, and introduces the aims of this thesis on the hazards of smoking and benefits of quitting in large prospective studies from Cuba, Mexico, and the United States.

1.1 History of the evidence: Before and after the British Doctors Study

There are two eras in the history of research into the health effects of smoking: what was known before the British Doctors Study, and what has followed as a result of Sir Richard Doll's landmark findings nearly 70 years ago (conducted with Sir Austin Bradford Hill). Although there had been several small studies that pointed to a link between tobacco and disease (most notably lung cancer), by the 1940's it remained unclear what association, if any, smoking had with mortality risk. However, it was noted that the number of deaths from lung cancer in England and Wales had increased approximately 15-fold between 1922 and 1947, and a rapid rise had been reported in several other (mostly high-income) countries.[10] Several factors were considered as possible causes of this apparent increase in lung cancer: standards of diagnosis for lung cancer had substantially improved over the preceding years; over the prior half-century, exhaust fumes from cars, tarred roads, gas-works, industrial plants, and coal fires had all become more prevalent; and finally, tobacco smoking (particularly in the form of cigarettes) had increased.[11]

At the time of publication of their case-control study in 1950, Doll and Hill reported that small case-control studies from Germany, the US, and England had found associations between smoking and lung cancer, with more definitive evidence from the US being published the same year as their own study.[10,12-15] In the largest contemporary US study, Wynder and Graham reported that, of 605 men with bronchial carcinoma, 1.3% were

non-smokers, while 51.2% had smoked >20 cigarettes/day over the same period (compared with male general hospital cases of a similar age, of whom 14.6% were non-smokers, and 19.1% smoked >20 cigarettes/day).[12] In Doll and Hill's case-control study, planned in 1947 and conducted in 1948-1949, 709 cases with lung cancer and 709 control cases without cancer were compared (selected to be matched on age and sex). In a testament to the epidemic of tobacco smoking among men in the country at that time, it was found that 0.3% of men with lung cancer were non-smokers, a statistically significant difference from the just 4.2% of controls who were non-smokers. Among women, 31.7% of lung cancer cases were non-smokers, compared with 53.3% of controls (indicating that lung cancer could commonly occur even among non-smokers). Whether comparing the most recent amount of tobacco consumed, the maximum amount consumed, or the total amount consumed, the emergent pattern was clear: "it must be concluded that there is a real association between carcinoma of the lung and smoking."[10] Two less certain findings were that smoking pipes was less strongly associated with lung cancer than smoking cigarettes, and that inhaling while smoking was more common in controls than in those with lung cancer (though the difference was not significant when age and sex differences were taken into account), indicating that inhalation had little association with lung cancer. The latter observation would famously lead to substantial criticism by, and much scepticism from, the renowned statistician R.A. Fisher, among others.[11,16] Nevertheless, there was sufficient evidence to motivate further research into the relationship between tobacco and lung cancer.

A series of further retrospective studies emerged following the findings of Doll and Hill's report in England, and (primarily) Wynder and Graham's in the US. Although the findings were generally similar across these studies, for many, uncertainty remained about the association being causal. Doll and Hill concluded that additional "retrospective studies of that same kind would seem to us unlikely to advance our knowledge materially… any

undetected flaw in the evidence that such studies have produced, it would be exposed only by some entirely new approach."[17] Their response was to employ a new study design: the prospective study. (Their report included the Oxford English Dictionary definition of the word 'prospective' in a footnote, with a quote from Leigh Hunt illustrating its use.) The result was the British Doctors Study, which began in 1951-1954 with 40 564 doctors who replied to the mailed questionnaire, and followed up for the next five decades. The preliminary report of 29 months of follow-up among the 24 389 men over the age of 35 provided a clear dose-response relationship between tobacco smoking and lung cancer, and evidence of a potential relationship with other diseases, including 'coronary thrombosis'. Evidence from the US was again published the same year as Doll and Hill's study, this time a large prospective study by Hammond and Horn (a precursor to the American Cancer Society's Cancer Prevention Studies). Despite being set up, in fact, to disprove the relationship between smoking and lung cancer,[11,16] the US study showed a similarly definitive relationship between smoking and lung cancer (and likely heart disease).[18] With these first clear results published, and meticulous prospective studies set up to continue following participants' vital status into the future, the era of large-scale research into the effects of smoking on human health had begun.

1.2 Cohort effects (and their relevance to smoking among women)

As evidence from other (generally high-income) populations began to emerge on the association between smoking and health,[19-24] subsequent updates to the British Doctors Study over the next 50 years[3,25-27] (and, notably, the large US Cancer Prevention Studies initiated in 1959[28] and 1982[29,30]) continued to shed light on the tobacco epidemic. While early efforts were focused on the association between smoking and lung cancer, initially in men and later in women,[31,32] three important points became clear over the following decades. The first was that smoking's effect on human health was not limited to lung cancer, but in fact led to increased risk of death from multiple diseases,

including death from cardiovascular, respiratory, neoplastic, infectious, and other causes. Second, that while there was now no doubt that smoking killed many of those who smoked, quitting worked, and could lead to large reductions in the excess risk associated with otherwise continuing to smoke. And the third, which took decades of committed work to show definitively, was the importance of the generational effects on smoking-related mortality.

One of R.A. Fisher's major criticisms about the plausibility of a causal relationship between smoking and lung cancer was that lung cancer rates in the UK had been rising much more rapidly in men than in women, even though female smoking habits had been increasing in recent decades, while those of men had generally plateaued.[33] What Fisher had missed, which Doll in particular clarified over the course of his half-century of evidence on British doctors, was the importance of generational or 'cohort' effects—that the hazards of smoking depend not only on recent smoking habits, but consistent, long-term smoking habits over many years. It is now clear that the impact of generational effects on the relationship between smoking and mortality, even within a population, is not necessarily static.

Over the course of the first 40 years of the British Doctors Study, the all-cause mortality rate ratio (RR) among men dying in middle age (here, ages 45-74) was approximately two, indicating that half of smokers were killed by their habit.[27] Overall, fifty years of observation from the British Doctors Study showed that, among those dying at age ≥60, those born in the 19th century had less than half the excess risk associated with smoking than those born in the 20th century, even when amount smoked was taken into account.[3] Among men aged 60 and over, compared to lifelong non-smokers, the all-cause mortality RR for current smokers was 1.46 among those born in the 19th century, and 2.19 among those born in the 20th century. This was due to the fact that smoking from a

young age and throughout adult life was not common until the 1900's, and especially after the 1920's (following the First World War). With successive decades of birth cohorts in the 20th century, the reduction in premature mortality observed in never smokers was offset among smokers by the increasing maturity of the smoking epidemic, leading to an increasingly wide gap in survivorship between smokers and non-smokers. At the 20-year follow-up of the British Doctors Study, the annual death rate for lung cancer was 10 among non-smokers and was 104 for current smokers (RR of 10.4); by the 50-year follow-up, the lung cancer RR was 15.9.[26]

Similar effects were observed in the United States over successive prospective studies (and successive generations of smokers): among US cohorts recruited in 1959-1965, 1982-1988, and 2000-2010, the smoker versus non-smoker lung cancer RRs for women were 2.73 (95% CI 2.07-3.61), 12.65 (23.17-28.40), and 25.66 (23.17-28.40), respectively.[34] Despite the substantial increase in lung cancer risk among smokers, the hazards of smoking may not yet have peaked among women in the most recent study.[8] Directly studying the full effects of lifelong smoking among women in the UK who had smoked substantial numbers of cigarettes throughout adult life was only recently possible, and the results were clear: "if women smoke like men, they die like men—but, stopping early enough gains 10 extra years of life expectancy for women or men."[35]

1.3 The relevance of different smoking behaviours to mortality

An important aspect of cohort effects is behaviour change, which in the case above, was primarily men (and later women) starting to smoke from younger ages at the population level. However, individual behaviour change is also important to take into account. From early on in the investigation of smoking in relation to mortality in the British Doctors Study, ascertaining long-term ('usual') smoking habits was prioritised. This proved to be an important feature of the study, given that the repeated longitudinal surveys

('resurveys') demonstrated that the study population substantially reduced its cigarette consumption over the follow-up period (and therefore, the harms of prolonged smoking would have been underestimated).[26] Another notable change over the course of decades was the shift from pipe or cigar smoking to cigarette smoking,[11] which again caused a shift in the association between smoking and mortality risk. The idea that there might be a 'safe' cigarette was ultimately found to be a chimera: though lower-tar or lower-nicotine cigarettes somewhat reduced the risk of lung cancer and heart disease compared with 'regular' cigarettes, the mortality risk of smoking such cigarettes was still several times that of not smoking at all (or long-term cessation).[36-39]

Two factors that did strongly relate to smokers' mortality risk, however, were the amount smoked and the duration of smoking (or, for lifelong smokers, the age at smoking initiation). As early as 1964, Hammond in the US had demonstrated that the hazards of smoking depended heavily on the age at smoking initiation, and that starting younger was associated with greater excess risk than starting older, examining initiation from as young as age <15 years.[28] It was observed that age at smoking initiation was particularly important because it had implications not only for potential duration of smoking exposure over a lifetime, but also because it was associated with heavier smoking. However, it appeared that the relevance of age at smoking initiation to mortality was likely greater than that of amount smoked when each was stratified by the other. In 1978, Doll observed that the incidence rate of lung cancer in non-smokers was proportional to age to the fourth power, while the excess risk of lung cancer in smokers was proportional to smoking duration to the 4^{th} power multiplied by amount smoked.[40] Although there is some evidence that the power of cigarettes smoked/day in relation to lung cancer incidence may be 2 instead of 1, the fact that the power of smoking duration was approximately 4 nevertheless suggests that age started (and stopped) smoking is particularly important, at least for lung carcinogenesis. This has been an argument

against the use of pack-years in studying the health effects of smoking (as this metric combines amount smoked and duration of smoking equally).[41,42]

From a public health perspective, the most important factor relating smoking to death is almost certainly not the product smoked, the amount smoked , or even the age at smoking initiation, but whether smokers manage to quit smoking (ideally, well before smoking results in ill health, though even afterwards significant health benefits likely accrue). While the benefits of quitting were convincing even from the 1960's, by now they have been shown definitively in several populations. The clearest evidence has come from the 50-year follow-up to the British Doctors Study, which showed that smokers died an average of approximately 10 years younger than lifelong non-smokers, but that cessation at age 60, 50, 40, or 30 resulted in gaining 3, 6, 9, or nearly up to all 10 of those years of life expectancy back, respectively.[3] In the aftermath of this observation, a number of studies have observed almost identical patterns: quitting by age ~40 (and ideally much younger than 40) avoided most of the excess risk associated with continued smoking, in British women,[35] in the US,[4] in Australia,[5] in Japan,[43] and in men across Asian countries,[44] among others. Observations from these and other populations suggest that the health benefits of cessation begin to accrue almost immediately, and continue to grow with duration of time since cessation. Among those who quit early enough, particularly while still healthy, the risk associated with ex-smoking may reach nearly the same level of risk as never smokers (ie, almost no excess risk). This is contingent, of course, on surviving long enough to quit smoking in the first place, doing so before the adverse health effects have become irreversible, and maintaining long-term smoking cessation into the future.

1.4 The current state of the evidence on smoking and resultant policies

In several high-income countries, the full harms of lifelong tobacco smoking have been realised, with lifelong smokers losing an average of ~10 years of life expectancy in comparison to never smokers. These losses in life expectancy are driven by much higher relative risks among smokers compared with non-smokers for certain causes of death. For example, in the Million Women Study, the 12-year relative risk of current versus never smoking was 35.3 (95% CI 29.2-42.5) for chronic lung disease, 21.4 (19.7-23.2) for lung cancer, 4.5 (4.2-4.8) for coronary heart disease, and 3.1 (2.8-3.3) for cerebrovascular disease (among many other major causes associated with smoking).[35] The relative risks of current versus never smoking among women aged ≥55 years in a recent US cohort (recruited 2000-2010) were similar in magnitude for these causes, including 22.4 (19.6-25.6) for COPD, 25.7 (23.2-28.4) for lung cancer, 2.9 (2.7-3.1) for ischaemic heart disease, and 2.1 (1.8-2.4) for stroke.[8] The relative risk for all-cause mortality comparing current and never smokers in each of these cohorts was approximately 3, indicating that, if these associations are largely causal, approximately 2/3 of deaths among smokers were due to smoking.

The widespread acknowledgement of the hazards of smoking in these countries has led to effective tobacco control policies like increased excise taxes on tobacco[2,45,46] advertising bans,[47,48] and better consumer information on the hazards of smoking and benefits of quitting,[49] resulting in large reductions in the prevalence of smoking in some countries. In 2000, Peto *et al.* demonstrated the early effects of such cessation: between 1950 and 2000 smoking prevalence had plummeted in the UK (driven largely by cessation among smokers, and some prevention of uptake among never smokers), particularly among men, and resulted in a halving of the national lung cancer mortality rate that would have been expected had ex-smokers continued to smoke.[50] Smoking prevalence has since continued to decline in the UK and in many other high-income

countries. For example, estimates suggest that between 1990 and 2015, overall age-standardised smoking prevalence has decreased from 27% to 19% in the UK, 22% to 13% in the US, and 24% to 15% in Australia.[1,51] Over the same time period, however, many LMICs did not benefit from such steep declines in smoking prevalence: prevalence decreased from 26% to 22% in China, remained steady at 25% in Russia, and rose from 22% to 26% in Indonesia.[1,51] These prevalence estimates mask sometimes divergent trends among men and women, especially in LMICs, where smoking remains much more prevalent among men than women.

Tobacco control policies have been accelerated by efforts at both the national and international levels. Internationally, the most significant effort to develop and enforce tobacco control policies has been the World Health Organisation's (WHO) 2003 adoption of the Framework Convention on Tobacco Control (FCTC), which entered into force in 2005.[52,53] An essential aspect of the FCTC is its underpinning by the MPOWER package, a set of tobacco control policies aimed at raising taxes and tobacco prices, banning advertising and promotion, protecting individuals from second-hand smoke, warning of the dangers of tobacco, offering assistance for tobacco cessation, and monitoring the tobacco epidemic (and the policies designed to curb it).[54] Subsequent monitoring of the uptake and enforcement of these policies indicates that progress has not been even across countries,[55] nor have the specific policies within the MPOWER package been equally enforced.[56] Where such policies have been enforced, they have been associated with considerable reductions in the prevalence of smoking.[57,58]

To build on this progress, tobacco control policies have been embedded in other international targets, including the WHO's 25x25 goal (a 25% reduction in premature non-communicable disease [NCD] mortality by 2025) and the Sustainable Development Goals (SDGs), particularly SDG 3.a: to "strengthen the implementation of the World

Health Organisation Framework Convention on Tobacco Control in all countries, as appropriate".[59] These efforts notwithstanding, in a number of LMICs today, tobacco control policies remain limited, if in effect at all, and require substantial further efforts where they are in place.[60] Smoking remains the leading cause of healthy life lost among men globally (an estimated loss of 124 million disability-adjusted life-years, or DALYs, in 2016), and a leading cause of such health loss among women.[61]

1.5 The need for continued research on smoking in diverse populations globally

In addition to the heterogeneous nature of smoking patterns and products smoked around the globe, the relationship between smoking and mortality is further complicated by population differences in typical exposures, outcomes, and genetic architecture, potentially limiting the generalisability of findings from one population to another. For example, smoking was found to be a major driver of tuberculosis-related mortality in India – causing nearly 10 times as many smoking-related deaths from tuberculosis as from lung cancer—a major divergence from what would have been predicted by studies of Western populations, where tuberculosis is uncommon.[62] Where such studies have been conducted in LMICs, however, they have generated simple public health messages rooted in clear evidence: tobacco kills as many as one in four men in countries like China,[9] India,[62] Bangladesh,[63] and South Africa,[64] and tobacco-attributable mortality may still be on the rise in such populations.

Local research produces local evidence, which allows for clear, population-specific policy on how best to prevent and control tobacco use. Policies grounded in locally-produced evidence are more credible for local policymakers, are more relevant to the populations they are designed to protect, and are better able to guide local disease control priorities. Furthermore, local evidence is critical for informing national, regional, and global comparative risk factor assessments, such as the Global Burden of Disease

(GBD) Study's quantification of tobacco-attributable mortality (and indirect estimates of risk-attributable mortality). These efforts are vital for tracking health information across populations globally, and are particularly important for assessing progress on national and international health targets such as the health-related Sustainable Development Goals.[65] Therefore, to inform country-specific disease control priorities, to inform global comparative risk factor assessments and progress on health policy goals, and to inform individual smokers and prospective smokers of the hazards of smoking and benefits of quitting, population-specific evidence is needed.

To address this evidence gap, in the late 1990's and early 2000's, researchers at the Clinical Trial Service Unit and Epidemiological Studies Unit (CTSU), now part of the Nuffield Department of Population Health (NDPH), worked with local collaborators to support the conduct of large prospective studies in China, India, Russia, Mexico, and Cuba, each consisting of more than 100 000 participants, and resurveying a subset to characterise long-term behaviours. Using these resources, the hazards of smoking have previously been assessed in China,[9,66] and are currently being assessed in India[67] and Russia[68], alongside other major risk factors. This thesis reports the hazards of smoking, and benefits of quitting, in the Cuban and Mexican prospective cohorts.

1.6 Outline of the thesis

There were an estimated ~50 million smokers in Latin America and the Caribbean in 2015.[1,51] However, the relevance of tobacco to mortality has not yet been established reliably with direct evidence from local populations in any country in this region (discussed in **Chapter 2**). Therefore, the predominant focus of this thesis is on the relationship between smoking and mortality in large prospective cohorts in Cuba (**Chapter 3**) and Mexico (**Chapter 4**), with supporting evidence from a prospective study in the US with a large number of Mexican Americans (**Chapter 5**). Finally, the relevance

of these prospective findings to other populations globally is discussed in the context of prior prospective evidence, as well as cross-sectional evidence from the Global Adult Tobacco Surveys in 30 low- and middle-income countries, in **Chapter 6**.

Chapter 2: Review of the literature

2.1 Chapter aims

As this thesis is focused on the relationship between smoking and mortality in three large, prospective studies conducted in Latin American populations (including those resident in the US), a systematic review was conducted to identify all of the relevant evidence previously reported on smoking and major health outcomes in these populations. The specific aims of this chapter are:

1) To report the methodology used to identify the relevant literature on smoking and major health outcomes in Latin American populations
2) To report the main findings of case-control studies on smoking and major health outcomes identified by this systematic review
3) To report the main findings of prospective cohort studies on smoking and major health outcomes identified by this systematic review
4) To summarize the available evidence and report gaps in the literature that the studies in this thesis may address

2.2 Methodology

A literature search was conducted to identify reports reporting on the relationship between tobacco smoking and major health outcomes (ie, fatal or life-threatening) in Latin American populations (including individuals of Latin American descent residing elsewhere, eg, Mexican Americans in the United States). I conducted a search in August 2019 for reports in the databases Global Health (1973-2019 week 32), Embase (1974-present), and Medline (including Epub ahead of print, in-process, and other non-indexed citations; 1946-present). The search relied on keywords and Medical Subject Headings (MeSH) terms related to the exposure of interest (smoking/tobacco), study design

(prospective cohort/case-control), and Latin American populations (see Table 2.1 for the specific search terms used).

There were 6085 reports identified by the search, of which 3994 were unique publications across databases (Figure 2.1). The titles and abstracts of each report were then screened according to pre-specified selection criteria: reports were required to report the relationship between smoking and a fatal or life-threatening health outcome of interest (eg, cardiovascular, respiratory, or neoplastic disease); had to be published in English or Spanish; had to report the specific methodology used (eg, what was adjusted for in analyses, if anything); had to pertain to adults (ie, those aged≥18 years); had to be a population-based prospective study or a case-control study; and the study had to report estimates for the specific population or populations of interest (eg, Mexican Americans, Brazilians). Finally, only studies large enough to report reliable estimates are reported in this chapter, defined (arbitrarily) as ≥1000 participants for a prospective study, or ≥1000 cases for a case-control study.

After exclusions based on screening of titles and abstracts, the remaining 241 relevant or potentially-relevant reports underwent full-text review. A total of 92 relevant reports were identified, of which 13 were unique studies described in 17 reports that met the criteria for study size. The flow diagram of the literature review process is shown in Figure 2.1.

2.3 Case-control evidence

12 reports describing 8 unique case-control studies including ≥1000 cases were identified (described in Table 2.2). Studies are described according to the outcome of interest. Relevant details about smoking patterns are presented where reported. However, few studies reported smoking characteristics beyond current (or ever) smoking prevalence.

2.3.1 Heart disease

Five reports describing 4 unique studies reported the association between smoking and non-fatal myocardial infarction (MI) or other ischaemic heart disease (IHD). Two reports on one unique study reported the association between smoking and MI in Costa Rica. Rossi *et al.* (2011) reported the risks associated with smoking among 2094 MI cases and 2094 controls (matched by age, sex, and area) recruited in 1994-2004 from 34 counties in Costa Rica.[69] Smoking prevalence was significantly higher among cases than among controls (41% and 22%, respectively; p<.001), and among current smokers, cases tended to smoke more heavily, have a longer duration of smoking, and to start smoking at a younger age than controls on average. The OR for current versus never smoking was 2.71 (95% CI 2.22-3.31), although for ex-smoking versus never smoking it was 1.01 (0.84-1.21). Compared with never smokers, the ORs associated with smoking <15, 15-24, or ≥25 cigarettes/day were 1.80 (1.42-2.28), 3.05 (2.34-3.99), and 8.18 (5.47-12.21), respectively. However, among those who had quit <1, 1-4, and ≥5 years previously, the ORs were 3.41 (1.12-10.40), 1.39 (0.93-2.08), and 0.93 (0.77-1.13), respectively, indicating little association with MI for those who quit smoking at least 1 year previously. Similar results were reported by Sotos-Prieto *et al.* (2016) for current versus never smoking when stratifying by genetic risk scores for MI.[70]

Piegas *et al.* (2003) reported the risks associated with smoking and other risk factors among 1279 cases and 1279 age- and sex-matched controls recruited from 104 hospitals in 51 cities in Brazil in 1997-2000.[71] Among cases, 34% had never smoked, 20% had quit at least 6 months previously, 4% smoked <5 cigarettes/day, and 42% smoked ≥5 cigarettes/day. In univariate analyses comparing smokers to never smokers, the ORs for having quit, smoking <5 cigarettes/day, and smoked ≥5 cigarettes/day were 1.11 (0.90-1.36), 1.52 (1.02-2.27), and 3.09 (2.55-3.76), respectively. After further adjusting for a number of other factors, the associations were similar, if somewhat

stronger (4.90 [3.53-6.81] for ≥5 cigarettes/day and 2.07 [1.14-3.76] for <5 cigarettes/day, compared with never smokers).

The multi-country INTERHEART study of MI reported results for 234 cases and 178 controls in Argentina, 313 cases and 364 controls in Brazil, 275 cases and 550 controls in Colombia, 322 cases and 672 controls in Chile, 85 cases and 107 controls in Guatemala, and 8 cases and 17 controls in Mexico, each matched by age and sex.[72] Overall, 48% of controls were ever (ie, current or former) smokers in the Latin American countries, and the OR for ever versus never smoking was 2.31 (95% CI 1.97-2.71), resulting in an estimated population attributable fraction associated with smoking of 38%. The ORs associated with ever versus never smoking varied by country: 2.3 (1.5-3.7) in Argentina, 2.4 (1.7-3.4) in Brazil, 3.1 (2.3-4.2) in Chile, and 1.4 (1.0-2.0) in Colombia (too few cases occurred in Guatemala and Mexico to be estimated separately). However, specific smoking patterns were not reported by country, so the likely reasons for these different associations are unclear. Another international case-control study of MI recruited 1060 cases and 1071 controls in Argentina, 323 cases and 314 controls in Cuba, 200 cases and 200 controls in Mexico, and 266 cases and 264 controls in Venezuela between 1992-1998 (matched by age, sex, and medical centre).[73] Current smoking prevalence among controls was 27% in Argentina, 39% in Cuba, 27% in Mexico, and 22% in Venezuela. Adjusting for age and sex, the ORs for current versus not currently smoking (ie, former or never smoking) were 2.40 (1.96-2.93) in Argentina, 2.04 (1.47-2.81) in Cuba, 2.12 (1.36-3.32) in Mexico, and 2.50 (1.68-3.73) in Venezuela. Again, however, specific smoking patterns were not described.

2.3.2 Cancer

Six reports describing four studies were identified that reported the association between smoking and cancer. A large study conducted by de Souza Moura *et al.* (2014) including

231 102 patients registered in the Cancer Hospital Registries in Brazil between 1998-2011, of whom 204 131 were cancer cases relating to 32 sites (each of which is known to be associated with tobacco use) and 26 971 were controls diagnosed with non-melanoma skin cancer (which is not related to tobacco use).[74] The smoking prevalence among controls was 31% overall (40% among men and 19% among women) and among cases was 48% (64% among men and 35% among women). After adjusting for age, sex, and alcohol consumption, smoking was strongly associated with a number of cancers, resulting in ORs of 8.1 (95% CI 6.2-10.6) for cancer of the piriform sinus, 7.9 (7.4-8.3) for the bronchi and lung, 6.3 (5.9-6.8) for the larynx, 5.7 (4.8-6.7) for the hypopharynx, and 5.2 (4.6-5.9) for the oropharynx, among others. Results were generally similar between men and women.

Two case-control studies of smoking and head and neck cancer were also conducted in Brazil. A study conducted by Kfouri *et al.* (2018) was conducted in five Brazilian cities from three regions in 1998-2003, during which time 1594 cases and 1292 controls were recruited. Current smoking prevalence among controls was 32% in Goiania, 39% in Rio de Janeiro and Sao Paulo, and 30% in Pelotas and Porto Alegre.[75] The head and neck cancer ORs for current and former versus never smoking were, respectively, 13.6 (95% CI 6.4-28.6) and 3.5 (1.6-7.4) in Goiania, 6.6 (4.5-9.7) and 2.3 (1.6-3.5) in Sao Paulo and Rio de Janeiro, and 8.8 (4.7-16.2) and 3.0 (1.6-5.7) in Pelotas and Porto Alegre. Another study (with some overlapping participants from the prior study), conducted by Boing *et al.* (2011), was conducted in 1998-2006 in Sao Paulo, Brazil, during which 1017 incident cases of oral, pharyngeal, and laryngeal cancer were recruited, along with 951 age- and sex-matched controls.[76] Smoking prevalence was 94% among cases and 64% among controls. Adjusting for age, sex, education, and alcohol consumption, the head and neck cancer ORs for smoking <21.8, 21.8-37, 37.1-57.7, and >57.7 pack-years were 2.67 (1.84-3.88), 6.25 (4.25-9.18), 7.92 (5.34-11.75), and 7.78 (5.22-11.60).

However, no information was given for the individual contributions of amount smoked and duration of smoking to analyses by pack-years. Another study of this population by Ferreira Antunes *et al.* (2013), slightly expanded to include 1144 cases of oral and oropharyngeal cancer recruited to 2008, reported ORs for ever versus never smoking of 3.50 (2.76-4.44), adjusting for age, sex, education, and alcohol.[77] (A similar report by Szymanska *et al.* [2011][78] also reported on these overlapping Brazilian populations, pooled with small numbers of cases and controls from Cuba and Argentina reported jointly, which is not described here.)

Finally, Sánchez-Zamorano *et al.* (2011) reported the relevance of smoking and other risk factors to breast cancer in a population-based case-control study conducted in three areas of Mexico between 2004-2007.[79] A total of 1000 incident cases of breast cancer were matched to 1074 controls by age, region, and health institution. Prevalence of ever smoking was 25% among cases and 21% among controls. After adjusting for age, type of healthcare, region, socioeconomic status, breastfeeding, body mass index, and a number of other lifestyle and familial risk factors, the OR for never versus ever smoking (ever smoking being the reference group) was 1.20 (95% CI 0.83-1.73) among premenopausal women and 0.82 (0.57-1.18) among postmenopausal women. As in other studies reported in this chapter, no further details relating to smoking habits were reported, limiting the interpretation of these and other findings.

2.4 Prospective evidence

Six reports describing 5 unique prospective cohort studies including ≥1000 participants were identified (described in Table 2.3). As there were relatively few studies (and they spanned several decades), the reports are described in order of their publication date, beginning with those published most recently. Two reports described smoking characteristics beyond whether or not a participant was a current smoker at recruitment

(limited, however, to the amount smoked/day), and no reports reported results for 'usual' (as opposed to baseline) smoking behaviours. A study by Llibre-Guerra *et al.* (2015) described 2916 adults aged ≥65 years who were recruited in Havana and Matanzas, Cuba, in 2003-2007 and resurveyed in 2008-2011 for incident cases of stroke, originally conducted as part of the 10/66 study.[80] The study sample comprised the 2703 patients who had not previously been diagnosed with stroke at recruitment and who agreed to be resurveyed (28 participants declined). At resurvey, participants with self-reported or suspected stroke were visited by a vascular neurologist to confirm incident cases, at which point 82 incident cases were confirmed (40 among men and 42 among women). Adjusting for sex, education, and 5-year age group, the hazard ratio for incident stroke was 1.3 (95% CI 1.2-1.4) among ever smokers compared with never smokers (annual stroke incidence/1000 person-years of 8.6 [6.4-11.7] and 6.9 [5.1-9.4], respectively).

A study by de Oliveira *et al.* (2015) reported the relevance of smoking and other risk factors to mortality among participants of the Bambui Cohort Study of Ageing in Brazil.[81] In 1997, adults aged ≥60 years were recruited in Bambui City, Minas Gerais State, Brazil, resulting in 1606 participants (of whom 1382 had sufficient information to be included in prospective analyses). Following recruitment, participants participated in face-to-face interviews annually, with a health examination additionally conducted in certain years. Deaths occurring in 1997-2002 were reported by next of kin during the annual interview and through the Brazilian System of Information on mortality, with death certificates obtained for 99% of deaths. Over follow-up (mean 5 years), 241 participants died and 37 participants' vital status could not be confirmed. At recruitment, 17% of participants were current smokers, and, following adjustment for age and sex, the hazard ratio for current versus not current smoking was 1.31 (95% CI 0.94-1.81) (further adjusted for hypertension and diabetes, the hazard ratio was 1.38 [0.99-1.92]). The

estimated population attributable fraction (PAF) for smoking-related mortality was 5.6% (0.6-10.2%) in the further-adjusted model.

Otiniano and colleagues (2003) reported the association between smoking and other risk factors to self-reported myocardial infarction among participants in the Hispanic Established Population for the Epidemiological Study of the Elderly (HEPESE) cohort.[82] A population-based study of Mexican Americans aged ≥65 years from Texas, California, New Mexico, Colorado, and Arizona, the cohort is comprised of 3050 participants recruited in 1993-1994 with follow-up surveys in 1995-1996, 1998-1999, and 2000-2001. During a household interview, respondents answered questions about lifestyle behaviours and medical history, with similar procedures at each follow-up interview. At recruitment, 42% of participants were ever smokers. Adjusting for age, sex, education, body mass index, living arrangement (living alone or with others), diabetes, hypertension, and stroke, the odds ratio for incident MI among ever versus never smokers was 0.68 (95% CI 0.46-0.99), 1.00 (0.71-1.40), and 1,44 (0.98-2.12) at the first, second, and third follow-up interviews, respectively (participants with/without heart attack at each follow-up: 139/2772, 170/2122, and 121/1679, respectively). Other factors such as amount smoked, age at smoking initiation, or age at smoking cessation were not taken into account.

A report by Wei *et al.* (1996) reported results from the San Antonio Heart Study, which focused on risk factors for cardiovascular and all-cause mortality among Mexican Americans.[83] Recruitment occurred in two phases (1979-1982 and 1984-1988), during which households were randomly sampled and all men and non-pregnant women aged 25-64 years invited to participate. A total of 2629 Mexican Americans were recruited (along with 1136 non-Hispanic Whites for comparison). After an average of 8 years of follow-up, vital status was ascertained for 98% of participants, for which deaths were

identified by next of kin and review of death certificates. Although current cigarette smoking prevalence was similar among Mexican Americans (31%) and non-Hispanic Whites (32%), the average cigarettes/day among Mexican Americans was lower than among non-Hispanic Whites (11 [95% CI 5-17] and 20 [11-29] cigarettes/day, respectively). Adjusting for age and sex, the all-cause mortality RR among current versus never smokers was 1.6 (1.1-2.5) among Mexican Americans and 2.5 (1.4-4.7) among non-Hispanic Whites. Further adjusting for socioeconomic status, diabetes, health status, body mass index, hypertension, hypercholesterolemia, high-density lipoprotein cholesterol, alcohol consumption, and exercise, the RRs were little changed (1.8 [1.2-2.6] among Mexican Americans and 2.3 [1.2-4.3] among non-Hispanic Whites). The fully-adjusted RRs for cardiovascular mortality were similar between the two groups: 2.0 (1.0-4.2) among Mexican Americans and 2.4 (0.8-6.6) among non-Hispanic Whites (though the estimates lacked precision).

Finally, a report by Sorlie *et al.* (1982) reported results from the Puerto Rico Heart Health Program), which was based on the Framingham Study and focused on risk factors for cardiovascular disease among Puerto Rican men.[84] Recruitment took place in 1965, during which a house-to-house census in seven areas was conducted, and all men aged 45-64 years were invited to participate (>12 000 invited overall). A total of 9824 men agreed to participate and underwent clinical examination, in addition to answering questions about lifestyle behaviours and medical history. Although participants were re-examined every 2-3 years, resurvey information was not reported. At recruitment, 23% of participants had never smoked, 24% had formerly smoked, 9% smoked pipes or cigars, and 44% were current cigarette smokers (22% smoked <1 pack/day, 14% smoked 1 pack/day, and 8% smoked >1 pack/day; on average, 8, 20, and 37 cigarettes/day, respectively). In rural areas, among those without coronary heart disease (CHD) at entry, the age-adjusted rate of incident myocardial infarction was 0.8% among

never smokers, 1.2% among smokers who quit >4 years prior, 1.5% among smokers who quit 0-4 years prior, and 2.6% among continuing cigarette smokers. In urban areas, the age-adjusted rate of incident MI was 1.9% among never smokers, 2.5% among smokers who quit >4 years prior, 2.4% among smokers who quit 0-4 years prior, and 3.8% among continuing cigarette smokers. Adjusting for age, education, physical activity, blood pressure, haematocrit, blood glucose, cholesterol, alcohol, and caffeine consumption, the odds ratio for versus never-smoking was 3.40 in rural areas (95% CI 2.70-4.10), and 1.82 (1.48-2.16) in urban areas (in 198 incident cases of MI among 8081 participants at risk). Another report examining the same population showed similar, if somewhat weaker, associations between smoking and non-MI incident heart disease.[85]

2.5 Summary of findings

This literature review of the case-control and prospective studies of smoking and major health outcomes in Latin American populations demonstrates both the paucity of the evidence and the limitations of the available evidence. A total of 17 relevant reports reflecting 13 unique studies were identified, of which 8 were case-control studies including ≥1000 cases and 5 were prospective studies including ≥1000 participants. The studies identified were generally small (the largest prospective study included just over 9000 participants), no study accounted for 'usual' smoking behaviours (ie, changes in amount smoked or smoking cessation), and only one prospective study and two case-control studies reported on the benefits of quitting in these populations. Furthermore, these studies were generally limited to current or ever versus never smoking, with little information on amount smoked, age started smoking, products smoked, or cessation habits (and such information was heterogeneous when reported).

Overall, the evidence on the relationship between smoking and health in Latin America is extremely limited. Where it has been produced, the relationships for relevant

outcomes are generally weaker than those reported in high-income populations (though few outcomes were reported). For example, a study of older adults in Brazil (aged ≥60 when recruited in 1997) reported a current versus not current smoker all-cause mortality RR of 1.31 (95% CI 0.94-1.81),[81] which is much lower than the current versus never all-cause mortality RRs of approximately 3 for contemporary UK[35] and US[4] populations. However, the precision of the estimates from these relatively small studies is limited, and given the imprecise information reported on smoking habits, it is not clear why these relationships may have been weaker. A reliable and representative summary estimate of the relationship between smoking and incident or fatal disease among Latin American populations could not be produced due to the substantial heterogeneity of both the exposures and outcomes reported, the low quality of the data (offering little information about smoking behaviours), and the limited statistical power of included studies.

In summary, to date, there has not been a large, prospective study on the relationship between smoking and all-cause and cause-specific mortality in a Latin-American population. As smoking behaviours are likely to change over time (particularly in terms of amount smoked and continued smoking versus cessation), and detailed information is needed on the relevance of different smoking habits to mortality risk, studies collecting comprehensive smoking information at recruitment and at resurvey are needed. Perhaps most importantly, population-specific information on the hazards of smoking *and* benefits of quitting is needed to estimate and communicate the health risks of smoking in different populations, alongside the substantial benefits of quitting. Therefore, the overarching aim of this thesis is to characterise the smoking (and quitting) behaviours and their relevance to mortality in three large prospective studies from Cuba, Mexico, and the United States. The detailed aims of each of these studies are described at the outset of each chapter.

Table 2.1: Literature review search terms

Keyword	Text words and MeSH terms	
Smoking	Smok$	Smoking/
	Tobacco$	Tobacco/
		Tobacco smoking/
Cohort study	Cohort$	Cohort study/
	Longitudinal$	Prospective$
Case-control study	Case control$	Case control study/
Latin America	Latin America$	Latin America/
	Mexic$	Bolivia$
	Cuba$	Dominica$
	Brazil$	Hondura$
	Colomb$	Paraguay$
	Argentin$	Nicaragua$
	Peru$	El Salvador$
	Venezuela$	Costa Rica$
	Chile$	Panama$
	Ecuador$	Puerto Ric$
	Guatemala$	Uruguay$
	Haiti$	Jamaica$
	Trinidad$	Tobag$
	Guyan$	Surinam$
	Bahama$	Beliz$

Searched (Smoking) AND (Cohort study OR case-control study) AND (Latin America)

Table 2.2: Case-control studies of smoking in Latin American populations (≥1000 cases), in the order described in text

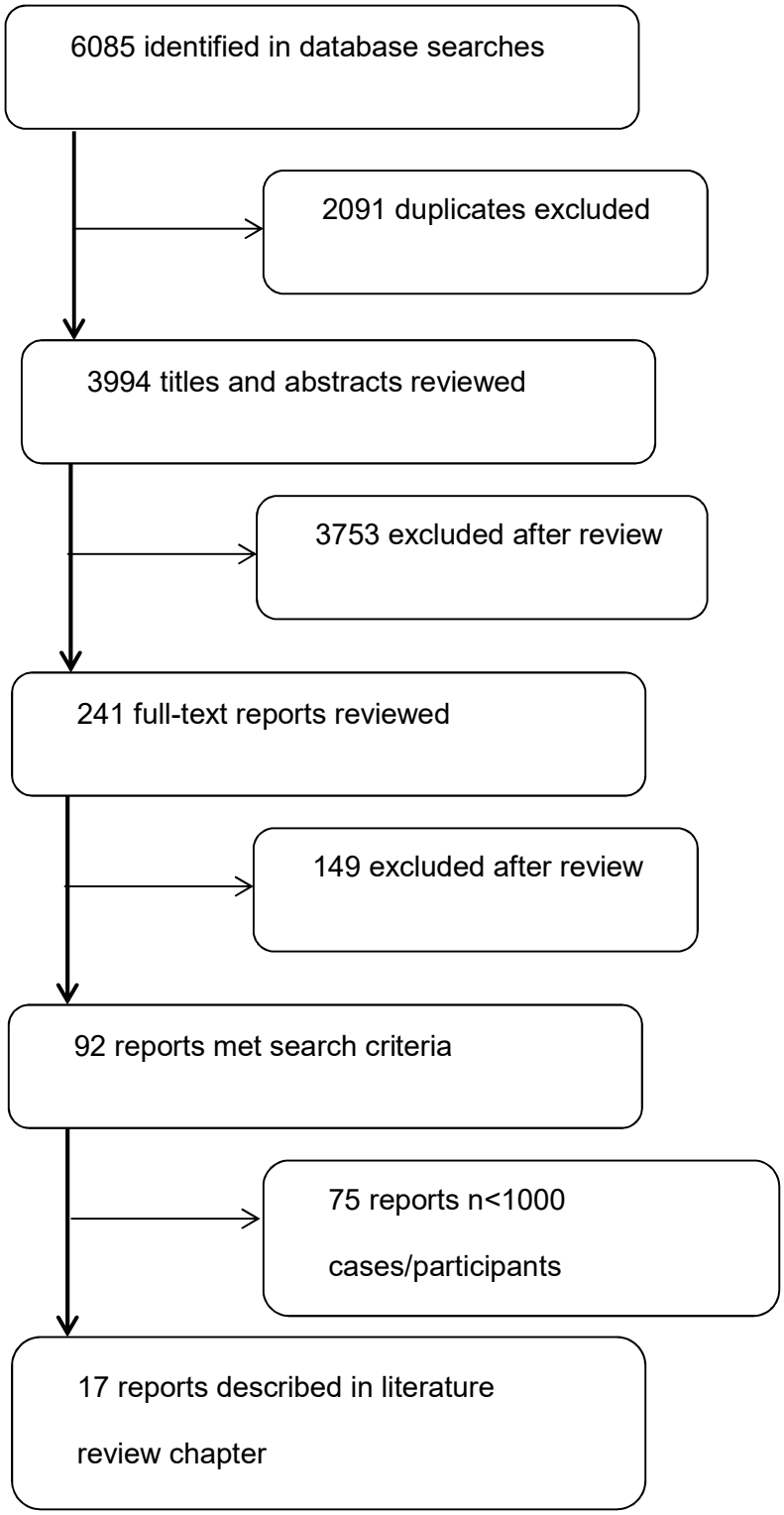
First author, year published	Location	Study years	Sample size (cases)	Male (%)	Outcome	Smoking prevalence (among controls, %)	Other smoking characteristics
Rossi 2011	Costa Rica	1994-2004	2094	74	MI	22 (current)	Mean 13 cigarettes/day
Piegas 2003	Brazil	1997-2000	1279	60	MI	46 (daily)	91% smoked ≥5 cigarettes/day
Lanas 2007	Argentina, Brazil, Chile, and Colombia	1999-2003	1237	75	MI	48 (ever)	-
Ciruzzi 2003	Argentina, Cuba, Mexico, and Venezuela	1992-1998	1849	74	MI	45 (daily)	-
De Souza Moura 2014	Brazil	1998-2011	204 131	46	Multiple cancers	31 (current)	-
Kfouri 2018	Brazil	1998-2003	1594	85	Head and neck cancer	34 (current)	37% were ex-smokers
Boing 2011	Brazil	1998-2006	1017	84	Head and neck cancer	64 (daily)	23% smoked >37 pack-years
Sanchez-Zamorano 2011	Mexico	2004-2007	1000	0	Breast cancer	21 (ever)	-

Studies predominantly resulting on overlapping populations are described only once (but additional reports based on these underlying populations are described in text)

Table 2.3: Prospective studies of smoking in Latin American populations (≥1000 participants), in order of publication date

Study Name	Location	Study years	Sample size	Male (%)	Outcome	Smoking prevalence (%)	Other smoking characteristics
Matanzas	Cuba	2003-2011	2802	35	Stroke	-	-
Bambui	Brazil	1997-2002	1382	38	All-cause mortality	17 (current)	-
HEPESE	US (Mexican Americans)	1993-2001	3050	43	Non-fatal MI	42 (ever)	-
San Antonio Heart Study	US (Mexican Americans)	1979-1988	2629	41	All-cause and CHD mortality	31 (current cigarette)	Mean 11 cigarettes/day
Puerto Rico Heart Health Study	Puerto Rico	1965-1973	9150	100	MI/CHD death	44 (current cigarette)	Mean 17 cigarettes/day

Figure 2.1: Flow chart of selection process for reports included in the systematic review



Chapter 3: Smoking and mortality in the Cuba Prospective Study

3.1 Chapter aims

The overarching goal of this chapter is to report the hazards of smoking and benefits of quitting in the Cuba Prospective Study. Specific aims are:

1) To assess the prevalence of smoking patterns across birth cohorts and by sex
2) To characterise the association between smoking behaviours and all-cause mortality
3) To determine whether other factors may modify the association between smoking and all-cause mortality
4) To characterise the association between smoking behaviours and cause-specific mortality
5) To estimate the smoking-attributable mortality in Cuba
6) To interpret findings in the context of prior evidence of smoking in Cuba and other populations globally

3.2 Baseline survey

3.2.1 Study design

The Cuba Prospective Study is a large, prospective, population-based study in Cuba consisting of 146 556 adults aged 30 or older recruited between 1996 and 2002 in five Cuban provinces.[86] The study was originally established to determine the relevance of tobacco to premature mortality in Cuba, but also included information on other risk factors including alcohol consumption, adiposity, and blood pressure. Although these risk factors are well-studied in high-income populations, their health effects (especially in relation to mortality) are not well-established in other settings, and their relative importance is expected to change over time. At 2010 mortality rates, approximately one in four Cuban men and one in six Cuban women will die between the ages of 35-69,

primarily from vascular diseases and cancer.[87] The Cuba Prospective Study was designed to provide information on the relevance of these risk factors to premature death (particularly from these major causes) in Cuba for at least the next several decades, and may generate insights into the relative importance of these risks in other populations. The study was designed to be sufficiently large to provide reliable estimates not only overall, but within sub-groups of interest (eg, men and women) and at different levels of these risks (eg, smoking in those who do or don't drink). The study was conducted by collaborators in Cuba (at the National Institute of Cardiology in Havana, the National Institute of Hygiene, Epidemiology, and Microbiology, and the Municipal Centres of Hygiene and Epidemiology of Pinar del Río, Matanzas, and Camagüey) and in the UK (the Clinical Trial Service Unit and Epidemiological Studies Unit [CTSU], University of Oxford). Research ethics approval for the study was granted by the National Institute of Cardiology and Cardiac Surgery Ethics Committee in Havana (ref: 0404134).

3.2.2 Study population

The baseline survey was conducted in five of Cuba's 14 provinces (at the time of data collection; there are now 15 provinces after an administrative change in 2011): Pinar del Río, Matanzas, Camagüey, La Habana, and Ciudad de La Habana. These provinces were selected because they represented geographic diversity, consisted of relatively settled populations that were likely to remain in the area, and local staff had both the capacity and commitment to carry out the survey. Within each province, a computer-generated random allocation sequence was used to randomly select primary care clinics for participation (in Cuba, these publicly-funded medical clinics each provide care for approximately 150 families in the surrounding catchment area). At participating clinics, local staff (typically the family doctor, but occasionally other trained health workers) visited households under their care and invited all adults aged 30 or older to participate

in the study. Of those registered with selected medical clinics and eligible to participate, 74% were ultimately recruited into the study: 64 743 men and 81 813 women.

3.2.3 Data collection

During the household visit, participants gave written consent, followed by their national identification number and information on age, sex, education, occupation, smoking and drinking habits, disease history, and medication use. Participants were asked separately about cigarette and cigar smoking, and beer and rum drinking, to estimate their individual and combined health effects. Blood pressure was measured twice (once towards the beginning and once towards the end of the interview) using a manual sphygmomanometer while the participant was seated. Participants' height and weight were measured on a separate visit to their local clinic, because at the time of the survey the mechanical scales typically used were not easily portable. Of those who completed the interview, 80% subsequently had anthropometric measurements taken. Because the survey was conducted during The Special Period, a time of severe deprivation in Cuba,[88] the survey was deliberately short in order to fit on a single sheet of paper, which limited the breadth and depth of information collected (see Appendix E for baseline questionnaire in English and Spanish).

3.3 Resurvey

A subset of baseline participants in two municipalities in the province of Matanzas (Colón and Jagüey Grande) were invited to be resurveyed in 2006-2008, approximately 6 years after the baseline survey. The resurvey was conducted primarily to correct for within-person variation in physical measurements (eg, blood pressure), as well as to quantify behaviour changes over time (eg, starting or quitting smoking, changes to cigarettes smoked per day).The resurvey procedures and instrument were the same as at baseline. A total of 24 345 participated in the resurvey physical measurements;

however, updated questionnaire information, including smoking habits, was only collected for participants in Jagüey Grande (4102 participants).

3.4 Follow-up

Participants have been followed-up to 1 January, 2017 through the Cuban Public Health Ministry's national death registry. Mortality records were matched electronically to study participants through national identification numbers, name, sex, and date of birth. Issues with matching (eg, matches where the identification number matched but the sex did not) were resolved in Cuba by reviewing the original death certificate. Participants who emigrated were lost to follow-up (but not those who moved elsewhere within Cuba); however, the proportion of Cubans who emigrate is estimated to be <1% annually. As of writing, 241 participants (0.2%) are known to have emigrated. Causes of death were coded according to the International Classification of Diseases (ICD) 9th edition (for deaths during 1996-2000) and 10th edition (2001-present). In Cuba, causes of death are certified for all adults, leading to few deaths being ill-defined. The ICD-9 and ICD-10 codes for the causes of death reported are given in Table 3.1. In part because electronic health records are not yet common in Cuba, follow-up for non-fatal events has not yet been possible for the Cuba Prospective Study.

3.5 Methods

3.5.1 Assessment of smoking

Smoking behaviour was self-reported by participants according to their responses to the standardised questionnaire. Participants were classified as ever-regular tobacco smokers if they answered "yes" to the question "Have you EVER smoked regularly (i.e. on most days for at least a year)?" Those who were ever-regular smokers then answered questions about the age at which they first started smoking tobacco regularly, whether or not they had ever stopped smoking for a year or more, and how many

cigarettes and/or cigars they smoked per day in the last month (number of cigarettes and cigars smoked per day were captured as two separate questions). Participants who reported smoking at least one cigarette and/or cigar per day in the past month were classified as current smokers. Ever-regular smokers were classified as ex-smokers if they reported smoking no cigarettes or cigars in the past month (and answered “yes” to having ever regularly smoked), and were subsequently asked the age at which they last stopped smoking. Due to the structure of questionnaire, those who smoked cigarettes less than daily would be classified as either never smokers (if they reported never having smoked regularly), or ex-smokers (if they reported ever-regular smoking, but smoking zero cigarettes or cigars daily in the past month). For ex-smokers, the age at which they began smoking regularly and last stopped smoking was recorded, but amount previously smoked was not. A subset of participants were asked identical questions at resurvey, which allowed for the quantification of changes in smoking patterns over time.

3.5.2 Selection of covariates

In each set of prospective analyses in this thesis, confounders of the association between smoking and mortality were initially identified through *a priori* knowledge based on literature and through descriptive analyses, and subsequently tested empirically as a sensitivity analysis. The main set of covariates included in prospective analyses aimed to adjust for the three main areas of potential confounding for the relationship between smoking and mortality: demographic factors (age and sex), socioeconomic factors (level of education and area of residence), and a common behavioural factor (alcohol consumption). Alcohol was included as a potential confounder because smoking and drinking are strongly correlated in many populations (including those studied in this thesis). Other potential lifestyle-related confounders were not included in the main analyses, however, in part because the effect smoking may have on these factors could cause them to be on the causal pathway to mortality. For example, smoking may directly

affect important risk factors, such as body mass index (BMI),[89,90] and in such cases these factors may mediate the relationship between smoking and mortality.

In order to maximise comparability of findings to existing literature, and to findings from the other cohorts in this thesis, the selected confounders were included in regression models for the main analyses regardless of whether their inclusion had a material effect on the relationship between smoking and mortality. The effect of smoking across different levels of potential confounders was then assessed individually to determine whether the relationship between smoking and mortality was materially affected by any of these factors. The findings of this sensitivity analysis are presented in the results section.

3.5.3 Participant exclusions

Of the 146 556 participants recruited into the study, those aged ≥90 (n=703) or with missing information on smoking or other covariates of interest (n=1933) were excluded. Of the remaining participants, 22 991 were aged 70 or older at recruitment and were excluded from the main analyses (but are shown in Appendix A for ages 70-89). The main analyses focused on the 120 929 participants aged 30-69 for two main reasons: to assess effects of smoking on premature death (which should be more avoidable than deaths at older ages),[46] and to improve the accuracy of cause-specific mortality (as causes of death tend to be less reliable at older ages).[91] Furthermore, the reliability of long-term smoking behaviour was uncertain in older participants in this study, which is further explored in Appendix A and the discussion section of this chapter.

To limit the effect of reverse causality, whereby smoking may have caused disease which in turn led to smoking cessation, participants with chronic disease at recruitment were also excluded from the main analyses of ex-smokers versus current or never smokers (n=1644 ex-smokers, 7500 never smokers, and 4664 current smokers).

Chronic disease was defined as ischaemic heart disease (IHD), stroke, cirrhosis, cancer, chronic kidney disease, or chronic obstructive pulmonary disease (COPD). The effect of including participants with chronic disease is shown in Appendix A.

3.5.4 Statistical analysis

Cross-sectional analyses of the prevalence of smoking behaviours are unadjusted, and are presented by sex and birth cohort (where relevant). For prospective analyses, Cox regression was used to calculate mortality rate ratios (RRs) and accompanying 95% confidence intervals (CIs) according to categories of smoking at recruitment. Follow-up was censored at the date of death, the end of the risk period under consideration (eg, age 70 for most analyses), date of loss to follow-up, or at the end of the follow-up period (1 January 2017). Analyses were stratified by age-at-risk (in 5-year age groups), sex, and province (meaning that the proportional hazards assumption was not assumed between 5-year age groups, sex, or province, but was assumed within these groups). The main analyses also adjusted for formal education completed (<primary, primary, secondary, high school/technical college, or university) and frequency of alcohol consumption (never, <weekly, ≥weekly).

Analyses of current versus never smoking were assessed overall, and by age at smoking initiation and amount smoked in order to examine the shape of these associations with premature mortality. Because of the relatively few cigar smokers identified in cross-sectional analyses, and due to the *a priori* expectation that the hazards of smoking cigarettes would be considerably larger than the hazards of smoking cigars, the main analyses considered those who smoked cigarettes irrespective of additional cigar smoking. However, sensitivity analyses consider these tobacco products separately. The association between cigarettes smoked per day was assessed by categorising smokers according to their reported number of cigarettes smoked per day

at baseline: <10, 10-19, 20, or >20 cigarettes per day (regardless of cigars smoked per day, if any). These cut-offs were determined in part to accommodate for digit preference found in the raw data for cigarette smokers (ie, high likelihood of reporting 'round' numbers like 10 or 20 cigarettes per day; Figures A1 and A2 in Appendix A), and in part for consistency with the prior literature. Although the baseline amount smoked was used to categorise smokers, the RRs associated with smoking at these baseline levels were plotted at the number of cigarettes smoked per day at resurvey (around 6 years later) among those who continued to smoke, which is taken as an estimate of the long-term smoking intensity (ie, the 'usual' amount, corrected for regression dilution bias).[35]

The association between age started smoking and premature mortality was assessed by categorizing smokers into the following ages at starting to smoke regularly: 5-9, 10-14, 15-19, and ≥20 years. As with amount smoked, the cut-offs for categorizing age started smoking were primarily based on digit preference and consistency with the (mostly cross-sectional) literature. To assess the relevance of smoking cessation to premature mortality, RRs by age at smoking cessation (25-34, 35-44, and 45-54 years) and years since quitting smoking (<5, 5-14, and ≥15 years) were calculated compared with never or continued smoking (at baseline). As described above, the main analyses of ex-smokers were limited to those without chronic disease at recruitment (but a sensitivity analysis including those with chronic disease is also shown).

To calculate 95% confidence intervals (CIs) for comparisons involving at least three categories of exposure, the group-specific variances and covariances of the log RR in each category were used to produce the CI.[92] These group-specific CIs (reflecting floating absolute risks) allow for statistical comparisons to be made between any pairwise categories, as opposed to comparisons only between an individual category and the arbitrary reference group (as with 'conventional' 95% CIs). Throughout this

thesis, these relative risks are plotted as squares, with the area of each square inversely proportional to the variance of the log risk in that category (to indicate how informative each individual category is). In some cases, such as comparisons between categories of amount smoked and never smoking, a grey shaded area is plotted to indicate a group-specific CI (eg, for the never-smoking group). Where relevant, the RR is given beside the plotted square and 95% CI, and the number of deaths in that category is given below the 95% CI. For clarity, however, any direct comparisons in the text of this thesis are given using conventional CIs (using never smokers as the reference group).

When testing for differential effects between subgroups, chi-squared (χ^2) tests for heterogeneity were calculated between subgroups with the appropriate number of degrees of freedom (ie, one less than the number of groups). To examine whether the length of follow-up time affected the RR estimates (ie, whether the proportional hazards assumption was violated), the RRs relating smoking habits to mortality were estimated separately for the following periods: 0-5, >5-10, >10-15, and >15-20 years after recruitment. However, the log RR estimated from the Cox model over the entire follow-up period may be thought of as a weighted average of the log RRs from each of these 5-year periods, and therefore, the overall RR may be an informative descriptive statistic even if the RRs do vary over the period of follow-up.[93]

In order to assess the importance of the excess risk of premature mortality associated with smoking in Cuba, cause-specific population attributable fractions (PAFs) were estimated. To capture the full impact of tobacco smoking on health (whether current or ex-smoking), the PAFs were calculated using the RRs for ever smoking associated with all medical causes of death. Assuming the relationship between tobacco smoking and premature mortality is largely or wholly causal, the PAF was estimated as P(RR-1) divided by RR, where P is the prevalence of smoking among those dying of the cause

under consideration. Among smokers, where P=1, the fraction of premature mortality attributable to smoking is simply (RR-1)/RR.

Statistical analyses were conducted using SAS v.9.3 (SAS Institute Inc., Cary, NC, USA) and plotted using R v.3.1.1 (R Foundation for Statistical Computing, Vienna, Austria).

3.6 Results

3.6.1 Participant characteristics

Descriptive analyses were initially conducted on all participants aged 30-89 at recruitment in the Cuba Prospective Study, after excluding those with missing information on smoking or covariates of interest (as described in the methods above), resulting in 143 920 participants (Table 3.2). Of these, 120 929 were age 30-69 at recruitment ('middle-aged' participants), and 22 991 were age 70-89 ('older' participants). Overall, the mean age of participants was 55 years, and 44% were male. The majority of participants were recruited in the provinces of Matanzas (44%) and Camagüey (40%), with the rest recruited in three other provinces (16%); this was similar among middle-aged and older participants. Overall, 36% of the cohort did not complete formal education past primary school; however, there were clear generational differences, with 29% of middle-aged participants educated only to primary school, compared with 74% of older participants. Within these age groups, education was similar between men and women. Among middle-aged participants, 4% had a self-reported history of ischaemic heart disease (IHD), 5% had diabetes, and 15% had any chronic disease at recruitment. Among older participants, 10% had a self-reported history of IHD, 11% had diabetes, and 32% had any chronic disease at recruitment.

3.6.2 Trends in smoking behaviour across birth cohorts

The prevalence of ever smoking was similar across male birth cohorts (approximately 60%) born before ~1950, and somewhat lower among those born after the 1950's. The

prevalence of ever smoking among women before 1940 was approximately 25%, and peaked at 40% among women born in the 1950's (Figure 3.1). Smoking cessation was uncommon overall: 15% of ever smokers reported having quit by recruitment. However, smoking cessation was more common among those born earlier (ie, it was positively associated with age), particularly among men. Among both men and women, tobacco was predominantly smoked in the form of cigarettes, with cigar smoking uncommon across birth cohorts. Very few women exclusively smoked cigars regardless of year of birth. Although 10% of men born before 1935 exclusively smoked cigars, only 2% of those born in the 1960's did.

The mean number of cigarettes smoked per day among smokers was similar across birth cohorts, ranging from 15 to 16 cigarettes/day among men and from 12 to 14 cigarettes/day among women. The mean age at smoking initiation also varied little across birth cohorts, ranging from age 18 among women born in the 1930's to age 17 among women born in the 1960's and from age 15 among men born in the 1930's to age 16 among men born in the 1960's (Figure 3.2).

3.6.3 Characteristics of participants at baseline and resurvey by smoking status

Table 3.3 shows the characteristics of the 120 929 participants age 30-69 at recruitment according to baseline smoking behaviour. At recruitment, 52% of men and 29% of women reported current smoking, primarily in the form of cigarettes (with 5% of smokers smoking exclusively cigars, and 7% of predominantly cigarette smokers also smoking cigars). Those who smoked exclusively cigars were on average older (mean age 57 [SD 9]) than primarily cigarette smokers (mean age 49 [SD 9]) or never smokers (mean age 51 [SD 10]). Weekly alcohol consumption was more prevalent among current cigarette smokers (27%) than ex-smokers (17%) or never smokers (9%). However, consuming

alcohol weekly was much more prevalent in male cigarette smokers (41%) than in female cigarette smokers (9%) overall (Tables 3.4 and 3.5).

More than one-third of smokers began smoking regularly before the age of 15: four percent of smokers began at ages 5-9, and 31% began at ages 10-14. Despite having a similar age at recruitment on average, those who began smoking at younger ages completed less formal education than those who began smoking at older ages, or had never smoked. Among those who began smoking at ages 5-9 or 10-14 years, 22% and 35% completed high school, respectively, compared with 44% of never smokers. Participants who began smoking before age 15 were also more likely to be male and to smoke more cigarettes per day than those who began at age 15 or older.

At resurvey, an average of 6 years after recruitment, 84% of cigarette smokers reported continuing to smoke, and those who continued smoking reported smoking a similar number of cigarettes per day at resurvey as at baseline, on average (Table 3.3). However, the mean number of cigarettes per day at resurvey increased among lighter smokers and decreased among heavier smokers (Table 3.6). Approximately one-quarter of ex-smokers at recruitment reported currently smoking at the time of the resurvey. As this observation is based on a very small number of ex-smokers at recruitment who were resurveyed (n=92), it may be only qualitatively informative about smoking resumption among ex-smokers. Participant characteristics were broadly similar among those aged 70-89 years at recruitment (Table A1 in Appendix A).

3.6.4 Association between current smoking and mortality

3.6.4.1 All-cause mortality

Among the 120 929 adults contributing to the main analyses, there were 1.7 million person-years of follow-up at ages 30-69 (mean 14 SD 6 per person), during which there were 8691 deaths. Of these, 3114 were vascular (1623 ischaemic heart disease [IHD],

1491 stroke or other vascular), 2993 neoplastic (1076 lung or upper-aerodigestive cancer ['respiratory cancers'], 1917 other), 666 respiratory (321 chronic obstructive pulmonary disease [COPD], 345 other), 1413 from other medical causes, and 505 from external causes.

Compared with never smokers, the all-cause mortality rate ratio (RR) at ages 30-69 among current cigarette smokers ('current smokers') was 1.66 (95% CI 1.59-1.74). However, the excess risk of premature death associated with smoking varied substantially depending on the age at smoking initiation and on the amount currently smoked (Figure 3.3). The all-cause mortality RR for those who started at ages 5-9, 10-14, 15-19, and ≥20 years was, respectively, 2.53 (2.23-2.87), 1.83 (1.72-1.95), 1.55 (1.46-1.65), and 1.49 (1.38-1.61), compared with never smokers. Similarly, the all-cause mortality RR for those who had reported smoking <10, 10-19, 20, or >20 cigarettes per day at recruitment was, respectively, 1.44 (1.33-1.55), 1.59 (1.50-1.70), 1.76 (1.66-1.88), 2.17 (1.98-2.38), compared with never smokers. These associations were little altered when excluding those reporting a history of chronic disease at recruitment (Figure 3.4). The RRs for each category of age started smoking and amount smoked were similar between men and women, with some evidence that women starting the youngest and smoking the most had slightly higher relative risk than men in comparable smoking categories (Figure 3.5 and Figure 3.6).

The effect of smoking within levels of each confounder adjusted for in the main analyses is shown in Figure 3.7. Overall, the associations of current smoking versus never smoking with all-cause mortality were similar at ages 30-54 and 55-69, in men and women, among residents of different provinces, and across levels of educational attainment and alcohol consumption, with no significant heterogeneity observed across these groups (considering a p-value of 0.05 as the significance threshold). Furthermore,

there was little evidence that these all-cause mortality RRs comparing current and never smokers varied much over the follow-up period; the RRs at 0-5, >5-10, >10-15, and >15-20 years after recruitment were 1.56 (95% CI 1.43-1.70), 1.52 (1.39-1.66), 1.81 (1.65-1.98), and 1.74 (1.56-1.95), respectively.

At ages 70-89, the relevance of age started smoking and the amount smoked were less clear. The higher excess risk of all-cause mortality associated with smoking initiation at ages 5-9 or 10-14 compared with ages 15-19 or ≥20 years or never smoking was again observed, but the all-cause RR for each category was considerably lower than at ages 30-69 (Figure A3 in Appendix A). There was little association between amount smoked and all-cause mortality at ages 70-89 compared with never smokers. As observed at younger ages, these associations were little changed when excluding participants reporting a history of chronic disease at recruitment (Figure A4 in Appendix A).

3.6.4.2 Cause-specific mortality

Dose-response relationships were observed for both age at smoking initiation and number of cigarettes smoked per day and several causes of death at ages 30-69, including cancer (both respiratory and other sites), respiratory (both COPD and other respiratory causes), and IHD (Figures 3.8 and 3.9). For each medical cause of death, smoking initiation before the age of 15 was associated with higher risk of premature mortality in adulthood (Figure 3.8). Compared with never smokers, the RRs for starting to smoke at ages 5-9 and 10-14 were, respectively, 6.71 (95% CI 4.91-9.18) and 4.85 (4.03-5.83) for respiratory cancers, 1.70 (1.23-2.36) and 1.51 (1.32-1.74) for other cancers, 7.87 (4.74-13.05) and 4.37 (3.17-6.02) for COPD, 2.79 (1.57-4.97) and 1.78 (1.30-2.44) for other respiratory causes, 2.73 (2.06-3.61) and 1.94 (1.68-2.24) for IHD, 2.55 (1.90-3.41) and 1.67 (1.43-1.95) for stroke and other vascular causes, and 1.43 (0.98-2.08) and 1.29 (1.10-1.52) for all other medical causes.

Similarly, heavier smokers were at particularly high excess risk of premature mortality from medical causes (Figure 3.9). Compared with never smokers, the RRs for smoking 20 and >20 cigarettes per day (as reported at recruitment) were, respectively, 5.02 (4.18-6.02) and 6.70 (5.32-8.44) for respiratory cancers, 1.35 (1.17-1.56) and 1.76 (1.42-2.18) for other cancers, 4.83 (3.53-6.60) and 6.21 (4.05-9.52) for COPD, 1.51 (1.09-2.09) and 2.54 (1.66-3.88) for other respiratory causes, 1.78 (1.54-2.06) and 1.97 (1.58-2.47) for IHD, 1.59 (1.36-1.86) and 1.72 (1.35-2.19) for stroke and other vascular causes, and 1.28 (1.09-1.51) and 1.41 (1.09-1.82) for other medical causes of death. However, the association between amount smoked and premature mortality from vascular causes was somewhat weaker than that of age started smoking and vascular mortality.

At ages 70-89, dose-response relationships were observed for age started smoking and amount smoked in relation to respiratory cancers and COPD, with limited evidence of dose-response relationships for vascular causes and age started smoking, and little clear association with other causes of death in either case (Figures A5 and A6 in Appendix A).

As few smokers exclusively smoked cigars (as opposed to smoking cigars in addition to cigarettes), there was insufficient power to determine reliably the association between cigar smoking and cause-specific mortality at ages 30-69. However, compared with never smokers, the all-cause mortality RR associated with smoking cigars only (1.30 [1.13-1.50]) was lower than that of smoking predominantly or exclusively cigarettes (1.66 [1.59-1.74]), but nevertheless was associated with excess risk of premature death (Figure 3.10).

3.6.5 Association between smoking cessation and all-cause mortality

Figure 3.11 shows the association between smoking cessation at different ages and for duration of time since quitting and all-cause mortality. Among ex-smokers without chronic disease at recruitment who quit at ages 25-54, quitting by approximately age 40 avoided most excess risk of continued smoking. Compared with never smokers, the all-cause mortality RR associated with quitting at ages 25-34, 35-44, and 45-54 years was 1.09 (0.82-1.44), 1.21 (0.98-1.50), and 1.39 (1.12-1.73), respectively. Similarly, the RR for all-cause mortality at ages 30-69 attenuated with quitting duration: quitting <5 years prior to baseline, 5-14, and ≥15 years prior to recruitment was associated with RRs of 1.43 (95% CI 1.15-1.84), 1.31 (1.07-1.62), and 1.02 (0.79-1.31), respectively, compared with never smokers. A sensitivity analysis including participants with chronic disease at baseline resulted in higher RRs for all-cause mortality, especially among those quitting at older ages and those quitting more recently, which may be evidence of the 'sick-quitter' effect (Figure A7 in Appendix A). As quitting smoking was uncommon in this population, there were insufficient deaths among ex-smokers to assess reliably the association between smoking cessation and specific causes of death.

3.6.6 Population attributable fraction of mortality associated with ever smoking

If the association between smoking and mortality at ages 30-69 was chiefly causal, then approximately one in four male deaths (27%) and one in five female deaths (19%) in this population were caused by smoking (Table 3.7). Among both men and women, smoking caused more than half of all respiratory cancer deaths and approximately one-quarter of IHD deaths. Among ever-smoking men, smoking caused 33% of IHD deaths, 70% of respiratory cancer deaths, 71% of COPD deaths, and 38% of premature deaths overall. Among ever-smoking women, smoking caused 50% of IHD deaths, 79% of respiratory cancer deaths, 73% of COPD deaths, and 42% of premature deaths overall. While ever-regular smoking was more prevalent among men than among women, the RRs

associated with premature death caused by smoking were somewhat higher among women than among men for several causes, particularly IHD and respiratory (ie, lung or upper-aerodigestive) cancers. As a result, the PAFs were generally similar by sex.

3.7 Discussion

3.7.1 Summary of findings

Of the 143 920 Cuban adults recruited into the study between 1996-2002, approximately half of men and one-quarter of women smoked at recruitment. In this population born before 1970, one-third of smokers began smoking regularly before the age of 15, and many began even before the age of 10. Adjusting for age-at-risk, sex, education, province, and alcohol consumption, smokers had a 66% higher risk of premature death than never smokers. The risks associated with premature mortality in middle age depended strongly on the age at smoking initiation, however: while those who had started smoking at age 20 years or older had a 50% higher risk of premature death compared with never smokers, those who had begun before the age of 10 had 150% higher risk of premature death.

Given the clear dose-response relationships between both age started smoking and amount smoked and mortality from causes of death strongly associated with smoking in other populations[3,4,35] (eg, respiratory diseases/cancers), it is reasonable to assume that the excess mortality observed among smokers was largely or wholly caused by smoking. If so, then smoking caused more than one-quarter of all premature deaths in men and approximately one-fifth of all premature deaths in women. More than half of the deaths among those who began smoking before the age of 10 or smoked more than 20 cigarettes per day were caused by smoking. However, even among those who had started to smoke at older ages (age 24 on average) or smoked relatively little (9 cigarettes/day on average), more than one-quarter were killed by their habit.

Although Cuba is known for its cigar production, tobacco was predominantly smoked in the form of cigarettes, which were associated with significantly greater risk of premature mortality than cigars. However, among the few (mostly older men) who smoked exclusively cigars at recruitment, cigar smoking was associated with a significant increase in premature mortality compared with never smoking.

Smoking cessation was uncommon prior to recruitment, among both men and women and older and younger participants. However, among the relatively few who had quit by the time of recruitment, the risk of premature death was substantially lower than among those who were current smokers (especially among those who had quit by the age of about 40). Furthermore, some of those who had quit may have done so only after developing disease from smoking (as opposed to doing so while still healthy), and some who had reported quitting prior to recruitment had resumed smoking by the time of the resurvey, resulting in likely under-estimation of the true benefits of quitting. The relevance of quitting smoking to specific causes of death could not yet be assessed reliably.

3.7.2 Strengths and limitations

A major strength of this study is its size, which was essential for reporting the effects of relatively uncommon, but important, smoking behaviours. Alongside the Mexico City Prospective Study (described in **Chapter 4**), this is the first large prospective study of the association between smoking and cause-specific mortality in a Latin American country, and the first study globally to demonstrate reliably the hazards of smoking from before the age of 10. This finding raises the hypothesis that, if smoking from before age 10 has occurred in other populations globally, it is likely to be associated with substantially higher excess risk than starting to smoke at older ages (or not at all). However, this has not yet been examined in sufficiently large studies prospectively, or

even commonly reported in cross-sectional studies, and is examined in **Chapter 5** and **Chapter 6**, respectively. Similarly, though smoking cessation was uncommon in this population, this study was of a sufficiently large size to demonstrate the benefits of quitting smoking in Cuba, which is critical for public health messages.

This study benefitted from several other strengths. The long follow-up time and low loss to follow-up allowed for sufficient accrual of events to examine cause-specific mortality reliably, even for relatively uncommon smoking behaviours (eg, smoking from before age 10), and reduced the likelihood of differential follow-up bias. The reliable linkage to the national registry of medically certified deaths further reduced the likelihood of bias from attrition during follow-up and enhanced the precision of causes of death. Though short, the questionnaire collected sufficiently detailed information on smoking behaviours to offer clear evidence on the risks of starting at different ages, smoking different amounts, smoking different products, and quitting at different ages. Finally, the resurvey conducted 6 years after baseline allowed for 'usual' smoking habits to be examined, resulting in more accurately characterised long-term smoking behaviours and associated risks. A particular strength of this resurvey is that it was conducted at approximately the midpoint of follow-up, which should have been conducted long enough after recruitment to allow for behaviour change, but not conducted so long after recruitment that it would lead to biased information (particularly as resurveys can only be conducted among surviving participants, whose behaviours may materially differ from those who died).

This study also had several limitations. Both the baseline survey and resurvey were deliberately brief, and therefore collected information on a limited set of potential confounders and effect modifiers, which meant that confounding or effect modification by potentially important factors could not be assessed. Smoking information was self-reported and collected at a single point in time, which could have resulted in exposure

misclassification. However, this was partially addressed by the resurvey conducted in a subset of participants (although information was again self-reported). Ideally, the protective effects of quitting should be examined separately in those who quit while still healthy versus those who quit after developing (particularly smoking-related) disease. However, information was not collected about reasons for quitting smoking, so the benefits of smoking cessation are likely to have been under-reported (as 'sick quitters' may have misleadingly higher risk than even continuing smokers, but may still benefit from quitting). However, to limit the effect of this reverse causality, participants with chronic disease were excluded from the main analyses of smoking cessation. The higher relative risks observed when including those with chronic disease at recruitment, especially among those who quit recently or at older ages, may be evidence of this 'sick-quitter' effect.

Although the study was not designed to be nationally representative, which could limit the generalisability of findings, participant characteristics (both demographic and behavioural) are consistent with those of a national survey conducted in 2001[94] (shortly after the midpoint of recruitment into this study). However, the baseline questionnaire was deliberately short, and therefore collected only limited demographic and behavioural information. Therefore, it is possible that there were other differences between study participants and the general population that were not captured in the survey. The observation that current versus never-smoker RRs were similar across levels of major likely confounders may lend further support to the generalisability of these findings.

Finally, one notable limitation of this study is that the recruitment period coincided with the Special Period in Cuba, in which there was severe economic hardship nationwide, which is likely to have severely impacted on both smoking habits and other lifestyle factors.[95] This may partly explain the lack of association between the reported number of

cigarettes smoked per day at recruitment and all-cause mortality at ages 70-89 (and for causes such as IHD), despite a similar pattern of greater excess risk among smokers starting at younger versus older ages and mortality at ages 30-69 and 70-89 (though the RRs were much lower in all categories at ages 70-89 than at ages 30-69 compared with never smokers). However, given that a dose-response relationship was observed between age at smoking initiation and both COPD and respiratory cancers at ages 70-89, but not IHD or stroke, the lower overall RRs may be driven in part by different effects of smoking duration on different systems in the body (eg, smoking from a young age may strongly affect the respiratory system even at older ages, but may have only a modest effect on the circulatory system at these ages). Regardless, given the robust association between both cigarettes smoked per day and age started smoking and mortality at ages 30-69, as observed in other populations globally, the information collected should be reasonably reflective of long-term behaviours and their associated risks.

3.7.3 Findings in the context of prior evidence

Placing the results from this Cuban study into national context is difficult, as little representative published evidence exists on adult health in Cuba. For instance, National Risk Factor Surveys were conducted in 1995-1996, 2001, and 2010-2011, but only results from the latter have been published (and can only be accessed as a hard copy report published in Spanish). However, some estimates from these surveys have been published in peer-reviewed literature, and I obtained a copy of the 2010-2011 National Risk Factor Survey from its Principal Investigator (Dr. Patricia Varona Peréz) while attending a conference in Havana, Cuba, in June 2018.

According to the 2010-2011 National Risk Factor Survey, the national prevalence of ever-regular smoking in Cuba in 2010 was 36%: 45% in men and 26% in women.

Current smoking prevalence was 24% (31% in men and 16% in women), and one-third of ever-smokers had quit at the time of the survey. Among current smokers, 4% began smoking regularly before the age of 10, and an additional 25% began at ages 10-14 (which was similar by sex), and 94% of smokers exclusively smoked cigarettes, while 6% smoked exclusively cigars, and 0.5% smoked both.[96] Current cigarette smokers smoked an average of 15 cigarettes/day (15/day among men and 14/day among women).

Results from this national survey ~12 years after the Cuba Prospective Study (or 'Cuban cohort') baseline survey indicate that the prevalence of smoking, the prevalence of smoking initiation in childhood, and the prevalence of smoking tobacco predominantly as cigarettes (and the average number of cigarettes smoked/day), were very similar to the prevalence and patterns of smoking in the Cuban cohort. One difference was the comparatively higher smoking prevalence in the Cuban cohort at recruitment (approximately half of men and one-quarter of women smoked), but published estimates of smoking prevalence in Cuba suggest that smoking prevalence was approximately 40% in 1995 and 35% in 2001 (among men and women combined),[95] which would be nearly identical to the smoking prevalence in the Cuban cohort baseline survey conducted around 2000. Further, the comparatively high smoking cessation prevalence reported in 2010 is inconsistent with the lack of reduction in cigarettes purchased per capita nationally over that time period,[97] suggesting smoking cessation may have been over-reported in the national survey compared with the Cuban cohort (Figure 3.12).

Therefore, it may be reasonable to assume that findings from the Cuban cohort should be broadly generalisable to the Cuban population. However, no other large prospective study has yet been set up in Cuba, and so validating these estimates directly in another Cuban population is not yet possible. However, I have discussed with local collaborators

and researchers in Oxford the possibility of linking the national surveys conducted in 1995, 2001, and 2010 and to the national death registry using identifying information collected at the time of the surveys (including a national identification number, as used to link participants to the death registry in the Cuban cohort), which could plausibly be used to validate these findings in the future.

Although prior evidence has indicated that smoking initiation at younger ages is more harmful than smoking initiation at older ages, few studies have assessed the relevance of smoking initiation before age 15 to premature mortality. A large US study of 90 278 adults recruited from 1997-2005 reported that 7% of smokers began smoking before age 13, and 11% began at ages 13-14. Although the study had limited power, compared with smokers who began at age ≥21, those who began at age 13-14 had an all-cause mortality RR of 1.24 (95% CI 1.12-1.36), and those who began at age <13 had an all-cause mortality RR of 1.18 (1.04-1.33), and much stronger associations were found in cross-sectional associations with chronic disease.[98] Another US study of 14 200 adults recruited in 1987-1989 reported a much stronger association between age at smoking initiation and mortality: compared with those who began smoking at age ≥22, the CVD mortality RR among those who started smoking at age ≤12 was 2.70 (1.89-3.86) after adjusting for age, sex, smoking-years, and cigarettes/day at baseline.[99] This CVD mortality associated with smoking at age ≤12 compared with age ≥22 was somewhat higher in White smokers (RR 3.19; 2.03-5.01) than in African American smokers (RR 2.17; 1.20-3.94).

While the Cuban Prospective Study is among the few studies to reliably assess the hazards of smoking initiation at ages 10-14, it is the first study globally to assess reliably the relevance of smoking initiation before age 10 to premature mortality. This finding in a Cuban population born before 1970 raised the hypothesis that smoking from childhood

(ie, age <15), and even from early childhood (ie, age <10), may be associated with similarly high excess risk in other populations globally. This hypothesis was explored in a large prospective US study in **Chapter 5**, and in the cross-sectional Global Adult Tobacco Surveys from 30 low- and middle-income countries in **Chapter 6**.

Other main findings from this study are similar to those reported previously in the literature. For example, although relatively few smokers in this study had quit by recruitment and therefore there were relatively few deaths among ex-smokers, the observation that stopping smoking by the age of approximately 40 avoided most of the excess risk associated with continued smoking is similar to findings from other populations globally (which had much larger numbers of ex-smokers).[4,5,35] Similarly, the observation that smoking exclusively cigars was associated with significantly increased risk of mortality compared with never smoking, but lower risk than predominantly cigarette smoking, was also in keeping with prior research in other populations globally.[100,101] These findings are discussed in further detail in the context of findings from **Chapter 4** and **Chapter 5**, as well as other populations globally, in **Chapter 6**.

Overall, however, the RRs associated with premature mortality are considerably lower than those of contemporary populations in which the smoking epidemic is mature. For example, the RRs associated with current cigarette smoking in this study were 3.89, 1.74, and 1.66 for COPD, IHD, and all-cause mortality, respectively, compared with RRs of 23.03, 2.86, and 2.80 for COPD, IHD, and all-cause mortality, respectively, in a contemporary US population.[8] The reason for these comparatively low RRs in a population in which smoking from a young age and heavy smoking were common remains unclear. It is possible that major fluctuations in smoking (and other related behaviours) occurred during The Special Period, that participant smoking patterns varied considerably in the decades before the study (eg, intermittent smoking, variable

amounts smoked), and potentially differences in product smoked (eg, it is possible that a number of participants transitioned from smoking cigars to cigarettes in the period prior to recruitment, under-estimating the harms of prolonged cigarette smoking).

Despite divergences in the RRs associated with smoking from those observed in high-income countries, indirect estimates of the impact of smoking on mortality nationally in Cuba are generally very similar to those observed in this study. For example, an estimated 27% of male deaths and 19% of female deaths from medical causes in this study were due to smoking. Estimates of tobacco-attributable mortality in Cuba from the Global Burden of Disease (GBD) Study in 2010 (ie, around the mid-point of the Cuban study) suggest that 27% of non-communicable disease deaths among men and 15% among women were caused by smoking.[61] Direct and indirect estimates were also similar by cause of death. For example, direct estimates from the Cuban cohort indicate that 59% of male and 46% of female COPD deaths were attributable to smoking, compared with GBD estimates of 64% of male and 40% of female COPD deaths attributed to smoking. Therefore, these indirect estimates of tobacco-attributable mortality in Cuba appear to be similar to direct estimates from this thesis. These direct estimates provide further evidence that tobacco smoking causes one in four premature deaths among men and one in five premature deaths among women in Cuba, underscoring the urgent need for tobacco control policies to combat this epidemic.

3.8 Conclusion

In this large prospective study in a Cuban population, smoking initiation at ages 10-14 was associated with nearly twice the risk of premature death compared with never smoking, and smoking before the age of 10 was associated with even greater risk of premature death. The prevalence of smoking initiation in childhood in this population, and prospective evidence of its harms, underscores the need for further study of the

prevalence of childhood smoking in diverse populations globally, and its relevance to premature death in adulthood. Although smoking caused an estimated two-fifths of the premature deaths among smokers in this population, as in other populations, quitting worked: those who quit smoking before the age of ~40 avoided most of the excess risk of premature death associated with continued smoking.

Table 3.1: ICD-9 and ICD-10 codes for cause-specific mortality endpoints

Endpoint	ICD-9 codes	ICD-10 codes
Cancer		
Lung or upper-aerodigestive	140-150, 161-162	C00-C15, C32-C34
All other	151-160, 163-208	C16-C31, C35-C97
Respiratory		
COPD	490-492, 494-496	J40-J44, J47
All other	460-489, 497-519, 010-012	A15-16, J00-J39, J48-J98
Vascular		
IHD	410-414	I20-I25
Stroke and all other	390-409, 415-459, 798	I00-I19, I26-I99
External	800-1002	S00-Y89
All-cause	000-1002	A00-Y89

Table 3.2: Baseline characteristics of the 143 920 participants in the Cuba Prospective Study, by age and sex

	Age 30-69			Age 70-89			
	Men	**Women**	**All**	**Men**	**Women**	**All**	**Total**
Participants	52 999	67 930	120 929	10 313	12 678	22 991	143 920
Age, years	50 (10)	50 (10)	50 (10)	77 (5)	77 (5)	77 (5)	55 (14)
Region							
Matanzas	44%	43%	44%	44%	43%	44%	44%
Camaguey	40%	40%	40%	41%	41%	41%	40%
Pinar del Rio	8%	8%	8%	6%	5%	6%	8%
Ciudad de La Habana	5%	6%	6%	7%	8%	8%	6%
La Habana	2%	2%	2%	2%	2%	2%	2%
Education completed							
University	11%	10%	10%	2%	2%	2%	9%
High school/technical college	37%	29%	32%	11%	6%	9%	29%
Lower secondary	28%	29%	29%	16%	14%	15%	27%
Primary	20%	26%	24%	53%	57%	55%	29%
Less than primary	4%	6%	5%	17%	21%	19%	7%
History of chronic disease†							
Ischaemic heart disease	4%	4%	4%	10%	11%	10%	5%
Stroke	1%	1%	1%	6%	6%	6%	2%
Diabetes	3%	6%	5%	7%	14%	11%	6%
Cancer	<0.5%	1%	1%	2%	2%	2%	1%
Other	7%	6%	6%	12%	8%	10%	7%
Any	14%	16%	15%	30%	33%	32%	18%

Percent or mean (SD) shown.
† Self-reported history of ischaemic heart disease (including myocardial infarction or angina pectoris), stroke, diabetes, cancer (any site), chronic obstructive pulmonary disease, chronic kidney disease, cirrhosis of the liver, and peptic ulcer.

Table 3.3: Baseline characteristics of the 120 929 participants aged 30-69 at recruitment, by smoking status

	Never smoker	Ex smoker	Smoker of cigars only	Cigarette smoker, by age started				
				Any	5-9	10-14	15-19	20+
Participants	66 536	7197	2559	44 637	1747	13 613	19 409	9868
Age, years	51 (10)	53 (10)	57 (9)	49 (9)	52 (10)	50 (10)	48 (9)	49 (9)
Male (%)	33%	53%	86%	57%	65%	61%	59%	45%
Completed high school	44%	42%	25%	42%	22%	35%	46%	49%
Consume alcohol weekly	9%	17%	27%	27%	27%	29%	29%	22%
BMI kg/m²	25 (4)	25 (4)	25 (4)	24 (4)	24 (4)	24 (4)	24 (4)	24 (4)
SBP (mmHg)	124 (15)	126 (16)	128 (16)	123 (15)	124 (17)	123 (15)	123 (14)	124 (15)
Smoking at entry								
Age started smoking (years)	-	17 (6)	17 (7)	17 (5)	8 (1)	12 (1)	17 (1)	24 (6)
Cigarettes/day*	-	-	-	15 (9)	18 (11)	17 (9)	15 (8)	13 (8)
Cigar smokers*	-	-	100%	7%	12%	9%	7%	5%
Smoking at resurvey†								
Current smokers at resurvey	4%	26%	80%	84%	86%	83%	86%	80%
Cigarettes/day among those still smoking	-	-	-	16 (10)	19 (10)	18 (10)	16 (9)	13 (9)
Cigarettes/day among all resurveyed	0 (3)	4 (9)	3 (7)	13 (11)	16 (12)	14 (11)	13 (10)	10 (10)

Percent or mean (SD) shown.
Excludes those age 70 or older at recruitment and those with missing information on smoking or covariates.
*Smokers of cigars only consumed 3.2 (SD 3.1) cigars/day, and the few cigarette smokers who also smoked cigars consumed 3.8 (SD 4.5) cigars/day.
† 3578 participants were resurveyed about 6 years after baseline. At baseline, 2233 participants reported never smoking, 91 ex-smoking, 75 cigar smoking only, and 1179 cigarette smoking (mean 16/day).

Table 3.4: Baseline characteristics of the included male participants aged 30-69 at recruitment, by smoking status

	Never smoker	Ex smoker	Smoke cigars only	Cigarette smoker, by age started				
				Any	5-9	10-14	15-19	20+
Participants	21 755	3785	2199	25 260	1135	8271	11 398	4456
Age, years	50 (10)	53 (10)	57 (9)	50 (10)	52 (10)	51 (10)	49 (9)	49 (9)
Completed high school	53%	47%	26%	45%	25%	38%	49%	55%
Consume alcohol weekly	20%	27%	28%	41%	38%	41%	42%	40%
BMI kg/m²	24 (4)	24 (4)	25 (4)	24 (4)	24 (4)	24 (4)	24 (4)	24 (4)
SBP (mmHg)	125 (13)	127 (15)	128 (16)	125 (14)	124 (16)	125 (14)	125 (13)	125 (14)
Smoking at entry								
Age started smoking (years)	-	16 (5)	17 (7)	16 (4)	8 (1)	12 (1)	17 (1)	22 (5)
Cigarettes/day	-	-	-	16 (9)	20 (11)	18 (9)	16 (8)	14 (8)
Cigar smokers	0%	0%	100%	11%	16%	13%	9%	9%
Smoking at resurvey†								
Current smokers at resurvey	7%	26%	83%	85%	86%	82%	87%	85%
Cigarettes/day among those still smoking	-	-	-	18 (10)	20 (10)	20 (10)	17 (10)	15 (9)
Cigarettes/day among all resurveyed	1 (4)	5 (10)	3 (6)	15 (11)	16 (12)	16 (12)	15 (11)	12 (10)

Percent or mean (SD) shown.
† Of the 3578 participants resurveyed, 1519 were men. Among men at baseline, 732 participants reported never smoking, 57 ex-smoking, 69 cigar smoking only, and 661 cigarette smoking (mean 18/day).

Table 3.5: Baseline characteristics of the included female participants aged 30-69 at recruitment, by smoking status

	Never smoker	Ex smoker	Smoke cigars only	Cigarette smoker, by age started				
				Any	5-9	10-14	15-19	20+
Participants	44 781	3412	360	19 377	612	5342	8011	5412
Age, years	51 (10)	52 (10)	57 (9)	48 (9)	50 (10)	48 (9)	47 (9)	49 (9)
Completed high school	39%	36%	18%	38%	15%	30%	42%	44%
Consume alcohol weekly	3%	6%	15%	9%	8%	11%	10%	7%
BMI kg/m²	25 (4)	25 (5)	25 (5)	24 (4)	24 (5)	24 (5)	24 (4)	24 (4)
SBP (mmHg)	124 (16)	124 (17)	128 (17)	122 (16)	122 (19)	121 (16)	122 (16)	122 (16)
Smoking at entry								
Age started smoking (years)	-	19 (7)	20 (10)	17 (6)	8 (1)	13 (1)	17 (1)	25 (7)
Cigarettes/day	-	-	-	13 (8)	16 (10)	15 (9)	13 (8)	12 (8)
Cigar smokers	0%	0%	100%	3%	4%	3%	3%	3%
Smoking at resurvey†								
Current smokers at resurvey	3%	26%	50%	81%	86%	83%	83%	76%
Cigarettes/day among those still smoking	-	-	-	14 (8)	18 (10)	14 (8)	14 (8)	12 (9)
Cigarettes/day among all resurveyed	0 (2)	3 (7)	5 (12)	11 (9)	15 (11)	12 (9)	11 (9)	9 (10)

Percent or mean (SD) shown.
† Of the 3578 participants resurveyed, 2059 were women. Among women at baseline, 1501 participants reported never smoking, 34 ex-smoking, 6 cigar smoking only, and 518 cigarette smoking (mean 13/day).

Table 3.6: Smoking characteristics at resurvey (mean 6 years later) according to baseline smoking status, ages 30-69 at recruitment

	Never smoker	Ex smoker	Smoke cigars only	Cigarette smoker, by cigarettes/day at recruitment				
				Any	< 10	10-19	20	21+
Participants	2243	92	76	1186	262	350	430	144
Smoking status at resurvey								
Never smoker	2070 (92%)	0 (0%)	0 (0%)	1 (<0.5%)	0 (0%)	0 (0%)	0 (0%)	1 (1%)
Ex smoker	73 (3%)	68 (74%)	15 (20%)	194 (16%)	63 (24%)	67 (19%)	50 (12%)	14 (10%)
Smoke < 20 cigs/day	62 (3%)	13 (14%)	10 (13%)	525 (44%)	168 (64%)	199 (57%)	126 (29%)	32 (22%)
Smoke 20 cigs/day	23 (1%)	5 (5%)	4 (5%)	317 (27%)	19 (7%)	57 (16%)	198 (46%)	43 (30%)
Smoke >20 cigs/day	6 (<0.5%)	4 (4%)	2 (3%)	121 (10%)	5 (2%)	18 (5%)	49 (11%)	49 (34%)
Smoke cigars only	9 (<0.5%)	2 (2%)	45 (59%)	28 (2%)	7 (3%)	9 (3%)	7 (2%)	5 (3%)
Average daily cigs								
At baseline	-	-	-	15.8 (9.5)	4.8 (2.0)	11.5 (2.2)	20.0 (0.0)	34.0 (8.6)
At resurvey, per smoker †	12.3 (9.6)	16.8 (11.9)	13.0 (8.9)	16.0 (9.6)	9.3 (7.5)	14.1 (7.6)	18.3 (7.5)	24.1 (13.0)
At resurvey, all participants ‡	0.5 (3.1)	4.0 (9.2)	2.7 (6.6)	13.0 (10.7)	6.8 (7.6)	11.1 (8.9)	15.8 (9.4)	20.7 (14.7)
Average daily cigars								
At baseline	-	-	2.4 (1.4)	0.1 (0.4)	0.0 (0.2)	0.1 (0.4)	0.1 (0.4)	0.1 (0.3)
At resurvey, per smoker †	0.3 (1.1)	0.3 (0.7)	1.9 (1.7)	0.1 (0.6)	0.2 (0.7)	0.1 (0.7)	0.1 (0.6)	0.1 (0.6)
At resurvey, all participants ‡	0.0 (0.2)	0.1 (0.4)	1.5 (1.7)	0.1 (0.6)	0.1 (0.6)	0.1 (0.6)	0.1 (0.6)	0.1 (0.5)

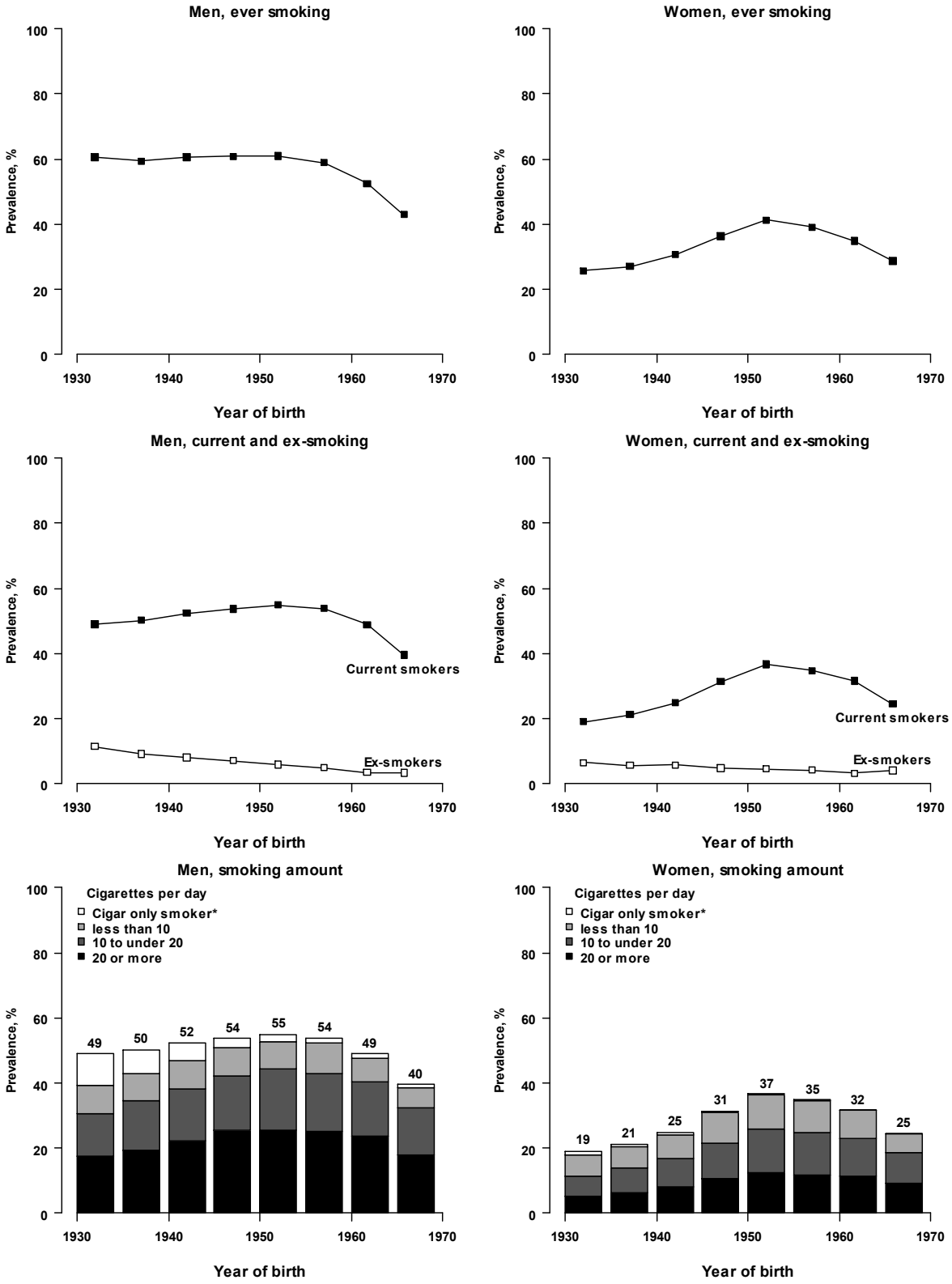
Includes only participants with non-missing smoking information, and non-missing blood pressure and BMI values.
Percent or mean (SD) shown.
† Among those self-reporting current smoking at resurvey
‡ Among all participants, i.e. those reporting never or ex smoking assigned zero cigarettes/day

Table 3.7: Cause-specific mortality rate ratios (RRs) and population attributable fractions (PAFs) for ever versus never smoking at ages 30-69, by sex

	Male			Female		
Cause of death	No. of deaths smokers/non-smokers	RR (95% CI)	PAF* (%)	No. of deaths smokers/non-smokers	RR (95% CI)	PAF (%)
IHD	665/307	1.51 (1.31-1.73)	23	324/327	2.02 (1.72-2.36)	25
Stroke and other vascular	529/224	1.55 (1.32-1.83)	25	329/409	1.58 (1.36-1.83)	16
Respiratory cancers	578/105	3.30 (2.67-4.08)	59	275/118	4.68 (3.77-5.83)	55
Other cancers	479/237	1.33 (1.13-1.56)	17	488/713	1.33 (1.18-1.49)	10
COPD	117/23	3.45 (2.19-5.43)	59	115/66	3.64 (2.68-4.95)	46
Other respiratory	110/62	1.22 (0.89-1.68)	12	80/93	1.85 (1.37-2.50)	21
Other medical	444/231	1.24 (1.06-1.46)	13	281/457	1.19 (1.02-1.38)	6
All medical	2922/1189	1.61 (1.51-1.73)	27	1892/2183	1.72 (1.61-1.83)	19

* Example calculation: for respiratory cancer in men, (2.30 / 3.30) x 578 ÷ (578 + 105) = 0.59

Figure 3.1: Smoking prevalence in the Cuba Prospective Study



Number above bar in plots 5 and 6 is percent current smokers
*: Cigar but not cigarette smokers

Figure 3.2: Mean cigarettes/day, and age started smoking, among 44 042 current daily smokers

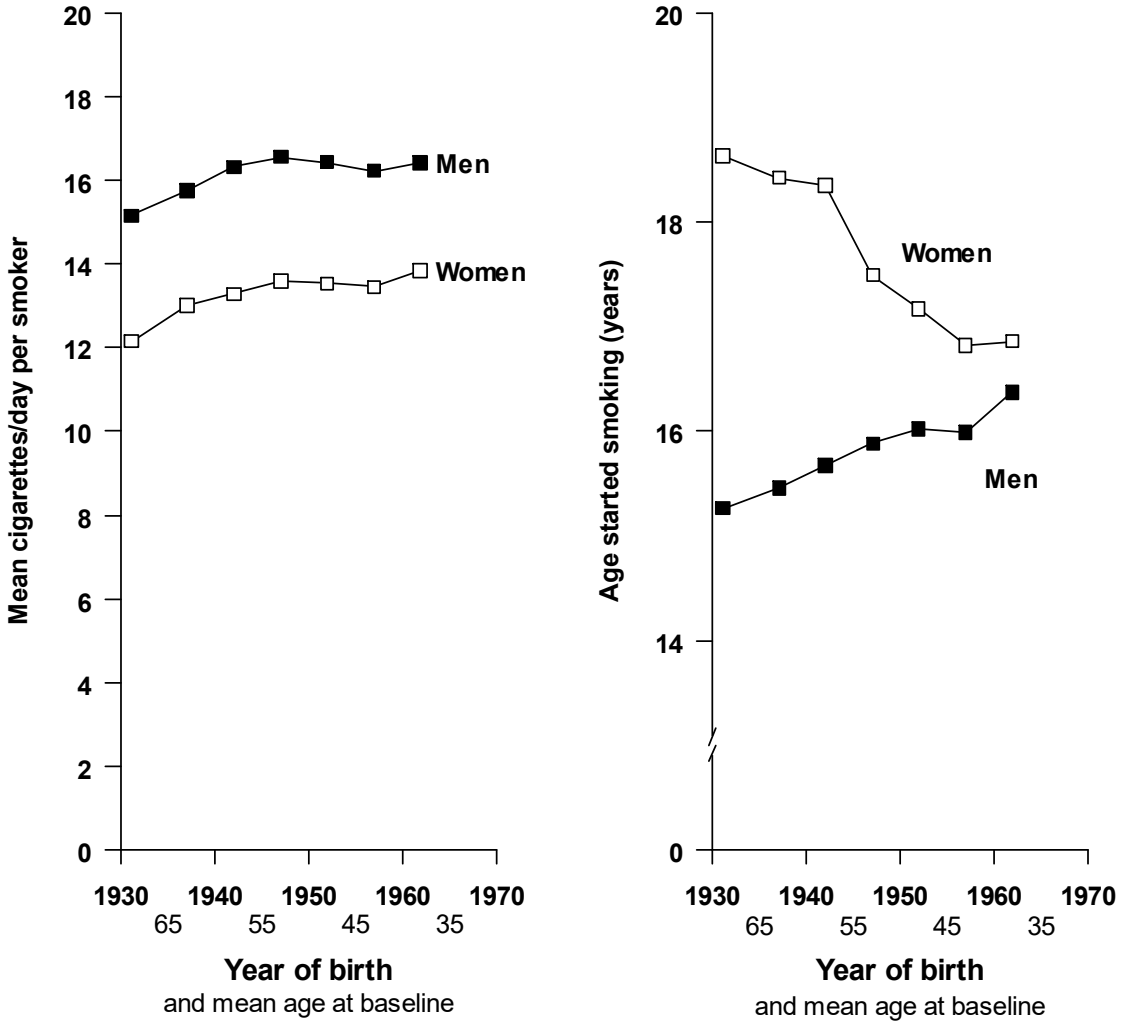


Figure 3.3: All-cause mortality rate ratio at ages 30-69, cigarette smoker vs never smoker, by age started smoking and usual amount smoked

Participants with chronic disease at recruitment were included in these analyses.

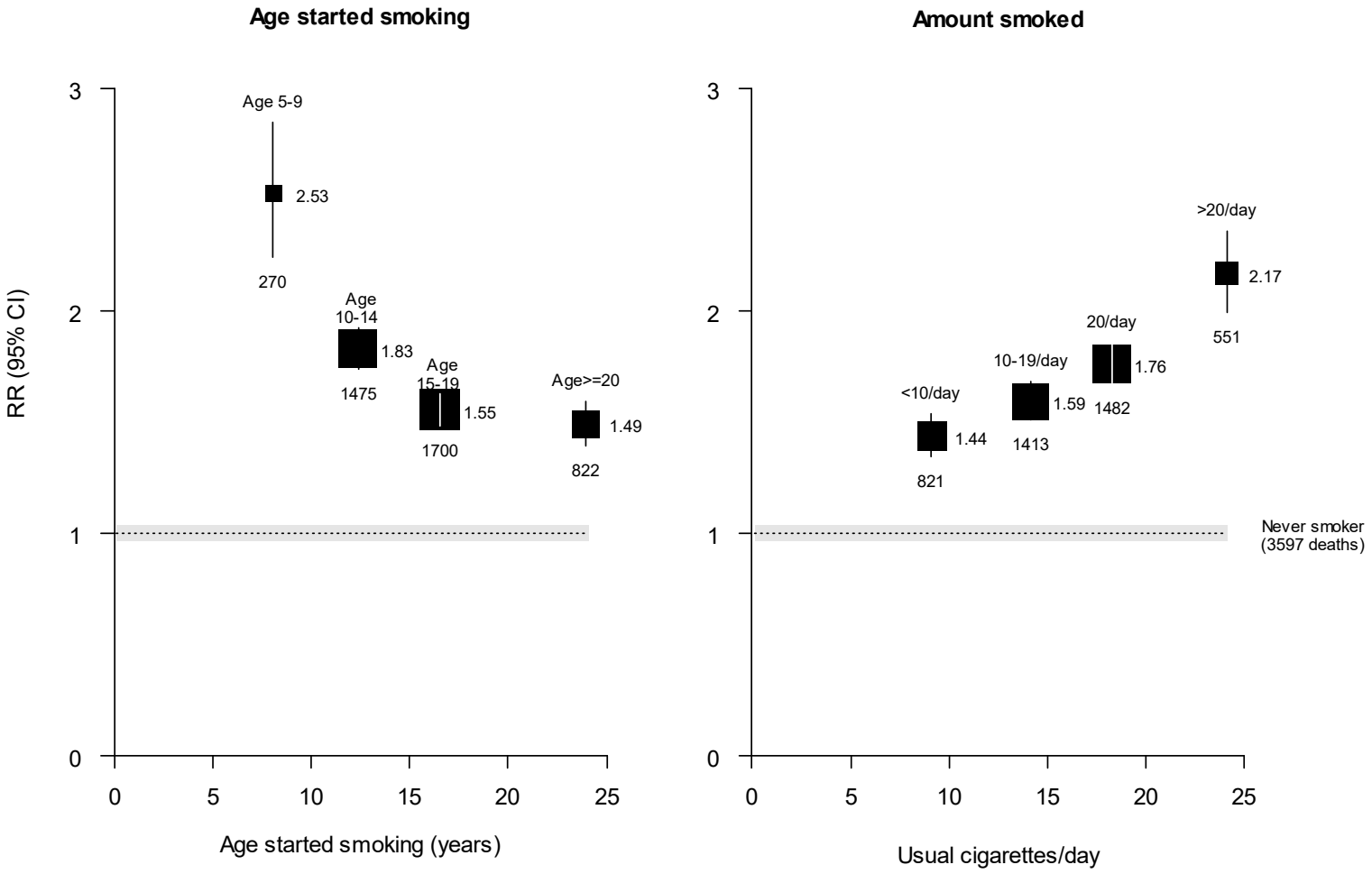


Figure 3.4: All-cause mortality rate ratio at ages 30-69, cigarette smoker vs never smoker, by age started smoking and usual amount smoked

Participants with chronic disease at recruitment were excluded from these analyses.

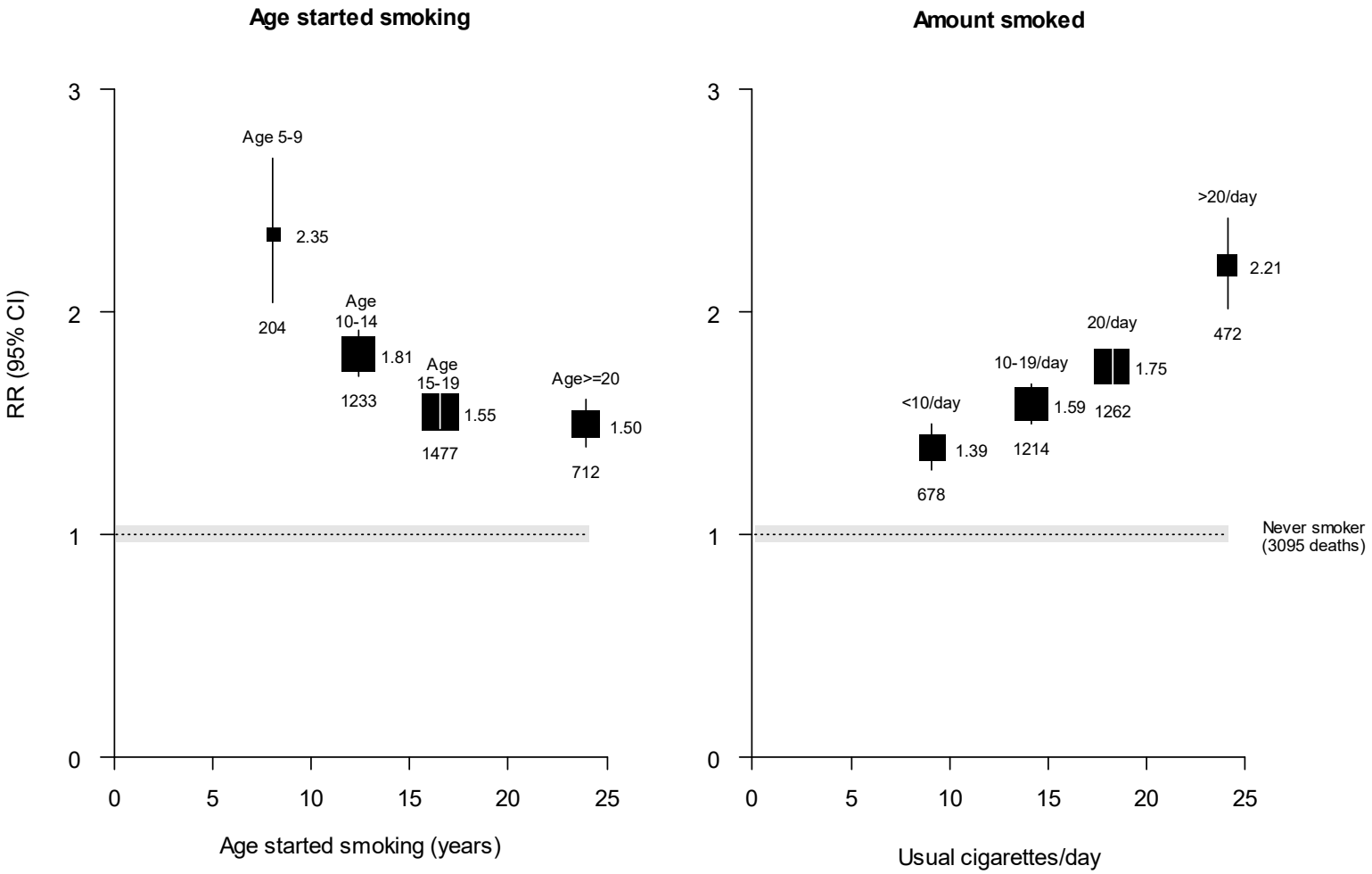


Figure 3.5: All-cause mortality rate ratio at ages 30-69, cigarette smoker vs never smoker, by age started smoking and sex

Participants with chronic disease at recruitment were included in these analyses.

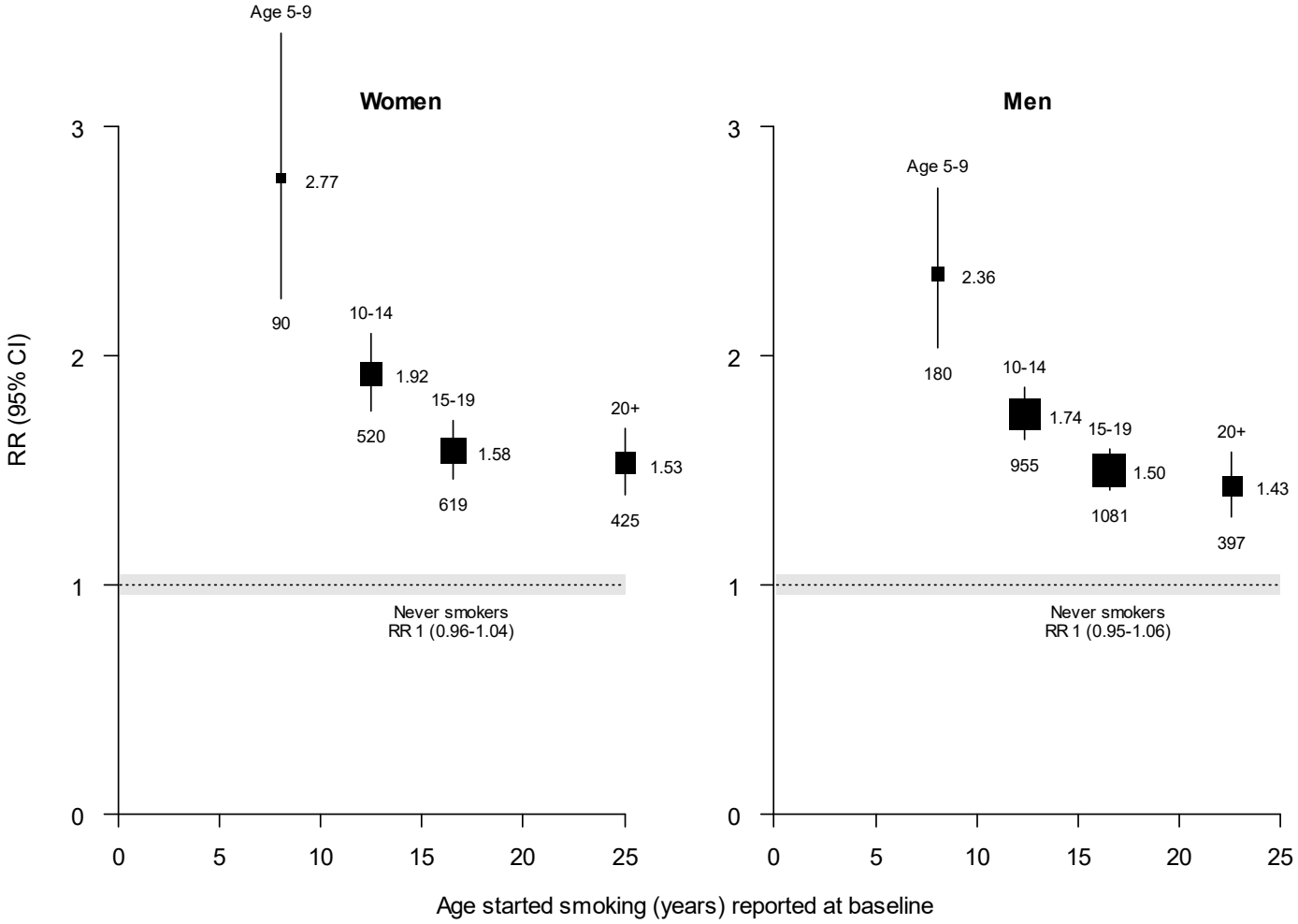


Figure 3.6: All-cause mortality rate ratio at ages 30-69, cigarette smoker vs never smoker, by usual amount smoked and sex

Participants with chronic disease at recruitment were included in these analyses.

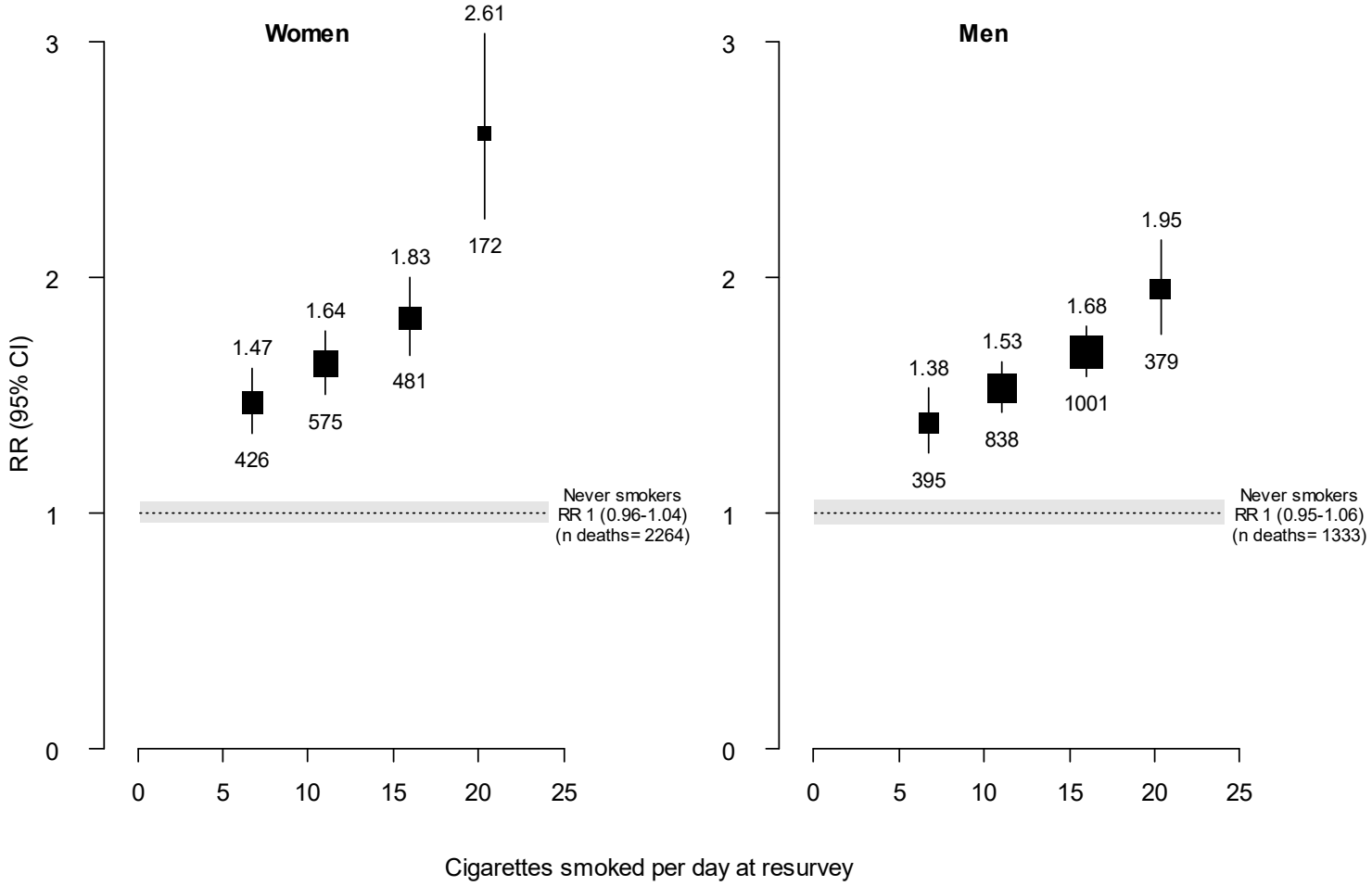
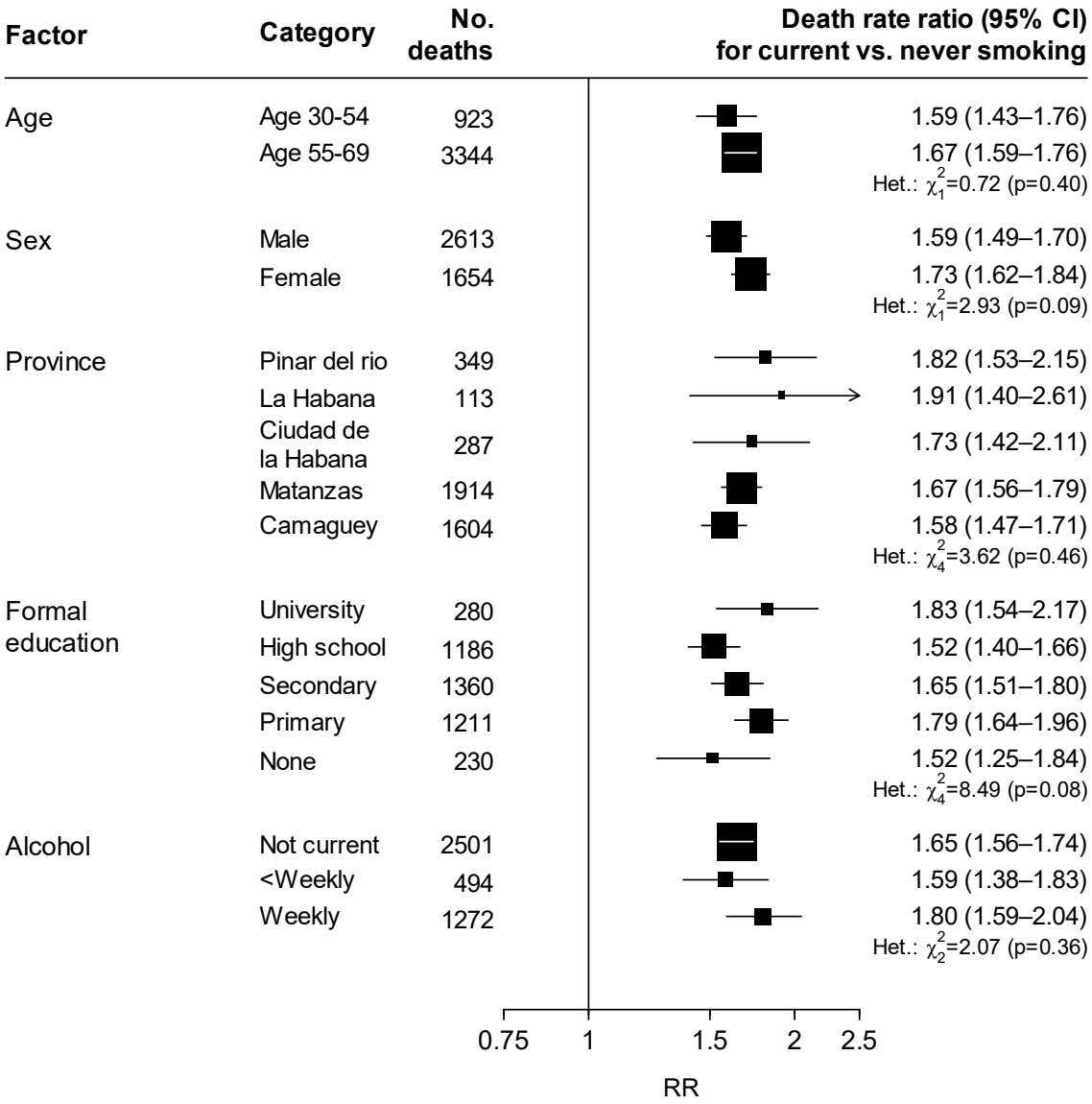


Figure 3.7: Risks associated with current cigarette vs never smoking by levels of factors adjusted for in main analyses

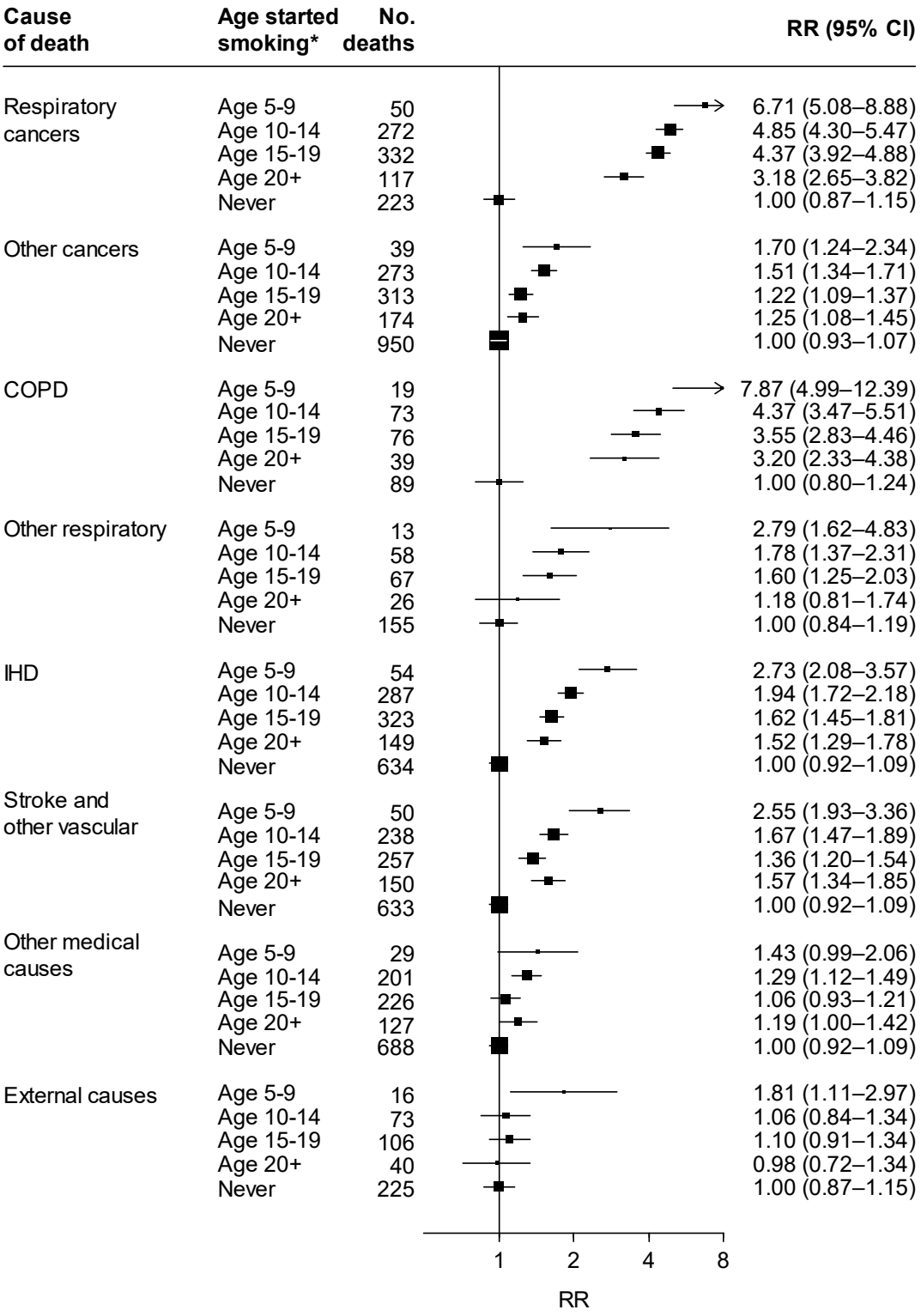
Participants with chronic disease at recruitment were included in these analyses.



RRs estimated separately by levels of each confounder shown, mutually adjusted for each other. Number of deaths among current cigarette smokers shown.

Figure 3.8: Cause-specific mortality rate ratio at ages 30-69, cigarette vs never smoker, by age started smoking

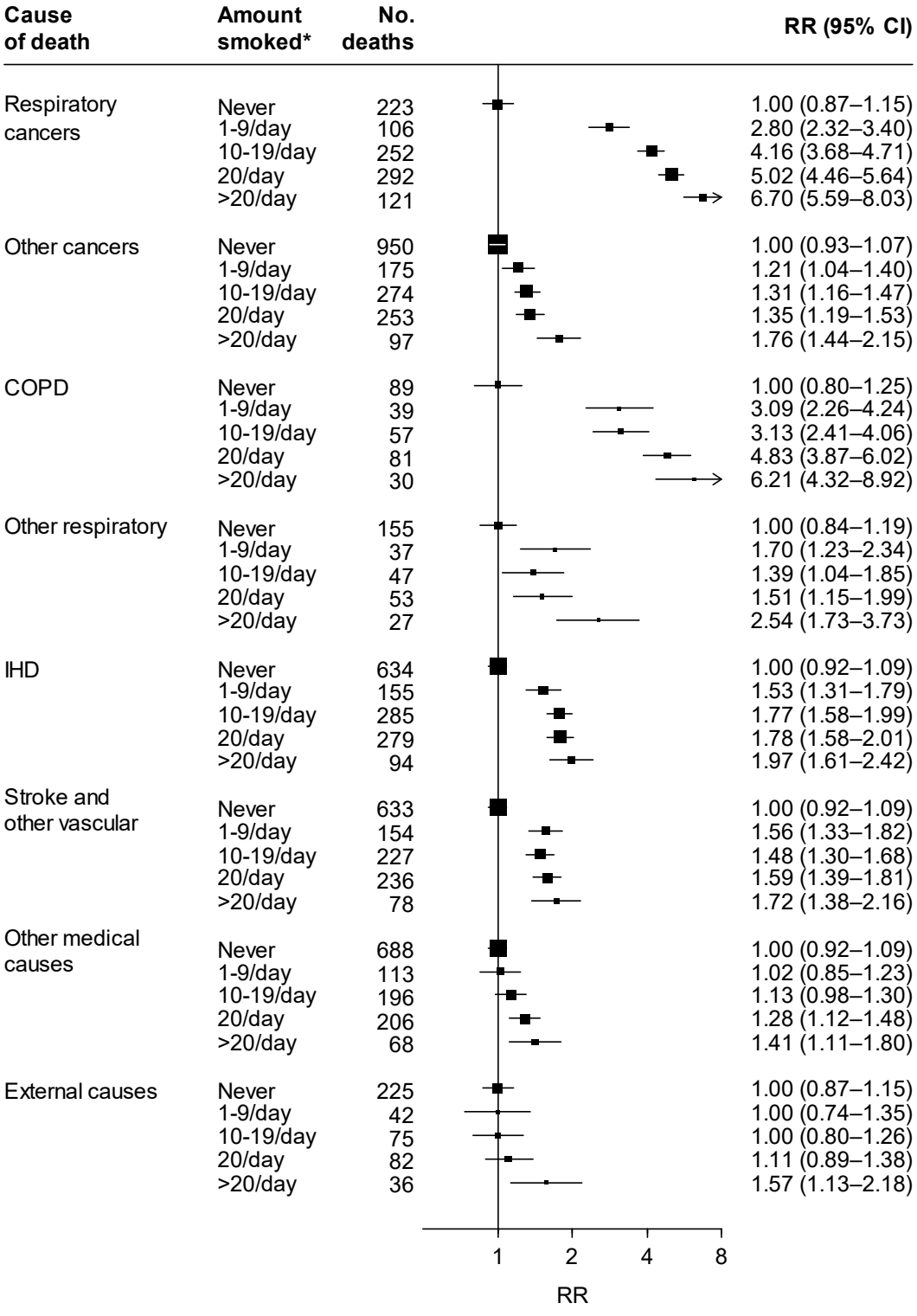
Participants with chronic disease at recruitment were included in these analyses.



*RRs associated with never smoking, compared to starting to smoke on average at ages 8, 12, 17, and 24, respectively.

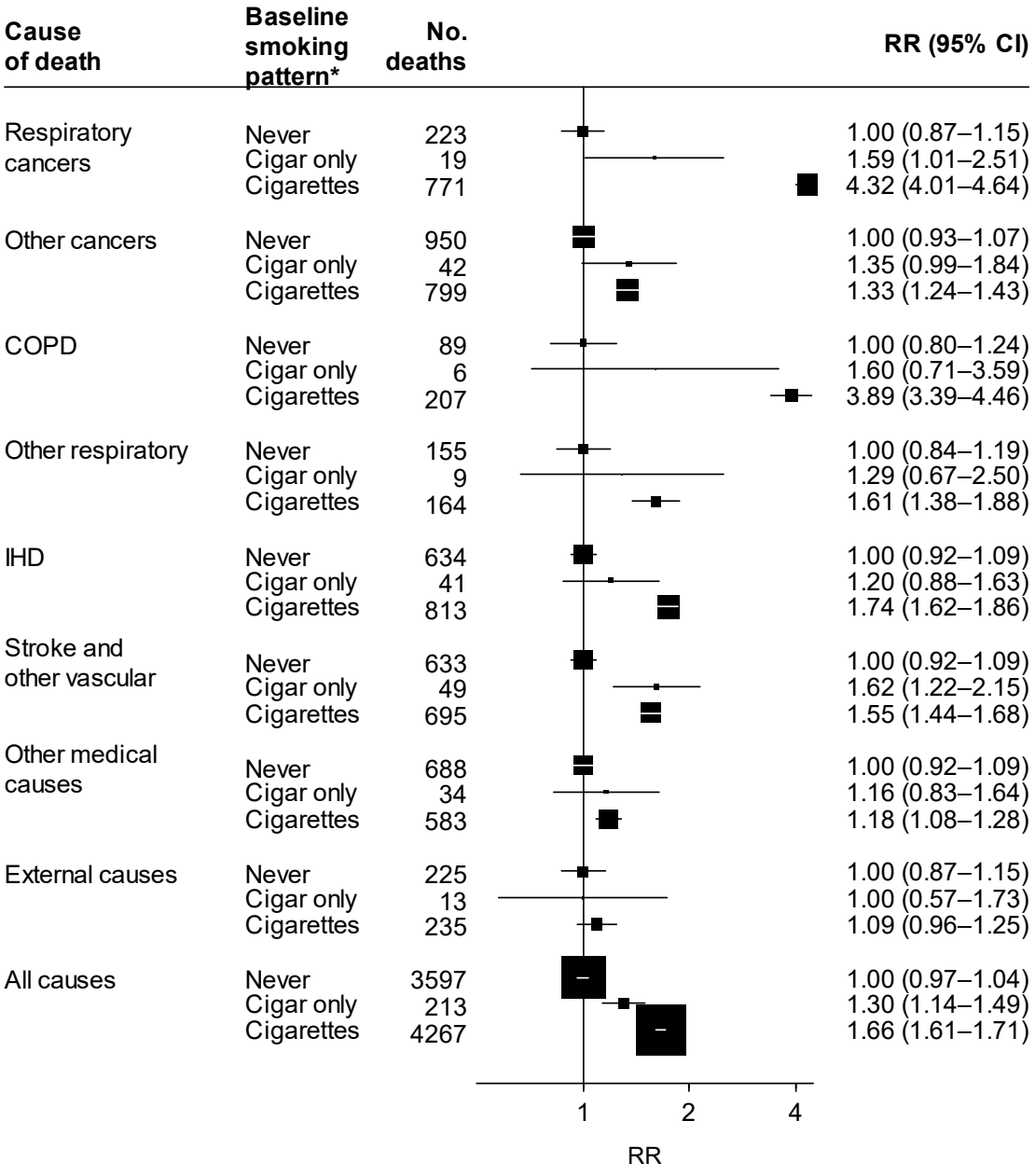
Figure 3.9: Cause-specific mortality rate ratio at ages 30-69, cigarette vs never smoker, by amount smoked

Participants with chronic disease at recruitment were included in these analyses.



*RRs associated with never smoking and smoking 9, 14, 18, and 24 cigarettes per day on average at resurvey, respectively.

Figure 3.10: Cause-specific mortality rate ratio at ages 30-69, cigarette or cigar smoker vs never smoker, by product smoked



*RRs associated with never smoking, smoking 3 cigars/day, or smoking 15 cigarettes/day on average at recruitment, respectively. Among cigarette smokers, 7% additionally smoked ~3 cigars/day.

Figure 3.11: All-cause mortality rate ratio at ages 30-69, cigarette smoker vs never smoker, by age at cessation and years since quitting

Participants with chronic disease at recruitment were excluded from these analyses.

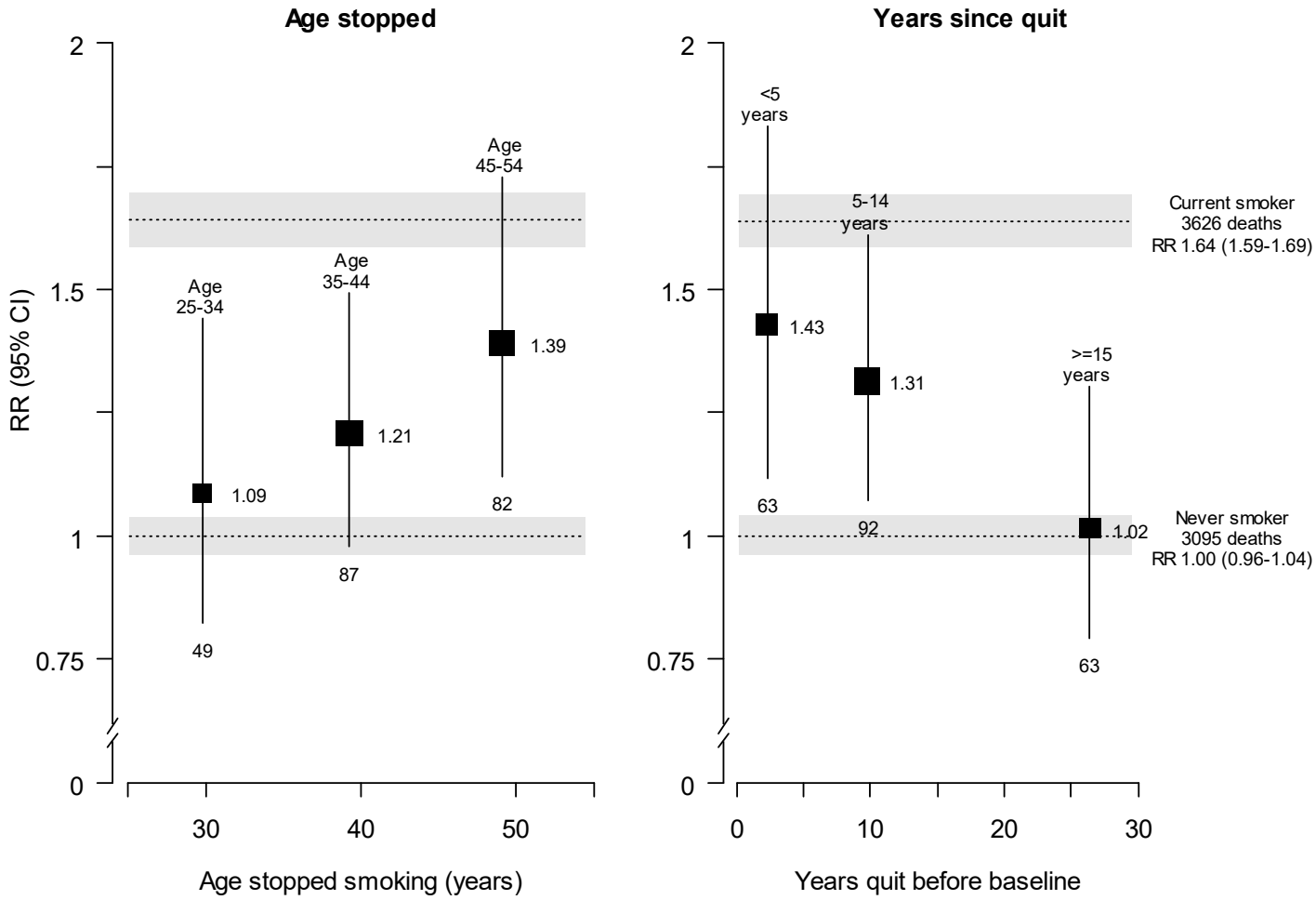
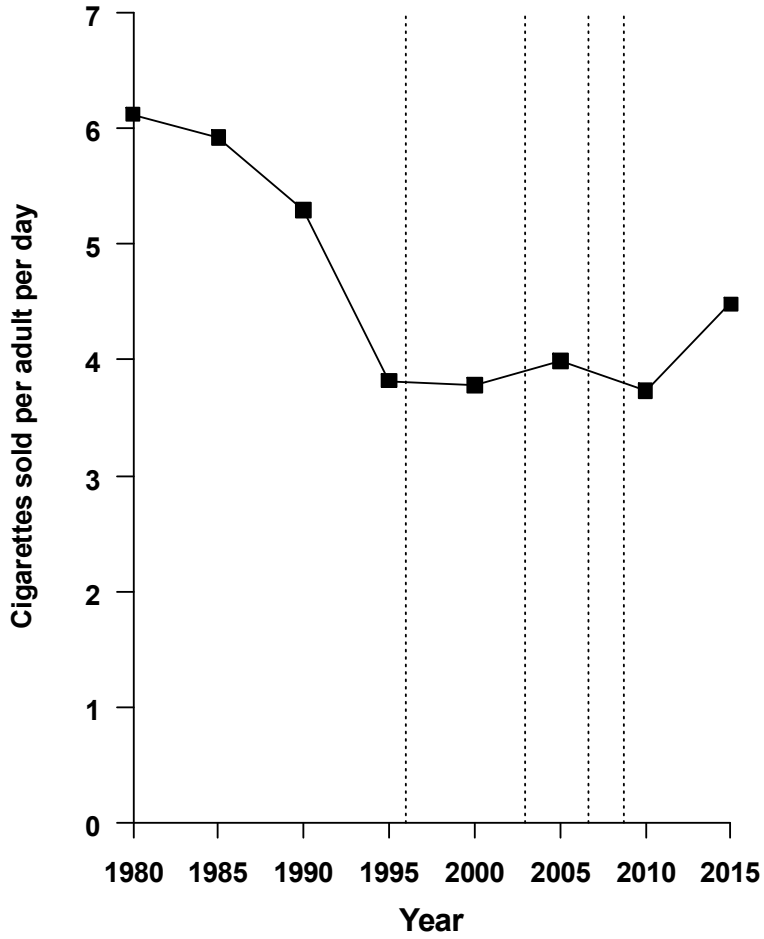


Figure 3.12: Cigarettes sold per adult per day in Cuba, 1980-2015



Vertical dotted lines refer to baseline survey (1996-2002) and resurvey (2006-2008) periods.

Source: Suarez-Lugo N. Consumo de cigarrillos y elasticidad precio-demanda. Cuba 2016. Horizonte Sanitario 2016; 16(3).

Chapter 4: Smoking and mortality in the Mexico City Prospective Study

4.1 Chapter aims

The objective of this chapter is to report the hazards of smoking and benefits of quitting in the Mexico City Prospective Study. Specific aims are:

1) To assess the prevalence of smoking patterns across birth cohorts and by sex
2) To characterise the association between smoking behaviours and all-cause mortality
3) To determine whether other factors may modify the association between smoking and all-cause mortality
4) To characterise the association between smoking behaviours and cause-specific mortality
5) To estimate the smoking-attributable mortality in Mexico
6) To interpret findings in the context of prior evidence of smoking in Mexico and other populations globally

4.2 Baseline survey

4.2.1 Study design

The Mexico City Prospective Study is a large, prospective, population-based study in Mexico City consisting of 159 755 adults aged 35 or older recruited between 1998 and 2004.[102] As with the Cuba Prospective Study, the Mexico City Prospective Study was originally established to determine the health effects of tobacco in a unique population, but grew to include information on other risk factors including alcohol consumption, diabetes, adiposity, and blood pressure. Notably, a 10 mL venous blood sample was taken for 98% of participants, making the study one of the largest blood-based prospective cohorts globally. At 2010 mortality rates, more than one in four Mexican men and one in five Mexican women will die between the ages of 35-69, primarily from

vascular diseases and cancer, as well as diabetes-related causes.[87] The study is a collaboration between researchers at the National Autonomous University of Mexico in Mexico City and the CTSU at the University of Oxford, with support from the Mexican Ministry of Health, the Mexican National Council of Science and Technology, the Wellcome Trust and the UK Medical Research Council. Research ethics approval for the study was granted by the Mexican Ministry of Health, the Mexican National Council for Science and Technology, and the University of Oxford.

4.2.2 Study population

The baseline survey was conducted in two contiguous urban districts of Mexico City: Coyoacán and Iztapalapa. To determine all potentially eligible individuals in these districts, a confidential register of all households in both districts was created through door-to-door interviews collecting the names and ages of all residents. Teams of 2-3 nurses then visited each household on the register and invited all men and women age 35 or older to participate in the study, with surveys being conducted in the home. A total of 112 333 households included at least one eligible inhabitant, of which individuals from 106 069 households ultimately consented (household response rate: 94%). From these households, a total of 159 755 individuals (107 111 women and 52 644 men) agreed to participate.

4.2.3 Data collection

After providing written consent, participants provided information on age, sex, occupation, education, lifestyle habits (smoking, drinking, physical activity, fruit and vegetable consumption), disease history, medication use, and reproductive history for women. Interviewers recorded this information using handheld recorders that represented, at the time, novel technology. These devices standardised data collection procedures by using scanned barcodes to record answers, whereby most information

would be categorised automatically. Most of these answers were already categorical (eg, yes, no, don't know), but some were numeric and stored as ordinal information (further described below). However, most numeric information, including participant age, age started smoking, cigarettes smoked per day, and income, were manually entered. Trained nurses then took a series of physical measurements, including standing height, waist and hip circumference, weight, and seated blood pressure. Finally, a 10 mL venous blood sample was taken, which was labelled with a barcode that was linked to information about the date and time of collection, as well as time of last meal, before being stored and processed at the end of each day (and later sent to Oxford for long-term storage over liquid nitrogen). See Appendix E for baseline questionnaire in English.

4.3 Resurvey

Between 2015 and 2019, on average 16 years after the baseline survey, a subset of surviving participants from both districts was invited to be resurveyed. The resurvey was conducted primarily to correct for variation in biological and behavioural changes, as in Cuba, and therefore collected similar information as at baseline. However, the resurvey was also enhanced to capture information on additional areas of health status to better characterise study participants and to address further hypotheses (see Appendix B for resurvey instrument). A total of 10 142 surviving participants were resurveyed. Unlike in Cuba, where the resurvey was conducted near the midpoint of the follow-up period, the Mexico City resurvey was conducted towards the end of the current follow-up period. Therefore, reported changes in behaviour should be interpreted more cautiously, particularly as resurvey participation was conditional on survival over the 16-year period between recruitment and resurvey. (A resurvey after about 5 years of follow-up had originally been planned, but for a variety of mainly logistical reasons was unfortunately not possible to complete.) See Appendix E for resurvey questionnaire in English.

4.4 Follow-up

Participants have been followed-up to 1 January 2018 through linkage to Mexico's national death registry. In Mexico, nearly all deaths are medically certified. Mortality records were probabilistically matched to study participants through a combination of name and date of birth, as collected at recruitment and as appearing on the death certificate. As there is a degree of uncertainty in probabilistic matching, matches were given a grade (A-E) based on the likelihood of being a correct match. Deaths graded 'A', 'B', and 'C' were those confirmed through fieldwork as being correct, those with an exact match (but not yet checked in the field), and those with a match on name and within 1 year of date of birth (but not yet checked in field), respectively. Deaths graded 'D' and 'E' were those with an age discrepancy of more than 1 year. Participants whose mortality match was graded 'D' or 'E' were excluded from all prospective analyses, as prior analyses of these matched deaths indicated a comparatively weak association between diabetes and mortality risk (which was strongly associated among matches graded A-C), suggesting non-trivial bias could be introduced by their inclusion. However, uncertain vital status at the end of the follow-up period was uncommon (0.8% of eligible participants). Although some matches graded B and C may also be incorrect, analyses suggest that the resulting bias is small (ie, the mortality risk associated with diabetes was similar among those graded A, B, and C). Furthermore, more than 7000 matched deaths were checked in the field, of which more than 95% were found to be correct matches.[93]

All diseases mentioned on the death certificate were categorised according to ICD-10 conventions. Study clinicians reviewed the death certificates of study participants who died during follow-up, and in some cases, recoded the underlying cause of death. This was done, primarily, to ensure that 'diabetes' was accepted as the underlying cause of death only in the case of death due to acute diabetic crisis.[103] The ICD-10 codes for the

categories used in analyses of cause-specific mortality are given in Table 4.1. Although electronic medical records have become routine in Mexico City over the follow-up period, linkage to these databases for non-fatal events is not yet possible. However, a health follow-up survey of all original participants is currently being planned, which will capture information about non-fatal health events (and validate fatal events) that have occurred over the follow-up period.

4.5 Methods

4.5.1 Assessment of smoking

Smoking behaviour was self-reported by participants according to their responses to the standardised questionnaire. Participants were classified as ever-smokers if they answered “Yes” to the question, “Have you ever smoked tobacco?” Those who were ever smokers then reported the age at which they first began smoking regularly, which was manually recorded as an integer. Ever-smokers were then asked, “Do you now smoke regularly?” Participants were considered current smokers if they answered “Yes” in response. Current smokers were subsequently asked with what frequency they smoked, with the accepted options being daily, weekly, monthly, or occasionally. There was some ambiguity as to how participants interpreted the frequency of ‘occasionally’ smoking, which is further discussed below. Those who reported smoking daily were considered daily smokers, while those who smoked less than daily were considered occasional smokers. Current smokers reported the number of cigarettes smoked per day on the days they smoked, which was recorded as an integer, as well as whether they inhaled when smoking, and whether they had ever quit in the past.

Information about smoking cessation was recorded somewhat imprecisely, as it was recorded according to pre-specified barcoded categories. Both current and ex-smokers were asked “How long ago did you quit the last time?” with the options “less than 1 year,

1 to 3 years, or more than 3 years," and "The last time that you quit smoking, how long did you give it up?" with the options "less than 1 year, 1 to 2 years, or more than 2 years." Participants who reported not smoking and having quit more than 3 years prior to recruitment were considered 'long-term' quitters, while those who reported having quit in the past 3 years were considered 'recent' quitters. However, among long-term quitters, it was not possible to determine the age at which they quit or the duration of time since they quit. Ex-smokers also reported the frequency with which they smoked in the past (with the same categories as were used for current smokers), and the number of cigarettes they used to smoke (again recorded as an integer).

Smoking questions at resurvey were generally similar to those at recruitment, but some differences are worth noting. At resurvey current smokers were asked "About how many days per week do you smoke?" which did not allow for smoking 'occasionally' or monthly. Therefore, those who reported smoking 7 days per week were considered daily smokers, those who reported smoking 1-6 days per week were considered 'occasional' smokers, and those who reported smoking zero days per week were considered ex-smokers. To determine amount smoked, current smokers were asked "On the days that you smoke, how many cigarettes do you normally smoke in a day?" At resurvey, ex-smokers were asked how many years ago they last quit, the age at which they last quit, and whether they quit due to illness. Information about the frequency and amount previously smoked was assessed as it was for current smokers.

4.5.2 Selection of covariates

The process for selecting covariates in each set of prospective analyses in this thesis was described previously in **Chapter 3**. Although the same set of demographic, socioeconomic, and behavioural factors were adjusted for as in **Chapter 3**, the exact

categories used to adjust for confounding differed slightly. The specific categorisations are defined below in section 4.5.4.

4.5.3 Participant exclusions

Of the 159 755 participants recruited into the study, 854 participants were aged 90 or older at recruitment and were excluded from all analyses, as were participants with missing information on smoking or other covariates of interest (n=213). Unlike in Cuba, the main analyses were focused on those aged 35-89 (but shown at ages 35-69 and 70-89 where relevant), in part because findings were similar at younger and older ages (shown below), and in part to improve the precision of cause-specific estimates. Also distinct from the Cuba cohort, in the Mexico City cohort major chronic disease (primarily diabetes) was common at recruitment, and presented an extremely high competing risk of mortality. Therefore, prospective analyses were initially conducted separately in those with and those without chronic disease at recruitment. The rest of the main prospective analyses were limited to those without chronic disease at recruitment following the initial findings, excluding 27 897 participants from subsequent analyses (further discussed below).

4.5.4 Statistical analysis

As in **Chapter 3**, cross-sectional analyses of the prevalence of smoking behaviours are unadjusted, and are presented by sex and birth cohort (where relevant). For prospective analyses, Cox regression was used to calculate mortality rate ratios (RRs) and accompanying 95% confidence intervals (CIs) according to categories of smoking at recruitment. Participants were followed up to the date of death, or else censored at the end of the risk period under consideration (eg, age 90 for most analyses) or the end of the follow-up period (1 January 2018). As described above, analyses were primarily conducted at ages 35-69 and 70-89 for all-cause mortality, and ages 35-89 for cause-

specific mortality. As in **Chapter 3**, prospective analyses were stratified by age-at-risk (in 5-year age groups), sex, and district, and further adjusted for formal education completed (<primary, primary, secondary, or college/university) and frequency of alcohol consumption (never, former, <weekly, 1-2 days/week, 3 or more days/week).

Analyses were initially conducted to determine the relevance of smoking less than daily, 1-9 cigarettes/day, and 10 or more cigarettes/day to all-cause mortality. Cross-sectional analyses determined that non-daily smoking was common in this cohort, and was difficult to quantify (eg, whether "occasional" smoking could mean one day per week or one day per year). Therefore, as non-daily smoking was considered likely to be reflective of very different long-term habits (and risks) compared with those who smoked daily (further discussed in the results section below), the main analyses were restricted to those who reported smoking daily at recruitment, with non-daily smokers considered in analyses as a separate group.

The association between cigarettes smoked per day and all-cause and cause-specific mortality was assessed by categorising daily smokers according to their reported cigarettes smoked per day at baseline: <10 or ≥10 cigarettes per day. These cut-offs were determined in part to accommodate for digit preference found in the raw data for daily smokers (Figure B1 in Appendix B), and in part for consistency with methods used in the Cuban cohort. Although baseline amount smoked was used to categorise smokers, the RRs associated with smoking at these baseline levels were plotted at the number of cigarettes smoked per day at resurvey (around 16 years later) among those who continued to smoke (ie, the 'usual' amount smoked). Although smoking cessation was common over the follow-up period, because the resurvey was conducted at the end of follow-up, the 'usual' amount smoked did not take into account those who had quit entirely (ie, did not assign a value of zero cigarettes/day among ex-smokers) in order to

avoid over-correction for these habits (further discussed below). The association between age started smoking and mortality was assessed by categorising daily smokers into two groups by the age at which they began smoking regularly: <18, or ≥18 years. These cut-offs were chosen to accommodate for digit preference, and to produce two roughly even groups in order to maximise statistical power for the comparison between these two groups.

Analyses of smoking cessation were limited due to the imprecise nature of the information collected at recruitment. To assess the relevance of smoking cessation to mortality, RRs by years since quitting prior to recruitment (1-3 years and >3 years) were calculated. However, the effects of 'heavier' versus 'lighter' past smoking behaviour was not examined, as duration of prior smoking and exact time since quitting were not known, and the effects of smoking 10 cigarettes per day for 40 years may be very different from that of smoking 20 cigarettes per day for 20 years (as discussed in **Chapter 1**).[41,42] These analyses were conducted separately in those with and without chronic disease at recruitment in order to assess the potential impact of the 'sick-quitter' effect. While reasons for quitting were not reported at recruitment, the resurvey collected information on whether ex-smokers had quit because they were already ill, providing information on a subset of participants who had survived over the follow-up period (~16 years later).

As in **Chapter 3**, group-specific CI's are plotted for comparisons involving at least three categories of exposure, but conventional CI's are reported for any direct comparisons made between two groups (and are used to describe comparisons in text). Similarly, the proportional hazards assumption was examined by estimating RRs in separate 5-year periods (ie, 0-5, >5-10, >10-15, and >15-20 years after recruitment). Methods for calculating the population attributable fractions (PAFs) were again similar to those used

in **Chapter 3** (using the RRs for ever smoking associated with all medical causes of death). However, in this chapter the overall PAFs are calculated at ages 35-89 to improve the precision of the estimates, but the PAFs associated with 'premature' mortality (deaths at ages 35-69) are shown in Appendix B.

4.6 Results

4.6.1 Participant characteristics

Descriptive analyses were initially conducted on all participants aged 35-89 at recruitment in the Mexico City Prospective Study, after excluding only those with missing information on smoking or covariates of interest (as described above), resulting in 158 688 participants. The characteristics of these participants are shown in Table 4.2. Of included participants, 138 045 were age 35-69 at recruitment ('middle-aged' participants), and 20 643 were age 70-89 ('older' participants). Overall, the mean age of participants was 52 years, and 33% were male. More participants were recruited in the district of Iztapalapa than in Coyoacan (60% and 40%, respectively), which was similar among both middle-aged and older participants and men and women. Overall, 39% of the cohort completed high school, but clear generational differences were observed: 10% of older women and 18% of older men completed high school, compared with 39% of middle-aged women and 54% of middle-aged men. Among middle-aged participants, 12% had a self-reported history of diabetes, 1% had ischaemic heart disease (IHD), and 16% had any chronic disease at recruitment. Among older participants, 23% had a self-reported history of diabetes, 4% had IHD, and 31% had any chronic disease at recruitment. Within age groups, the prevalence of chronic disease was similar between men and women.

4.6.2 Trends in smoking behaviour across birth cohorts

The prevalence of ever smoking among men was similar across birth cohorts (approximately 80%), while the prevalence of ever smoking was higher among women in more recent birth cohorts: approximately 20% among women born in ~1920 and more than 40% among women born in the 1960's (Figure 4.1). However, smoking cessation was common among both men and women, and was inversely associated with year of birth (ie, was more common among older participants). Among men born before 1940 and women born before 1945, more than half of ever-smokers had quit smoking by recruitment. Overall, 47% of ever smokers had quit by recruitment, which was similar between men and women. Among current smokers at recruitment, 35% were non-daily smokers (39% of women and 32% of men). However, non-daily smoking was more common among middle-aged than among older participants. Among current smokers born before 1930, 28% of women and 22% of men were non-daily smokers, while among current smokers born after 1960, 43% of women and 38% of men were non-daily smokers.

Among daily smokers, the mean number of cigarettes smoked per smoker per day was similar across birth cohorts (Figure 4.2). Among men, consumption was approximately 8 cigarettes per day among men born before 1930 or born after 1950, but was slightly higher (10 per day) among men born in the 1940's. Among women, consumption was approximately 6-7 cigarettes per day among women born before 1950, and was somewhat lower (5 per day) among women born after 1960. The mean age at smoking initiation differed notably between men and women. Among male daily smokers, the mean age at smoking initiation was somewhat lower among younger birth cohorts than among older birth cohorts (17 and 18 years, respectively). Among female daily smokers, however, the mean age at smoking initiation was 27 years among women born before 1930, but just 19 years among women born after 1965.

4.6.3 Characteristics of participants at baseline and resurvey by smoking status

Table 4.3 shows the characteristics of the 130 791 participants aged 35-89 and without chronic disease at recruitment according to baseline smoking behaviour. (Table B1 in Appendix B shows characteristics for the 27 897 participants with chronic disease at recruitment.) At recruitment, 30% of men and 13% of women reported daily smoking, and a further 15% of men and 8% of women reported non-daily smoking. Among both male and female current smokers, those who smoked less than daily were on average younger, more likely to be resident in Iztapalapa, less likely to inhale when smoking, and less likely to live with a smoker, than those who smoked daily. Compared with never smoking men, men who smoked daily were less likely to have completed college/university (22% and 31%, respectively), while women who smoked daily were more likely to have completed college/university than never smoking women (18% and 10%, respectively).

Among daily smokers, 63% of men and 77% of women smoked <10 cigarettes/day. Compared with 'lighter' smokers (<10 cigarettes/day), 'heavier' smokers (≥10 cigarettes/day) on average started at a younger age (mean 16.8 [SD 5.3] vs. 18.2 [SD 6.6], respectively, among men, and 19.8 [SD 7.0] vs. 21.6 [SD 8.5], respectively among women), were more likely to inhale while smoking (84% vs. 76% among men and 75% vs. 67% among women), and had a lower body mass index (mean 26.9 [SD 4.4] vs. 27.6 [SD 4.4], respectively, among men and 27.9 [SD 5.3] vs. 28.9 [SD 5.2], respectively, among women). Overall, the mean number of cigarettes smoked/day among daily smokers was 9 (SD 8) in men and 6 (SD 6) in women.

Among both men and women, ex-smokers who were 'recent' quitters (quit in the past 1-3 years) and ex-smokers who were 'long-term' quitters (quit >3 years prior to resurvey) were generally similar within each sex, including amount previously smoked and age

started smoking (Table 4.3). One notable difference was that recent quitters were on average younger than long-term quitters at recruitment: mean age 50 (SD 12) and 56 (SD 14) years among men, respectively, and 47(SD 11) and 51 (SD 13) years among women, respectively. Table B2 in Appendix B shows the baseline smoking characteristics for those aged 35-69 at recruitment without chronic disease.

Table 4.4 shows the smoking characteristics at resurvey (rows) according to baseline smoking behaviours (columns) an average of 16 years after recruitment. Among participants who smoked ≥10 cigarettes/day at recruitment, 68% reported continuing to smoke daily, and the average number of cigarettes smoked per day per smoker decreased from 15.4 (SD 7.2) to 10.5 (SD 6.9). Of these, 26% reported quitting smoking, of whom half reported quitting due to illness (49/101). Among participants who smoked daily but <10 cigarettes/day at recruitment, 39% at resurvey reported continuing to smoke daily, and the average number of cigarettes smoked per day per smoker was unchanged (approximately 4 cigarettes/day). Of these, 42% reported quitting smoking over the 16-year period, of whom 34% reported quitting due to illness (147/436). When considering those who had formerly smoked daily but had quit by resurvey as smoking zero cigarettes/day at resurvey, the mean number of cigarettes smoked at resurvey was 2.1 (SD 3.4) among those smoking <10/day at recruitment, and 7.5 (SD 7.5) among those smoking ≥10 cigarettes/day at recruitment.

Few participants who reported non-daily smoking at recruitment continued to smoke at resurvey, with 8% reportedly smoking daily, 18% still smoking less than daily, and 72% not smoking at all (2% provided incomplete information for categorisation). Few participants who reported never or ex-smoking at recruitment had changed their behaviour by resurvey: 1% of never smokers and 6% of ex-smokers reported any smoking (less than daily or daily) at resurvey. Of the 65% of ex-smokers who answered

whether they had quit due to illness at resurvey (a quarter reported being never smokers, and did not answer the question), one in five reported having quit due to illness (221/1164). Tables B3 and B4 in Appendix B show resurvey values among those aged 35-69 and without chronic disease at recruitment, and among those age 35-89 with chronic disease at recruitment, respectively.

4.6.4 Association between smoking and mortality among those with and without chronic disease at recruitment

Among the 158 688 adults initially contributing to the overall analyses, there were 2.3 million person-years of follow-up at ages 35-89 (mean 15 SD 4 per person), during which there were 19 091 deaths (10 773 among participants without diabetes or other chronic disease at recruitment, and 8318 among participants with diabetes or other chronic disease). There were notable differences in the relevance of daily smoking to mortality among those with versus those without chronic disease at recruitment (Figure 4.3). Among participants without chronic disease at recruitment, compared with never smokers, the all-cause mortality rate ratio (RR) for ex-smokers who quit >3 years prior to recruitment ('long-term quitters') was 1.11 (95% CI 1.06-1.17) and for those who quit 1-3 years prior to recruitment ('recent quitters') was 1.14 (1.02-1.27), while the RR for non-daily current smokers was 0.98 (0.91-1.07), and for daily smokers was 1.17 (1.10-1.25) among those smoking <10 cigarettes/day and 1.54 (1.43-1.67) among those smoking ≥10 cigarettes/day. Among those with chronic disease at recruitment, compared with never smokers, the all-cause RR for ex-smokers who were long-term quitters was 1.07 (1.01-1.13) and for recent quitters was 1.18 (1.06-1.32), while the RR for non-daily smoking was 0.96 (0.87-1.05), and for daily smokers was 1.00 (0.92-1.09) among those smoking <10 cigarettes/day and 1.09 (0.98-1.22) for those smoking ≥10 cigarettes/day. However, the absolute risk of mortality at ages 35-89 was considerably higher in those

with chronic disease at recruitment than among those without chronic disease regardless of smoking category.

It was hypothesized that this difference in the relevance of daily smoking to mortality between those with and those without prior chronic disease may be due in part to differences in the causes of death resulting from these pre-existing chronic diseases (in particular deaths associated with a baseline history of diabetes), and in part due to changes in smoking behaviour as a result of these potentially serious illnesses (further explored below). Although the RRs associated with ex-smoking and with non-daily smoking were similar among those with and without prior chronic disease, the main exposure of interest was daily smoking, which differed materially between these two groups. Therefore, to better capture the true relationship between daily smoking and mortality, subsequent analyses for all-cause and cause-specific mortality were restricted to the 138 045 participants without chronic disease at recruitment. However, sensitivity analyses showing the association between daily smoking and cause-specific mortality among those with chronic disease at recruitment are shown in Appendix B.

4.6.5 Association between daily smoking and mortality among participants without prior chronic disease at recruitment

4.6.5.1 All-cause mortality

Among those without chronic disease at recruitment (as in the rest of analyses for the Results section, except where indicated), compared with never smokers, the all-cause mortality RR at ages 35-89 among current daily smokers ('daily smokers') was 1.27 (1.19-1.35). However, the excess risk of death associated with daily smoking depended on the amount smoked and age started smoking. Compared with never smokers, the all-cause mortality RRs at ages 35-69 associated with smoking <10 and ≥10 cigarettes/day were 1.09 (1.00-1.20) and 1.52 (1.36-1.70), respectively, and at ages 70-89 they were 1.24 (1.13-1.37) and 1.55 (1.37-1.74), respectively (Figure 4.4). While the RRs by

amount smoked were similar at younger and older ages (particularly ≥10 cigarettes/day), the RRs for age started smoking depended on the age at risk. Compared with never smokers, the all-cause mortality RRs at ages 35-69 associated with starting to smoke regularly at age <18 and ≥18 years were 1.32 (1.20-1.46) and 1.14 (1.03-1.26), respectively, while at ages 70-89 they were 1.32 (1.19-1.47) and 1.35 (1.21-1.49) (Figure 4.5). For both the 35-69 and 70-89 age-at-risk groups, the association between amount smoked and all-cause mortality was similar in men and women (Figure 4.6). However, while smoking initiation before age 18 was associated with higher risk of all-cause mortality at ages 35-69 than smoking initiation at age 18 or older among men and women, age at smoking initiation had little relationship with mortality at ages 70-89 in men and women (Figure 4.7).

The effect of smoking according to levels of each confounder adjusted for in the main analyses is shown in Figure 4.8. Overall, the associations of daily smoking versus never smoking with all-cause mortality at ages 35-89 were similar at ages 35-59, 60-74, and 75-89 (χ_2^2=1.25, p=0.54), in men and women (χ_1^2=0.53, p=0.47), and across levels of educational attainment (χ_3^2=1.68, p=0.64) and alcohol consumption (χ_4^2=6.06, p=0.19). However, significant heterogeneity was observed by district, where the mortality risk associated with daily versus never smoking was higher among those resident in Iztapalapa than among those resident in Coyoacán (χ_1^2=6.65, p=0.01). There was little evidence that these all-cause mortality RRs comparing daily and never smokers varied over the follow-up period; the RRs at 0-5, >5-10, >10-15, and >15-20 years after recruitment were 1.34 (95% CI 1.20-1.51), 1.30 (1.17-1.44), 1.25 (1.13-1.38), and 1.23 (1.00-1.51), respectively.

4.6.5.2 Cause-specific mortality

As hypothesized, the causes of death differed substantially between those with and those without chronic disease at recruitment. Among those without chronic disease at recruitment, there were 6945 deaths among never or daily smokers at ages 35-89. Of these, 2255 were vascular (1554 ischaemic heart disease [IHD], 519 stroke, and 182 other vascular), 1438 neoplastic (160 lung or upper-aerodigestive cancer ['respiratory cancer'], 1278 other), 971 respiratory (375 chronic obstructive pulmonary disease [COPD], 596 other), 518 renal, 560 hepatobiliary, 412 gastrointestinal, 108 acute diabetic, 440 from other medical causes (or ill-defined), and 243 from external causes.

Among those with chronic disease at recruitment, there were 5058 deaths among never or daily smokers at ages 35-89. Of these, 1689 were vascular (1191 ischaemic heart disease [IHD], 390 stroke, and 108 other vascular), 506 neoplastic (43 lung or upper-aerodigestive cancer ['respiratory cancer'], 463 other), 550 respiratory (150 chronic obstructive pulmonary disease [COPD], 400 other), 1151 renal, 228 hepatobiliary, 218 gastrointestinal, 329 acute diabetic, 309 from other medical causes (or ill-defined), and 78 from external causes. Although the proportion of deaths by cause among those with versus without chronic disease at recruitment was similar for vascular deaths (33% and 32% of deaths, respectively), it differed substantially for several causes, including neoplastic deaths (10% and 21%, respectively), renal deaths (23% and 7%, respectively), and acute diabetic deaths (7% and 2%, respectively).

Among those without diabetes or other chronic disease at recruitment, dose-response relationships were observed with the number of cigarettes smoked per day for several causes of death at ages 35-89, including respiratory cancer, COPD, gastrointestinal, cardiac, and stroke (Figure 4.9). Compared with never smokers, the RRs for smoking <10 and ≥10 cigarettes/day, respectively, were 1.59 (95% CI 1.01-2.50) and 6.15 (4.14-

9.14) for respiratory cancers, 2.20 (1.65-2.92) and 5.84 (4.42-7.72) for COPD, 1.40 (1.06-1.84) and 2.20 (1.61-3.02) for gastrointestinal, 1.30 (1.13-1.49) and 1.52 (1.28-1.81) for IHD, and 1.24 (0.96-1.58) and 1.46 (1.07-2.01) for stroke. There was also some evidence of higher risk among those smoking ≥10 cigarettes/day compared with never smoking for deaths from other vascular causes, other cancers, and other respiratory causes. Dose-response relationships with age at smoking initiation were also observed for several causes of death, although the evidence was less consistent (Figure 4.10). Compared with never smokers, the RRs for starting to smoke regularly at age <18 and ≥18 were, respectively, 4.01 (2.70-5.97) and 2.12 (1.35-3.33) for respiratory cancers, 3.38 (2.55-4.48) and 3.15 (2.38-4.17) for COPD, 1.79 (1.34-2.40) and 1.50 (1.11-2.02) for gastrointestinal, 1.43 (1.23-1.67) and 1.31 (1.12-1.53) for IHD, and 1.29 (0.97-1.71) and 1.32 (1.01-1.72) for stroke. Although there was limited power to assess the relevance of the amount smoked and age started smoking to mortality at younger and older ages separately, restricting analyses to death at ages 35-69 resulted in associations similar to those observed at ages 35-89 (Figures B2-3 in Appendix B).

Among those with chronic disease at recruitment, the relationship between the number of cigarettes smoked per day and cause-specific mortality was less clear. Compared with never smokers, the RRs for smoking ≥10 cigarettes/day were 3.81 (1.59-9.12) for respiratory cancers, 3.59 (2.21-5.81) for COPD, 2.44 (1.36-4.38) for other vascular, and 1.76 (1.24-2.49) for other respiratory (Figure B4 in Appendix B). However, the amount smoked as reported at recruitment was little associated with gastrointestinal, IHD, stroke, or other cancers, and was associated with a reduced risk of renal mortality. Similarly, compared with never smokers, the RRs for smoking initiation at age <18 were 4.60 (2.22-9.50) for respiratory cancers, 1.78 (1.07-2.96) for COPD, 1.81 (1.04-3.16) for other vascular, and 1.47 (1.08-2.01) for other respiratory (Figure B5 in Appendix B).

However, age at smoking initiation was little associated with gastrointestinal, IHD, stroke, or other cancers, and was also associated with a reduced risk of renal mortality.

4.6.6 Association between smoking cessation and all-cause and cause-specific mortality among those without chronic disease at recruitment

Among those without chronic disease at recruitment, compared with never smokers, the all-cause mortality RR at ages 35-89 for ex-daily smokers who quit >3 years prior to baseline was 1.15 (95% CI 1.08-1.22), while the RR for continuing daily smokers at recruitment was 1.28 (1.21-1.36). Quitting smoking >3 years prior to recruitment was associated with lower cause-specific mortality in comparison to continued daily smoking for several causes, including respiratory cancers, COPD, gastrointestinal, IHD, and other vascular causes (Figure 4.11). Compared with never smokers, the RRs for quitting >3 years prior to recruitment and continued smoking, respectively, were 1.31 (0.86-2.01) and 2.91 (2.03-4.16) for respiratory cancers, 1.40 (1.08-1.83) and 3.22 (2.53-4.09) for COPD, 1.21 (0.95-1.54) and 1.61 (1.27-2.05) for gastrointestinal, 1.27 (1.12-1.43) and 1.37 (1.21-1.55) for IHD, and 1.19 (0.83-1.70) and 1.31 (0.92-1.87) for other vascular causes. For causes of death little associated with daily smoking, ex-daily smoking was also little associated with mortality.

Among those without chronic disease at recruitment, smoking cessation in the 1-3 years prior to recruitment was associated with somewhat higher risk of mortality than smoking cessation >3 years prior to recruitment for respiratory cancers and COPD (Figure 4.12). However, the risk of mortality among ‘recent’ and ‘long-term’ quitters was similar for other causes of death. Among those with chronic disease at recruitment, the all-cause mortality RR was also generally similar among ‘recent’ and ‘long-term’ quitters, with two notable exceptions: compared with never smokers, the RRs for quitting in the last 1-3 years were 7.14 (3.26-15.65) for respiratory cancer and 1.32 (1.04-1.68) for renal causes of death (Figure B6 in Appendix B).

As it was not known precisely when those who had quit >3 years prior to recruitment had stopped smoking, the relevance of quitting smoking at different ages, or for more precise periods of time, could not be assessed. Similarly, although amount previously smoked was reported, this information was not used, as its interpretation could be misleading without information on quitting duration. (For example, among those who had quit >3 years prior to recruitment, some may have quit 5 years prior after previously smoking 10 cigarettes/day, while others may have quit 40 years prior after previously smoking 20 cigarettes/day.)

Given that many participants who smoked daily at recruitment had quit smoking by resurvey (on average 16 years later), and more than one-third of resurveyed daily smokers who had quit by resurvey did so due to illness, the benefits of smoking cessation in comparison to continued daily smoking are likely underestimated.

4.6.7 Population attributable fraction of mortality associated with ever smoking

If the association between smoking and mortality at ages 35-89 was mostly or wholly causal, then approximately one in 10 male deaths and one in 25 female deaths at ages 35-89 in this population were caused by smoking (Table 4.5). The difference in population attributable fractions (PAFs) was driven by higher prevalence of ever-regular smoking among men than among women, as the RR for all death from any medical cause was similar in men and women. Among all men (ie, including both ever smokers and never smokers), the PAF by cause was highest for COPD (47%), respiratory cancers (43%), other vascular (30%), and stroke (23%). Among women, the PAF was highest for respiratory cancers (19%), COPD (15%), and gastrointestinal causes (8%).

Among ever-smokers, smoking caused 12% of deaths from any medical cause in both men and women (calculated using the RRs in Table 4.5). Among ever-smoking men,

smoking caused more than half of COPD deaths (53%), half of respiratory cancer deaths (49%), and one-third of gastrointestinal deaths (35%). Among ever-smoking women, smoking caused 42% of COPD deaths, 41% of respiratory cancer deaths, and 24% of gastrointestinal deaths.

The PAFs associated with premature mortality (ie, mortality at ages 35-69) from specific causes could not be assessed reliably by sex. However, these estimates are broadly consistent with those at ages 35-89, and are shown in Table B5 in Appendix B.

4.7 Discussion

4.7.1 Summary of findings

Of the 159 755 Mexican adults recruited into the study between 1998 and 2004, approximately half of men and one-quarter of women smoked at recruitment. Of these current smokers, however, one-third smoked less than daily, and consumption among daily smokers was on average just 9 cigarettes/day among men and 6 cigarettes/day among women at recruitment. Adjusting for age-at-risk, sex, education, district, and alcohol consumption, daily smokers without chronic disease at recruitment had 28% higher risk of all-cause mortality at ages 35-89 than those who had never smoked. Smoking as few as 4 cigarettes/day on average at recruitment was associated with 17% higher risk of all-cause mortality at ages 35-89 compared with those who had never smoked. These are underestimates of the true risks of continued daily smoking, however: of those who did smoke daily at recruitment, 40% of those resurveyed had quit by the time of resurvey an average of 16 years later. Ex-daily smokers who had quit >3 years prior to recruitment had approximately half the excess risk of death compared with those who continued to smoke daily.

Assuming that the excess mortality among ever smokers was chiefly caused by smoking, smoking (just 9 cigarettes per day among men and 6 per day among women at

recruitment) caused 10% of deaths in men and 4% of deaths in women at ages 35-89. Among those without chronic disease who smoked ≥10 cigarettes/day at recruitment more than one-third of the deaths were caused by smoking, which was observed at ages 35-69 and 70-89 years and among both men and women. Among both men and women who had ever regularly smoked, more than two-fifths of deaths from respiratory cancers or COPD were caused by smoking.

Although occasional smokers had no detected excess risk of death compared with never smokers, occasional smokers' average long-term exposure was very low: at resurvey, 71% of those who smoked less than daily at recruitment had either quit or reported never regularly smoking, while 7% reported beginning to smoke daily (and whether this continued long-term is uncertain). These findings underscore the importance of population-specific measures of tobacco use, and particularly the importance of considering 'usual' smoking habits and distinguishing between 'current' and 'daily' smoking. Furthermore, in those with major chronic disease at recruitment, daily smoking at recruitment was little associated with excess mortality compared with never smoking, in large part because of higher risk of death from causes unrelated to smoking.

Smoking cessation was common prior to recruitment, among both men and women, and was positively associated with age. Although smoking cessation information was not well characterised, the benefits of quitting were clear: the excess risk of death associated with smoking was substantially lower among those who had quit >3 years prior to recruitment than among continuing daily smokers for several causes of death. Furthermore, one-fifth of ex-smokers at recruitment reported at resurvey that they had quit smoking due to illness, while only 39% of 'light' smokers (<10 cigarettes/day) and 68% of 'heavy' smokers (≥10 cigarettes/day) continued to smoke daily at resurvey.

Therefore, both the benefits of quitting smoking while healthy, and the harms of continued daily smoking, are likely to have been underestimated.

4.7.2 Strengths and limitations

A major strength of this study is its size, which was especially important for capturing the hazards of smoking in this low-consumption population, in which smoking cessation and non-daily smoking were common. Alongside the Cuba Prospective Study (described in **Chapter 3**), this is the first large prospective study of the association between smoking and cause-specific mortality in a Latin American country, and among the first to demonstrate reliably the hazards of smoking at relatively low levels of consumption. This finding raised the hypothesis that low levels of consumption may lead to comparatively lower mortality risk associated with daily smoking (but still significant excess risk) in the Mexican-American population in the United States, and in other low-consumption populations globally. However, this has not yet been examined in sufficiently large studies, and is examined in **Chapter 5**.

There were several other strengths of this study. The long follow-up period allowed for sufficient accrual of events to examine reliably cause-specific mortality associated with daily smoking, despite the relatively few daily smokers in this population. The medically certified deaths in this population, for which the underlying death certificates were further reviewed by study clinicians, enhanced the accuracy of the causes of death in this study. The questionnaire collected information on a range of smoking behaviours, notably separating out daily and non-daily smokers, which was a critical methodological detail for this population in which non-daily smoking was common. Finally, the resurvey conducted towards the end of the follow-up period allowed for changes in smoking behaviour to be characterised, resulting in a clearer interpretation of the long-term smoking behaviours and associated risks in this study.

This study also had several limitations. Efforts to avoid paper-based data collection, which was a novel endeavour in a low- or middle-income setting at the time, resulted in some continuous variables being categorised somewhat crudely. This limited the extent to which some of the information collected was informative, most notably for duration of smoking cessation. However, this handheld tablet-based data collection also allowed for automatic data checks, and reduced inaccuracies in other aspects of the baseline survey. Smoking information was self-reported and collected at a single point in time, which is likely to have resulted in some exposure misclassification. This could partially be addressed by information collected in the resurvey of a subset of surviving participants. However, the resurvey was conducted towards the end of the follow-up period, and therefore may not represent 'average' behaviour changes over the follow-up period (ie, this information could lead to over-correction for 'usual' habits). Furthermore, because blood samples were stored for 98% of participants, exposure misclassification could, in the future, be further explored through the use of blood-based biomarkers such as cotinine (a metabolite of nicotine). These blood samples could also provide additional information on other possible mediators of the smoking-mortality relationship (and other major risk factors or clusters of risk factors), such as information on lipids or genetic architecture, to further refine risk mediation and stratification in this population. However, these blood-based data had not yet been produced when analyses for this thesis were conducted.

Ascertainment of death over the follow-up period relied on probabilistic matching, as national identification numbers were not in widespread use at the time of the baseline survey. Although there is some uncertainty in the matching process, and it is therefore possible that there was some follow-up bias, this potential bias was reduced by subsequent visits to the homes of participants who appeared in the death registry, and by excluding participants with uncertain vital status from analyses. A full resurvey of all

households to validate mortality information, and collect information on non-fatal events, is currently being piloted, which will further address this study limitation.

The study was not designed to be nationally-representative, and collected information on a settled population from two contiguous districts in Mexico City. Although the participation rate in the study was high, which should make the study reasonably representative of adults in these two districts, the generalisability of findings to the rest of Mexico remains unclear. Although no significant heterogeneity was observed in the association between smoking and mortality according to age, sex, education, or alcohol consumption, significant heterogeneity was observed by district (with higher RRs in the lower-income district of Iztapalapa). This observation may indicate that the relevance of smoking to mortality could vary according to socioeconomic status (or other uncaptured factors relating to differences between these districts).

Finally, diabetes was much more common than in most high-income populations, and importantly was poorly controlled in this population at recruitment.[103,104] Prospective results have shown that diabetes is a cause of at least one third of all deaths at ages 35-74 years of age in this population, primarily from renal disease, vascular disease, infection, and acute diabetic crises.[103,104] Because these diseases are little associated with smoking (with the exception of vascular disease), and poor diabetic control and resulting health outcomes may have materially impacted on other health behaviours (including smoking), analyses were performed separately among those with and without chronic disease (primarily diabetes) at recruitment. Therefore, while the impact of smoking on health among those without chronic disease at recruitment is clear, it is difficult to interpret the impact of true long-term smoking behaviours among those with chronic disease, and particularly diabetes, in this population. As 42% of deaths among daily or never smokers (the main comparator groups of interest) occurred in participants

with chronic disease at recruitment, the main results are relevant to only a portion of the study population. However, results are also presented for those with chronic disease at recruitment, and indicate that the mortality rate in daily smokers was higher than that of never smokers for several causes of death. Furthermore, the PAFs associated with smoking in this population incorporate the RRs associated with ever smoking among both those with and without chronic disease at recruitment, taking into account the high competing risk from diabetes in this population. Although there is presently limited statistical power to further stratify the relevance of smoking to mortality among those with different diseases at recruitment, particularly screen-detected diabetes, known diabetes, and other major chronic diseases separately (shown in Table 4.6), further accrual of events in this study may provide sufficient power to clarify these associations in the future.

4.7.3 Findings in the context of prior evidence

The prevalence of ever smoking in the Mexico City Prospective Study ('the Mexican cohort') was somewhat higher than in the contemporary National Health Survey 2000, a nationally-representative health survey conducted in Mexico[105]: approximately 80% of men and 37% of women had smoked, compared with 60% of men and 19% of women, respectively. In both studies, cessation among ever-smokers was strongly correlated with age, with the majority of ever-smokers having quit by the age of ~50. However, the National Health Survey 2000 did not distinguish between current daily and non-daily smokers, nor did it report information about amount smoked among daily smokers.

Mexico's first Global Adult Tobacco Survey (GATS), a nationally-representative survey of smoking habits, was conducted in 2008-2009, approximately 6 years after the average year of recruitment in the Mexico City Prospective Study. The GATS survey reported that 25% of men and 8% of women smoked, but that just over half of these

smokers smoked less than daily,[106] in comparison to 43% of men and 20% of women in the Mexican cohort, of whom approximately one-third smoked less than daily. However, the number of cigarettes smoked per smoker in the 2009 GATS was similar to, if slightly higher than, in the Mexican cohort: among daily smokers, an average of 10/day among men and 8/day among women, compared with 9/day among men and 6/day among women in the Mexican cohort. Age at smoking initiation was also generally similar. Therefore, the risks associated with daily smoking compared with never smoking are likely to be broadly generalisable to the Mexico City, and possibly the general Mexican population (particularly the urban, middle-income Mexican population). However, because smoking was more common in the Mexican cohort than in the general population, the overall PAFs associated with ever smoking may be overestimated depending on the 'usual' smoking behaviours in the general population. Furthermore, as this study sampled only residents of the urban (and generally middle-income) districts of Coyoacán and Iztapalapa, PAFs may be less generalisable in more rural areas or those of mostly high- or low-income backgrounds. However, given the similarity of estimated PAFs of this study population and those indirectly estimated from other sources (described below), these findings should be at least broadly generalisable to the general Mexican population.

The relative risks associated with daily smoking in Mexico have not previously been directly estimated in a large prospective study. Given that smoking behaviours in the Mexican cohort and the general population are broadly similar, the excess risk among smokers may be reasonably generalisable, despite the limitations outlined above. The generalisability of these results may eventually be confirmed or clarified in the Mexican Teachers Cohort, a prospective cohort study of 117 475 teachers across Mexico (98% women) recruited in 2006-2010.[107] However, given that the Mexican Teachers Cohort is very young (mean age 44 at recruitment) and the follow-up period to date is

comparatively short, it will be several more years before the cohort can be used to reliably estimate the hazards of smoking and benefits of quitting in another large study in the Mexican population (especially considering the relatively few smokers in this predominantly female cohort).

The risks associated with daily smoking in this population are considerably lower than those in high-income populations, and are more similar to the risks in populations in which the tobacco epidemic is not yet mature.[8,9,35,62] However, this is likely to have been driven, in part, by a lower propensity to smoke heavily or to smoke continuously over the life course in this population. A small study comparing Mexican Americans (n=2629) and non-Hispanic Whites (n=1136) in the United States, recruited between 1979-1988, reported all-cause mortality RRs for current versus never smoking of 1.6 in Mexican Americans (95% CI 1.1-2.5) compared with 2.5 (1.4-4.7) in non-Hispanic Whites, controlling for age and sex.[83] The authors hypothesised this was driven by a lower average number of cigarettes smoked/day between the groups: 11.2 (SE 3.2) among Mexican Americans, and 20.2 (SE 4.5) among non-Hispanic Whites. While the study relied on self-reported smoking behaviours at baseline, the cigarettes/day is somewhat higher than in the Mexican cohort, and the excess risk associated with smoking approximately twice as large (though the difference was not statistically significantly different). Although national estimates in the United States have also shown considerably lower daily cigarette consumption among Mexican-American smokers than among non-Hispanic White smokers, the risks associated with these different behaviours have not yet been reliably assessed in a large prospective study (see **Chapter 5**).

The NIH-AARP Study, a large prospective study of older adults in the US (median age 70 at recruitment) conducted in 1995-1996, examined ‘low-intensity’ versus ‘high-intensity’ smoking behaviours in 290 215 participants, including 1493 of whom reported

consistently smoking 1-10 cigarettes/day.[108] The all-cause RR associated with consistently smoking 1-10 cigarettes/day compared with never smoking was 1.87 (1.64-2.13), adjusting for sex, race/ethnicity, education, alcohol, and physical activity, using baseline age as the underlying time variable. While again higher than the risks associated with smoking <10 cigarettes/day in the Mexican cohort, the NIH-AARP study relied on self-reported smoking information at a single point in time using recall about smoking habits across 9 prior periods over the life course, which may be prone to substantial recall bias in this older population.

Although the observed mortality RRs comparing daily and never smokers are lower than would otherwise be expected in a population in which the smoking epidemic is mature, the Mexican population appears to be unusual due to the combination of several factors. First, Mexico has experienced large reductions in national smoking prevalence over the past few decades,[109] which is likely driven in large part by smoking cessation, which would lead to underestimation of the hazards of smoking if not adequately taken into consideration. Second, both national surveys and the Mexican cohort indicate that a high proportion of 'current' smokers in Mexico smoke relatively infrequently, and it may be more common than in other populations to switch between daily and non-daily smoking over time, again potentially leading to underestimation of the long-term risks of persistent daily smoking. Third, both national surveys and the Mexican cohort report low smoking intensity among those who do smoke daily, which may also lead to greater variability in smoking behaviours (including intermittent smoking cessation or non-daily smoking). Finally, given that the baseline survey was conducted shortly before the introduction of the FCTC, which substantially strengthened tobacco control in Mexico,[110] reductions in smoking behaviour (and subsequent mortality risk) may have been especially pronounced during this period, which would further contribute to under-estimation of the hazards of persistent smoking in this population.

Overall, in the Mexico City Prospective Study, smoking caused an estimated 10% of male and 4% of female adult deaths. Using indirect methods, the GBD Study reported that an estimated 15% of male deaths from non-communicable diseases and 6% of female deaths from non-communicable diseases in Mexico were caused by smoking. Conversely, Reynales-Shigematsu *et al.*[111] indirectly estimated that approximately 7% of male deaths and 2% of female deaths in ~2007 were caused by tobacco in Mexico nationally. For COPD mortality, smoking caused an estimated 47% of male deaths and 15% of female deaths in the Mexican cohort, compared with an estimated 56% of male deaths and 22% of female deaths in the GBD Study. Therefore, indirect estimates of tobacco-attributable mortality in Mexico are broadly similar to direct estimates reported here, though estimates are somewhat higher or lower depending on the methods and data used.

Although the excess mortality risk associated with daily smoking in Mexico is much lower than that of high-income countries, the excess risk observed is similar to those of other low- and middle-income countries. For example, results from the China Kadoorie Biobank of 500 000 Chinese adults recruited in 2004-2008 showed that, compared with never smokers, smoking <15 cigarettes/day (mean 8) among rural men was associated with an all-cause RR of 1.48 (1.39-1.59), lung cancer RR of 2.28 (1.90-2.74), and COPD RR of 2.94 (1.99-4.35). However, the age at smoking initiation among these Chinese men was 25 years among those born in ~1930 and 20 years among those born in ~1970, on average several years older than the age at which male smokers began in the Mexican cohort. These findings are discussed in further detail in the context of findings from **Chapter 3** and **Chapter 5**, as well as other populations globally, in **Chapter 6.**

4.8 Conclusion

Although four-in-five men and nearly two-in-five women in this population had ever regularly smoked, daily smoking at recruitment was relatively uncommon due to smoking cessation and non-daily smoking, and daily smokers smoked relatively few cigarettes/day. This study found that in Mexican adults without major chronic disease at recruitment, daily smoking was associated with 28% higher risk of death at ages 35-89 than never smoking (and 54% higher risk among those smoking ≥10 cigarettes/day). However, smoking even as few as 4 cigarettes/day on average was associated with significantly higher risk of death compared with never smoking, and smoking cessation was associated with significantly lower risk of death compared with continued daily smoking at recruitment. Furthermore, the risks associated with continued daily smoking are likely to have been underestimated due to the large proportion of smokers who had quit by resurvey, and therefore the comparative benefits of smoking cessation in this Mexican population are likely to have been even greater than those reported.

Table 4.1: ICD-10 codes for cause-specific mortality endpoints

Endpoint	ICD-10 codes
Respiratory cancers	C33X C340 C341 C342 C343, C348 C349 D380 D381 D382 D383 D384 D385, D386, C000, C001, C002, C003, C004, C005, C006, C008, C009, C01X, C020, C021, C022, C023, C024, C028, C029, C030, C031, C039, C040, C041, C048, C049 C050, C051, C052, C058, C059, C060, C061, C062, C068, C069, C07X, C080, C081, C088, C089, C090, C091, C098, C099, C100, C101, C102, C103, C104, C108, C109, C110, C111, C112, C113, C118, C119, C12X, C130, C131, C132, C138, C139, C140, C142, C148, C320, C321, C322, C323, C328, C329, D370
COPD	J40X, J410, J411, J418, J42X, J430, J431, J432, J438, J439, J440, J441, J448, J449
Gastrointestinal	A000, A001, A009, A010, A011, A012, A013, A014, A020, A021, A022, A028, A029, A030, A031, A032, A033, A038, A039, A040, A041, A042, A043, A044, A045, A046, A047, A048, A049, A050, A051, A052, A053, A054, A058, A059, A060, A061, A062, A063, A064, A065, A066, A067, A068, A069, A070, A071, A072, A073, A078, A079, A080, A081, A082, A083, A084, A085, A09, A090, A099, A09X, A183, A213, A222, A360, B054, B462, B573, B770, B780, B811, B813, B814, B818, B820, B829, D000, D001, D002, D010, D011, D012, D013, D014, D017, D019, D030, D040, D100, D101, D102, D103, D104, D105, D106, D107, D109, D110, D117, D119, D120, D121, D122, D123, D124, D125, D126, D127, D128, D129, D130, D131, D132, D133, D139, E340 E341, E743, P75X, P760, P761, P762, P768, P769, P77X, P780, P781, P782, P783, P788, P789, Q351, Q353, Q355, Q356, Q357, Q359, Q360, Q361, Q369, Q370, Q371, Q372, Q373, Q374, Q375 Q378, Q379, Q380, Q381, Q382, Q383, Q384, Q385, Q386, Q387, Q387, Q388, Q390, Q391, Q392, Q393, Q394, Q395, Q396, Q398, Q399, Q400, Q401, Q402, Q403, Q408, Q409, Q410, Q411, Q412, Q418, Q419, Q420, Q421, Q422, Q423, Q428, Q429, Q430, Q431, Q432, Q433, Q434, Q435, Q436, Q437, Q438, Q439, Q458, Q459, all ICD ‘K’ codes not listed under ‘Hepatobiliary’
IHD	I051, I059, I070, I071, I080, I081, I091, I099, I110, I119, I200, I210, I211, I219, I221, I251, I258, I259, I270, I272, I279, I301, I330, I340, I350, I351, I358, I38X, I420, I471, I472, I489, I48X, I490, I500, I501, I509, I518, I519, Q210, Q231, R570
Stroke	I600, I609, I610, I615, I619, I620, I629, I634, I639, I64X, I671, I672, I673, I674, I678, I679, I694, I698
Other vascular	All vascular diseases (ICD codes I00-I99) not listed under ‘IHD’ or ‘Stroke’
Other cancers	All cancers (ICD codes C00-D48) not listed under ‘Respiratory cancers’
Renal	E112, E142, I120, I129, I131, I132, N002, N039, N059, N12X, N151, N179, N180, N189, N19X, N200, N201, N289, N390
Hepatobiliary	B169, B171, B181, B182, B190, I850, I859, K701, K702, K703, K704, K709, K716, K720, K721, K729, K745, K746, K750, K754, K766, K767, K769, K800, K801, K803, K810, K811, K819, K821, K829, K830, K852, K858, K859, K85X, K861
Other respiratory	A162, A169, B206, J09, J157, J159, J180, J181, J189, J209, J22X, J459, J64X, J679, J680, J80X, J81X, J841, J849, J852, J980, J984, J985, J988, Q311
Acute diabetic	E110, E111, E140, E141
Other medical	All non-external causes of death not listed previously
External	T782, V093, V099, V299, V499, V580, V878, V892, V899, W018, W100, W108, W130, W134, W135, W138, W139, W179, W184, W188, W190, W314, W748, W784, W849, W878, X090, X094, X599, X650, X680, X700, X702, X740, X950, X954, X990, X994, Y044, Y200, Y240, Y249, Y338, Y340, Y344, Y349, Y579
Any medical	All causes of death other than ‘External’

Table 4.2: Baseline characteristics of the 158 688 participants aged 35-89 at recruitment, by age and sex

	Age 35-69			Age 70-89			
	Men	Women	All	Men	Women	All	Total
Participants	44 717	93 328	138 045	7574	13 069	20 643	158 688
Age, years	49 (10)	49 (10)	49 (10)	76 (5)	76 (5)	76 (5)	52 (13)
District							
Resident in Iztapalapa	56%	61%	59%	64%	64%	64%	60%
Resident in Coyoacan	44%	39%	41%	36%	36%	36%	40%
Education attained							
University	26%	13%	17%	8%	3%	4%	15%
High school	28%	26%	27%	10%	7%	8%	24%
Elementary school	40%	50%	47%	53%	48%	50%	47%
Other	6%	11%	10%	29%	42%	37%	13%
History of chronic disease†							
Diabetes	13%	12%	12%	21%	24%	23%	14%
Ischaemic heart disease	2%	1%	1%	5%	4%	4%	2%
Stroke	1%	1%	1%	3%	3%	3%	1%
Cancer	<0.5%	1%	1%	2%	2%	2%	1%
Other	1%	1%	1%	3%	2%	2%	1%
Any	16%	15%	16%	30%	32%	31%	18%

Mean (SD) or percent shown.
† Self-reported history of diabetes, ischemic heart disease, stroke, cancer (any site), chronic kidney disease, cirrhosis, or emphysema.

Table 4.3: Baseline characteristics of the 130 791 included participants aged 35-89 at recruitment, by sex and smoking status

	Men						Women					
	Never	**Ex >3 years**	**Ex 1-3 years**	**< Daily**	**Daily < 10**	**Daily 10+**	**Never**	**Ex >3 years**	**Ex 1-3 years**	**< Daily**	**Daily < 10**	**Daily 10+**
Participants	9013	12 387	2369	6232	8158	4811	54 686	12 312	2613	7092	8563	2555
Smoking behaviour												
Mean cigarettes/day ‡	-	8.7 (12.6)	8.3 (11.1)	0.2 (0.5)	4.1 (2.2)	16.3 (7.5)	-	4.0 (7.7)	4.4 (7.4)	0.1 (0.3)	3.7 (2.0)	14.5 (6.0)
Age started smoking (years)	-	17.6 (5.7)	18.0 (6.1)	18.9 (6.5)	18.2 (6.6)	16.8 (5.3)	-	21.8 (8.1)	22.9 (9.4)	23.8 (9.6)	21.6 (8.5)	19.8 (7.0)
Ever tried to quit	-	-	-	40%	43%	34%	-	-	-	42%	53%	43%
Inhale smoke	-	-	-	64%	76%	84%	-	-	-	52%	67%	75%
Others smoke inside home	16%	18%	22%	26%	34%	42%	30%	32%	38%	44%	55%	60%
Age, years	52 (13)	56 (14)	50 (12)	47 (10)	50 (12)	50 (11)	52 (13)	51 (13)	47 (11)	45 (9)	46 (10)	48 (10)
Resident in Iztapalapa	57%	57%	51%	59%	56%	53%	64%	59%	53%	60%	57%	48%
Education attained												
University	31%	23%	25%	26%	21%	24%	10%	17%	18%	16%	17%	22%
High school	24%	23%	28%	33%	30%	29%	22%	27%	32%	36%	34%	32%
Elementary school	36%	43%	38%	37%	41%	39%	51%	44%	41%	42%	43%	40%
Other	9%	11%	9%	5%	8%	7%	16%	12%	8%	6%	6%	6%
Drinking behaviour												
Drink at least weekly	48%	46%	47%	51%	48%	46%	51%	56%	59%	66%	61%	56%
Drink less than weekly	29%	41%	41%	35%	39%	41%	13%	22%	22%	16%	19%	22%
Former drinker	9%	8%	8%	11%	10%	9%	3%	5%	6%	8%	8%	8%
Never drinker	14%	5%	3%	2%	3%	4%	33%	17%	14%	10%	12%	14%
Anthropometry and blood pressure												
BMI kg/m²	28.0 (4.3)	28.2 (4.1)	28.5 (4.5)	28.6 (4.3)	27.6 (4.4)	26.9 (4.4)	29.6 (5.2)	29.8 (5.5)	29.6 (5.1)	29.7 (5.3)	28.9 (5.2)	27.9 (5.3)
Waist-to-hip ratio	0.95 (0.07)	0.96 (0.07)	0.95 (0.06)	0.95 (0.06)	0.95 (0.06)	0.95 (0.07)	0.88 (0.07)	0.87 (0.07)	0.87 (0.07)	0.87 (0.07)	0.86 (0.07)	0.86 (0.07)
Systolic BP (mmHg)	128 (16)	130 (16)	128 (15)	127 (14)	127 (15)	127 (15)	127 (17)	126 (16)	123 (16)	122 (15)	122 (15)	123 (16)
Diastolic BP (mmHg)	84 (10)	85 (10)	84 (10)	84 (10)	84 (10)	84 (10)	83 (10)	82 (10)	81 (10)	81 (10)	80 (10)	81 (10)

Mean (SD) or percent shown.
Analyses excluded data from participants with previously diagnosed chronic disease (chronic kidney disease, ischemic heart disease, stroke, cirrhosis, cancer, emphysema, or diabetes).
† Cigarettes per day among non-daily smokers determined by calculating number of cigarettes smoked per occasion multiplied by smoking frequency (e.g. daily, weekly) to determine average cigarettes smoked per day over a 1-month period.

Table 4.4: Smoking characteristics at resurvey according to baseline smoking status, among those without chronic disease at baseline, ages 35-89

	Smoking status at recruitment (1998-2004)				
	Never smoker	**Ex smoker**	**Smoke < daily**	**Smoke < 10 cigs/day**	**Smoke 10+ cigs/day daily**
Participants	4695	1779	898	1049	389
Smoking status at resurvey					
Never smoker	4295 (91%)	469 (26%)	172 (19%)	55 (5%)	8 (2%)
Ex smoker (not for illness)	295 (6%)	943 (53%)	407 (45%)	289 (28%)	52 (13%)
Ex smoker (for illness)	26 (1%)	221 (12%)	67 (7%)	147 (14%)	49 (13%)
Smoke < daily	35 (1%)	59 (3%)	158 (18%)	127 (12%)	14 (4%)
Smoke < 10 cigs/day	10 (<0.5%)	45 (3%)	66 (7%)	356 (34%)	109 (28%)
Smoke 10+ cigs/day	1 (<0.5%)	6 (<0.5%)	3 (<0.5%)	57 (5%)	155 (40%)
Incomplete info $	33 (1%)	36 (2%)	25 (3%)	18 (2%)	2 (1%)
Average daily cigs					
At baseline	-	5.5 (9.7)	0.1 (0.3)	3.8 (2.1)	15.4 (7.2)
At resurvey, per smoker †	1.8 (4.4)	2.7 (3.9)	1.3 (1.9)	4.0 (3.8)	10.5 (6.9)
At resurvey, all participants ‡	0.0 (0.5)	0.2 (1.2)	0.3 (1.1)	2.1 (3.4)	7.5 (7.5)
Inhale smoke					
At baseline	-	-	517 (58%)	760 (72%)	316 (81%)
At resurvey	27 (1%)	81 (5%)	160 (18%)	394 (38%)	205 (53%)
Ever tried to quit §					
At baseline	-	-	329 (37%)	483 (46%)	150 (39%)
At resurvey	30 (1%)	84 (5%)	168 (19%)	399 (38%)	210 (54%)

N (%) or mean (SD) shown.
$ Where information given is insufficient to categorise exposure, e.g. when participant reports ex-smoking but not years since quitting, or current smoking but not frequency of smoking
† Among those self-reporting current smoking at resurvey
‡ Among all participants, i.e. those reporting never or ex smoking assigned zero cigarettes/day
§ Only current smokers asked about quit behaviour at baseline.

Table 4.5: Cause-specific mortality rate ratios (RRs) and population attributable fractions (PAFs) for ever versus never smoking at ages 35-89, by sex

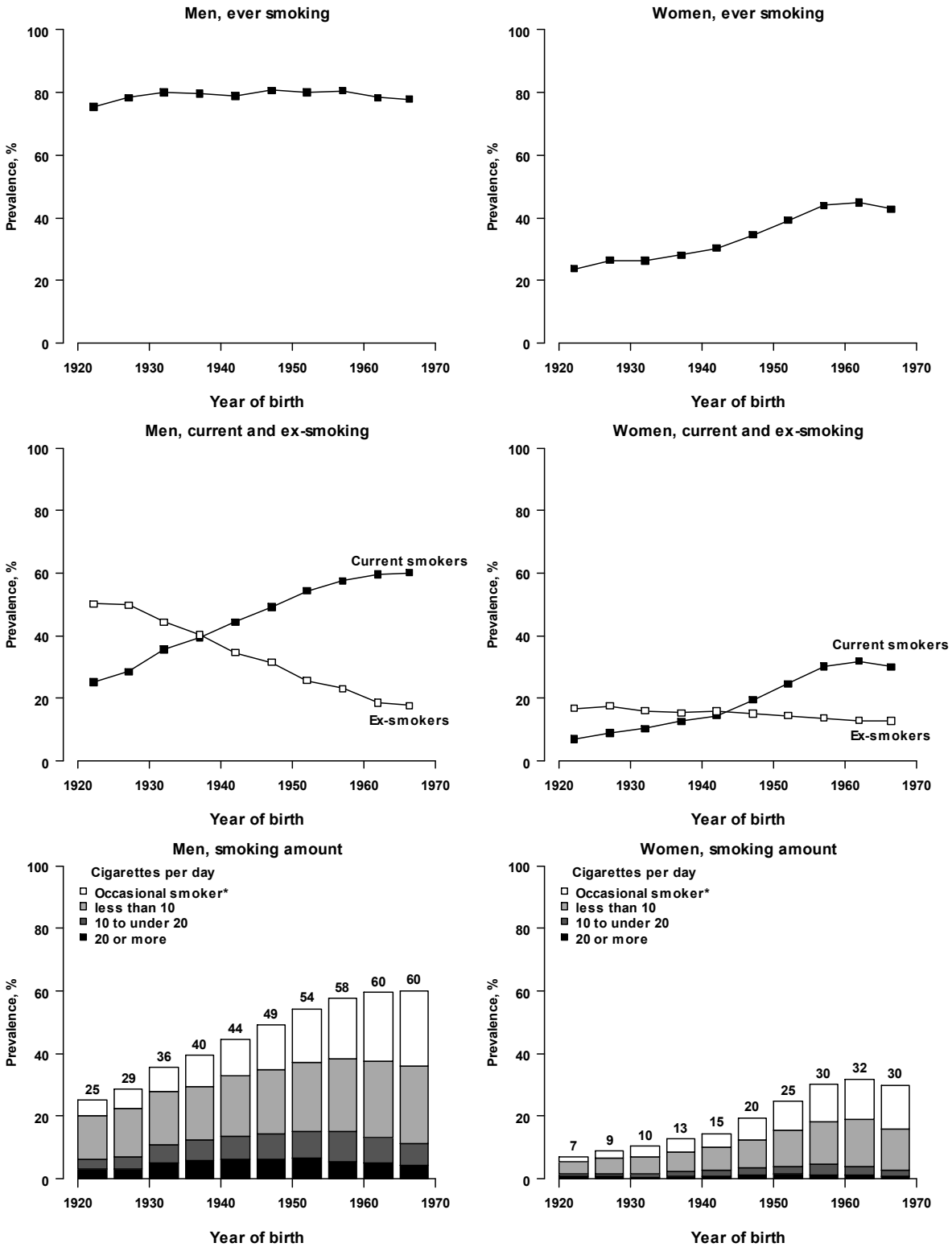
	Male			Female		
Cause of death	No. of deaths smokers/non-smokers	RR (95% CI)	PAF* (%)	No. of deaths smokers/non-smokers	RR (95% CI)	PAF (%)
Respiratory cancers	157/21	1.97 (1.24-3.12)	43	61/74	1.70 (1.20-2.43)	19
COPD	367/47	2.11 (1.55-2.87)	47	157/266	1.72 (1.40-2.10)	15
Gastrointestinal	315/76	1.04 (0.81-1.34)	3	203/380	1.32 (1.10-1.57)	8
IHD	1649/382	1.15 (1.02-1.28)	10	754/1633	1.16 (1.07-1.27)	4
Stroke	477/92	1.38 (1.10-1.73)	23	282/607	1.14 (0.99-1.32)	4
Other vascular	183/31	1.53 (1.04-2.26)	30	76/172	1.12 (0.85-1.48)	3
Other cancers	752/212	0.93 (0.80-1.08)	-6	541/1117	0.95 (0.86-1.06)	-2
Renal	928/226	1.04 (0.89-1.20)	3	505/1056	1.12 (1.00-1.25)	3
Hepatobiliary	570/132	1.00 (0.82-1.21)	0	218/392	1.18 (1.00-1.40)	6
Other respiratory	552/126	1.16 (0.96-1.42)	12	251/611	0.99 (0.85-1.15)	0
Acute diabetic	252/54	1.16 (0.86-1.56)	11	134/271	1.23 (0.99-1.53)	6
Other medical	374/90	1.11 (0.88-1.40)	8	203/418	1.15 (0.96-1.37)	4
Any medical	6576/1490	1.14 (1.08-1.21)	10	3385/7002	1.14 (1.09-1.19)	4

* Example calculation: for respiratory cancer in men, (0.97 / 1.97) x 157 ÷ (157 + 21) = 0.43

Table 4.6: All-cause mortality rate ratios (RRs) and population attributable fractions (PAFs) for ever versus never smoking at ages 35-89, by chronic disease status at recruitment

Chronic disease status at recruitment	Deaths (smoker/non-smoker)	Never	Ex	Daily	Ever	PAF (ever vs. never)
No illness	5213/4264	1.00 (0.96-1.04)	1.12 (1.08-1.16)	1.30 (1.24-1.36)	1.15 (1.10-1.21)	7.2%
Screen-detected diabetes	693/603	1.00 (0.90-1.11)	1.06 (0.96-1.18)	1.27 (1.11-1.44)	1.12 (0.98-1.28)	5.7%
Known diabetes	3092/2938	1.00 (0.95-1.05)	1.07 (1.02-1.12)	1.08 (1.01-1.16)	1.07 (1.00-1.13)	3.4%
Known other chronic disease	876/563	1.00 (0.90-1.11)	1.16 (1.07-1.25)	0.94 (0.81-1.08)	1.07 (0.95-1.21)	4.0%
Known diabetes and other chronic disease	384/465	1.00 (0.88-1.14)	1.07 (0.97-1.19)	1.00 (0.81-1.25)	1.01 (0.86-1.19)	0.4%

Figure 4.1: Baseline smoking prevalence in the Mexico City Prospective Study



Number above bar in plots 5 and 6 is percent current smokers
*: Non-daily smokers

Note: Figure excludes those born <1920 for clarity (2% of cohort)

Figure 4.2: Mean cigarettes/day at recruitment and age started smoking among the 28 122 daily smokers

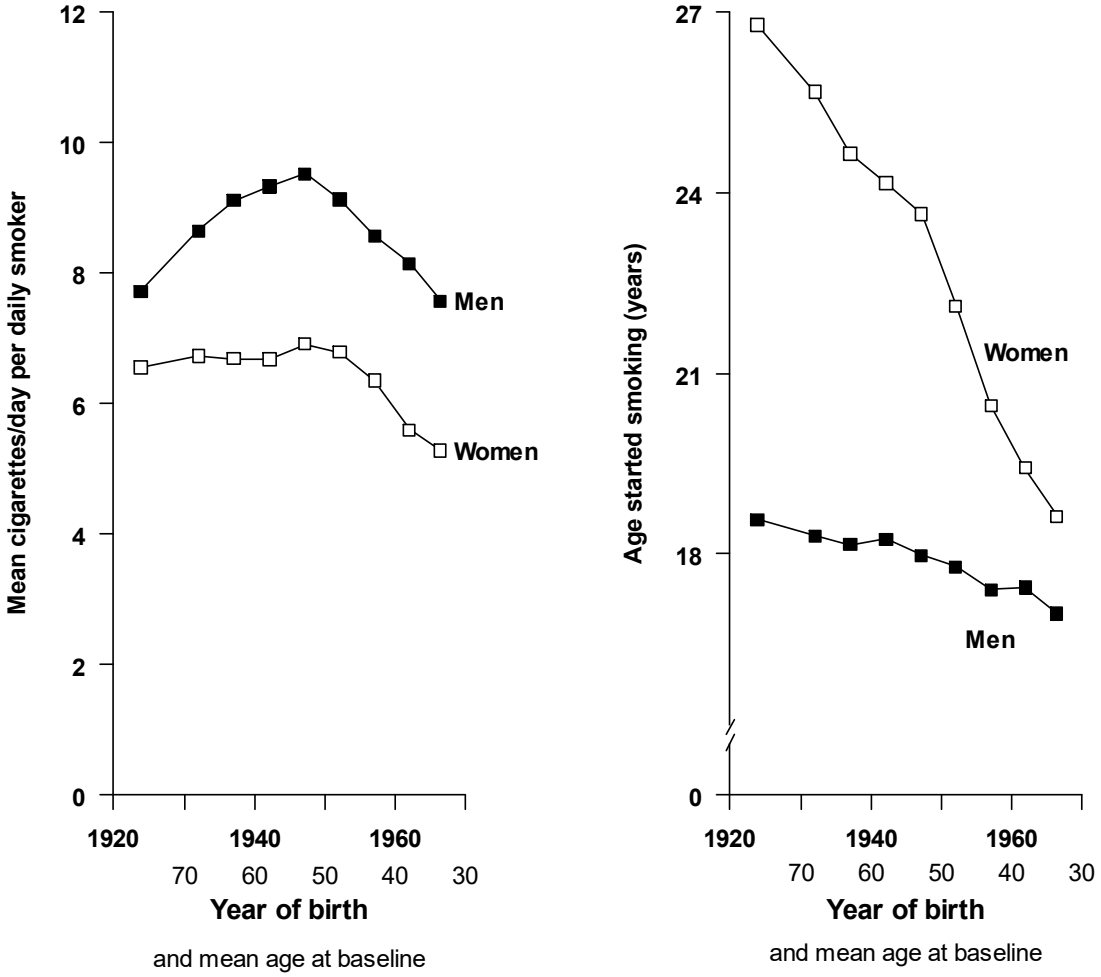
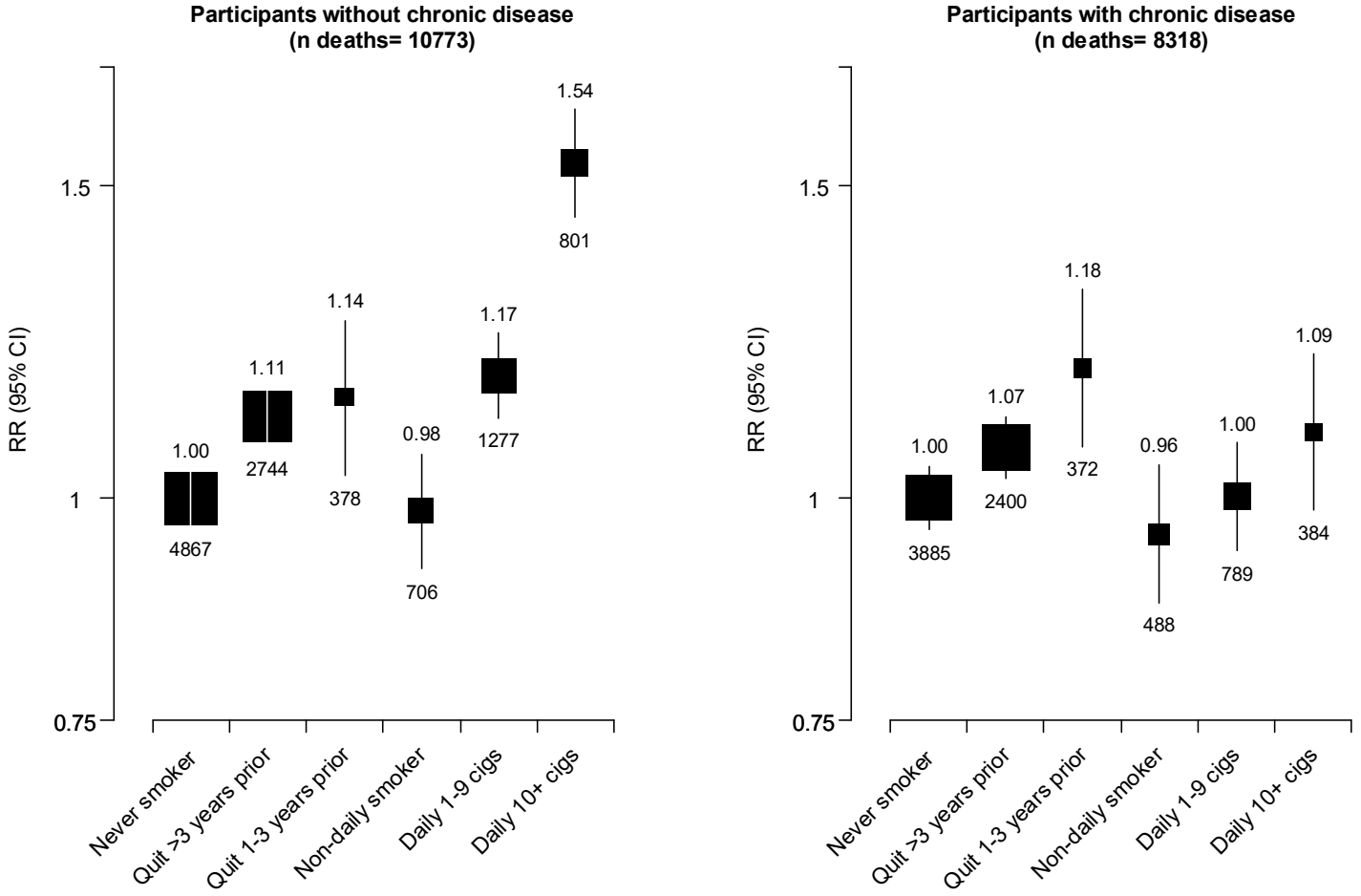


Figure 4.3: Association between smoking habits and all-cause mortality among those with vs without chronic disease at recruitment by smoking status, age 35-89



Chronic disease included those with ischaemic heart disease, stroke, cancer, cirrhosis, chronic kidney disease, or diabetes.

Figure 4.4: All-cause mortality rate ratio at ages 35-69 and 70-89, daily smoker vs never smoker, by usual amount smoked

Participants with chronic disease at recruitment were excluded from these analyses.

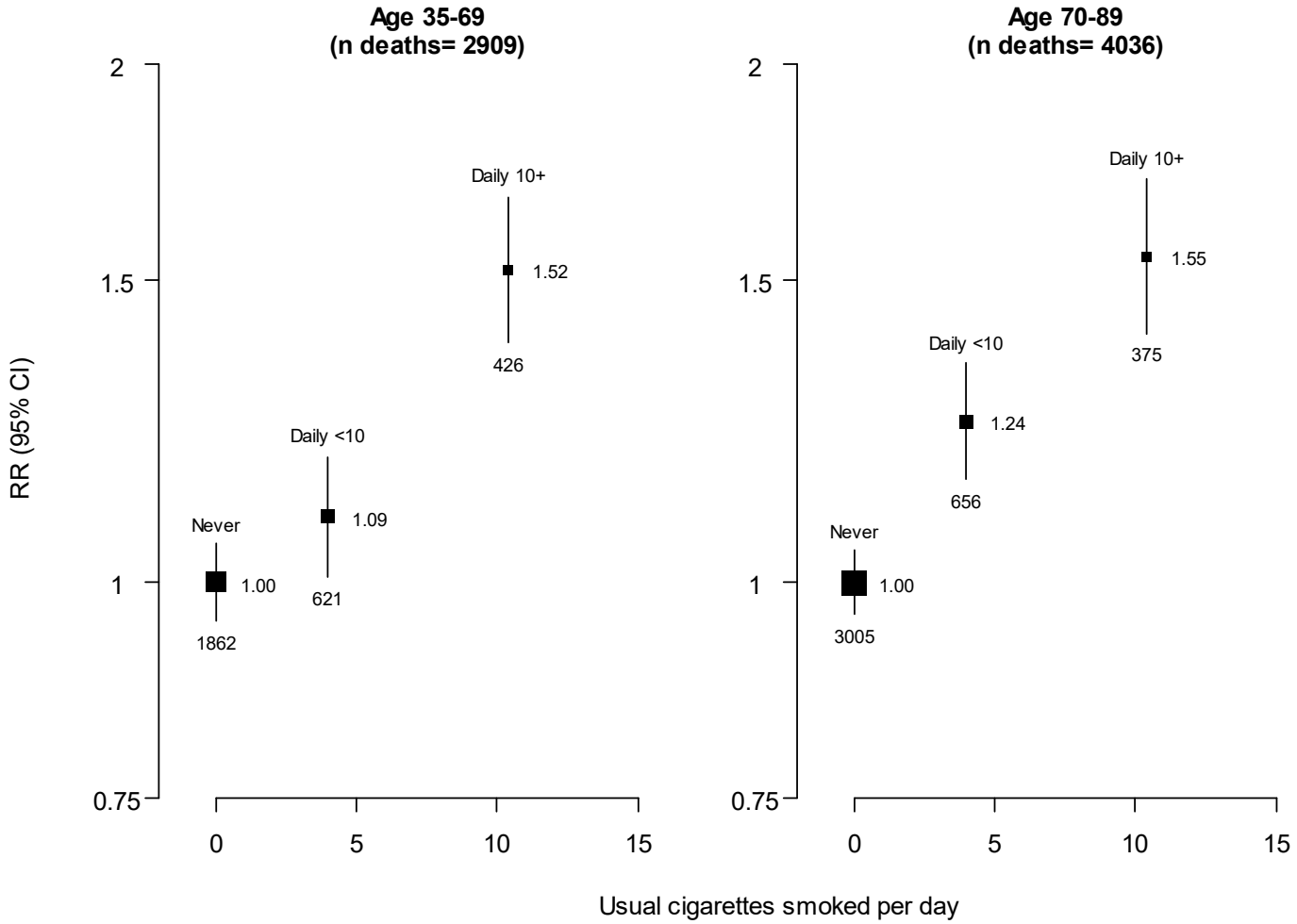


Figure 4.5: All-cause mortality rate ratio at ages 35-69 and 70-89, daily smoker vs never smoker, by age started smoking

Participants with chronic disease at recruitment were excluded from these analyses.

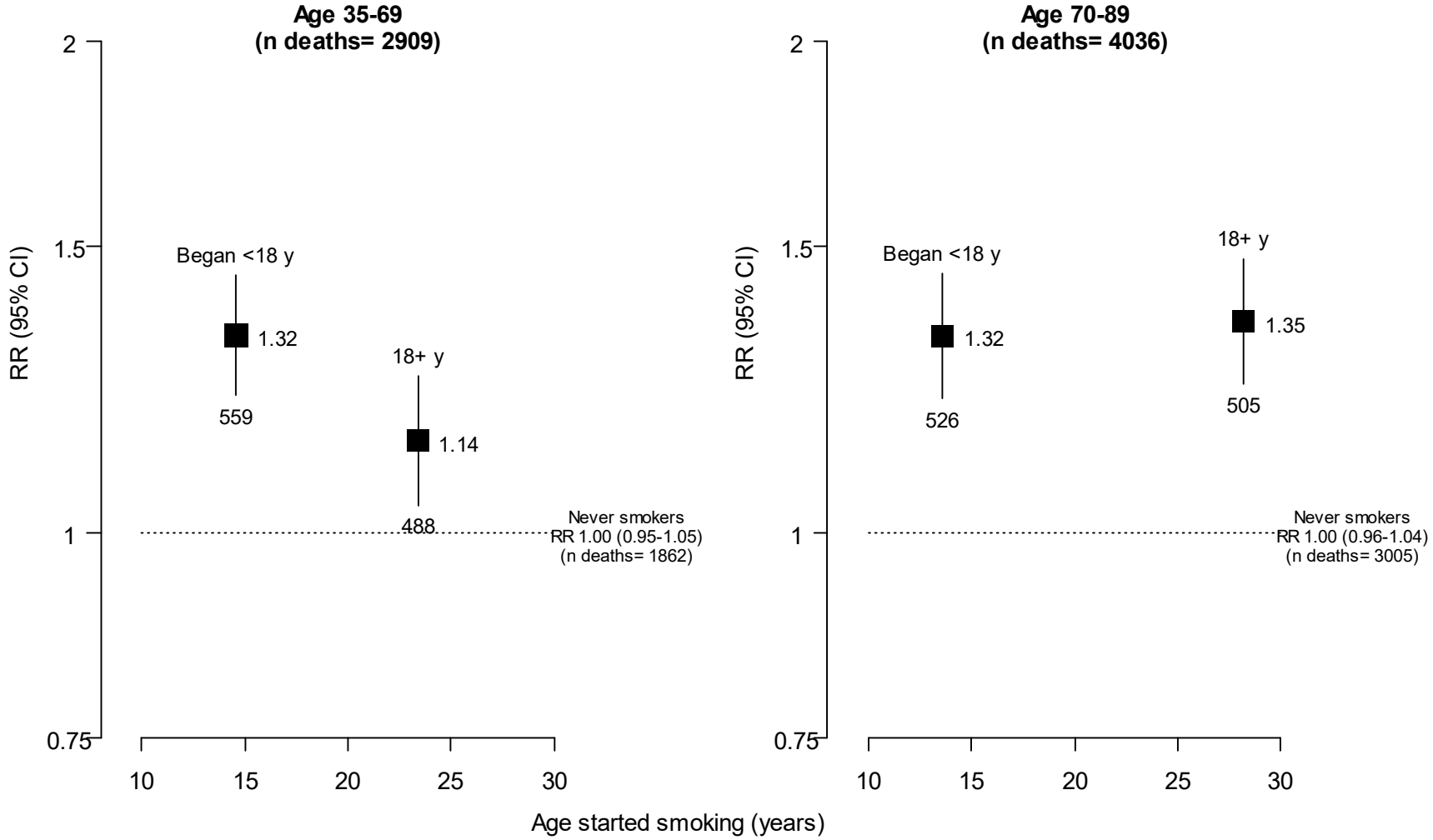
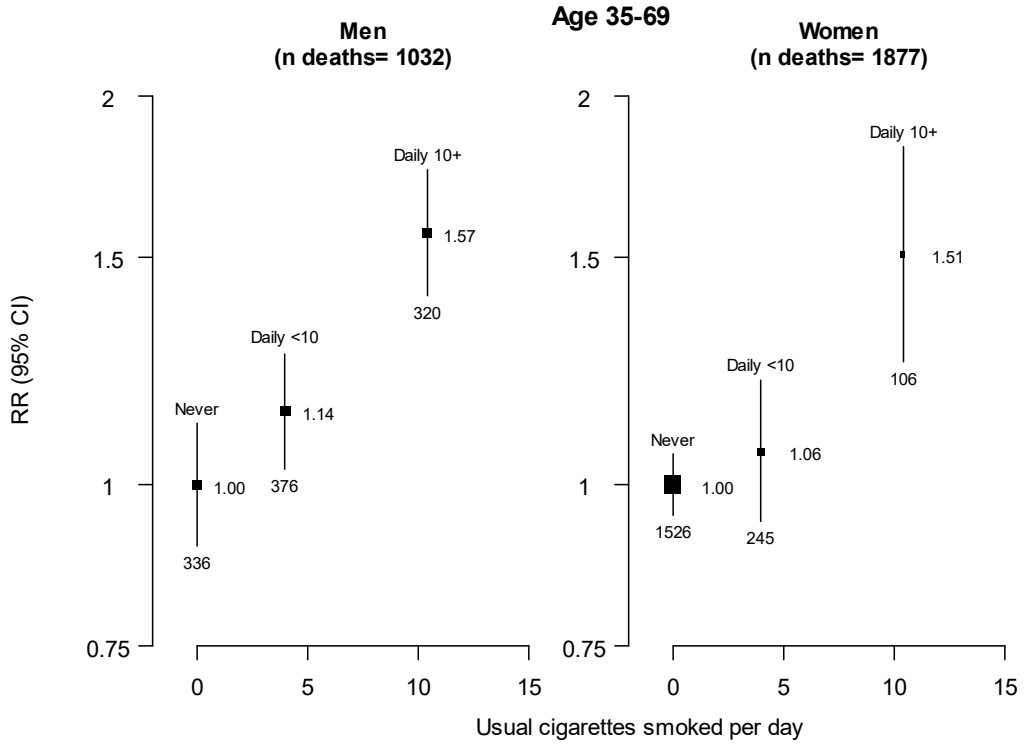


Figure 4.6: All-cause mortality rate ratio at ages 35-69 and 70-89, daily smoker vs never smoker, by usual amount smoked and sex

Participants with chronic disease at recruitment were excluded from these analyses.



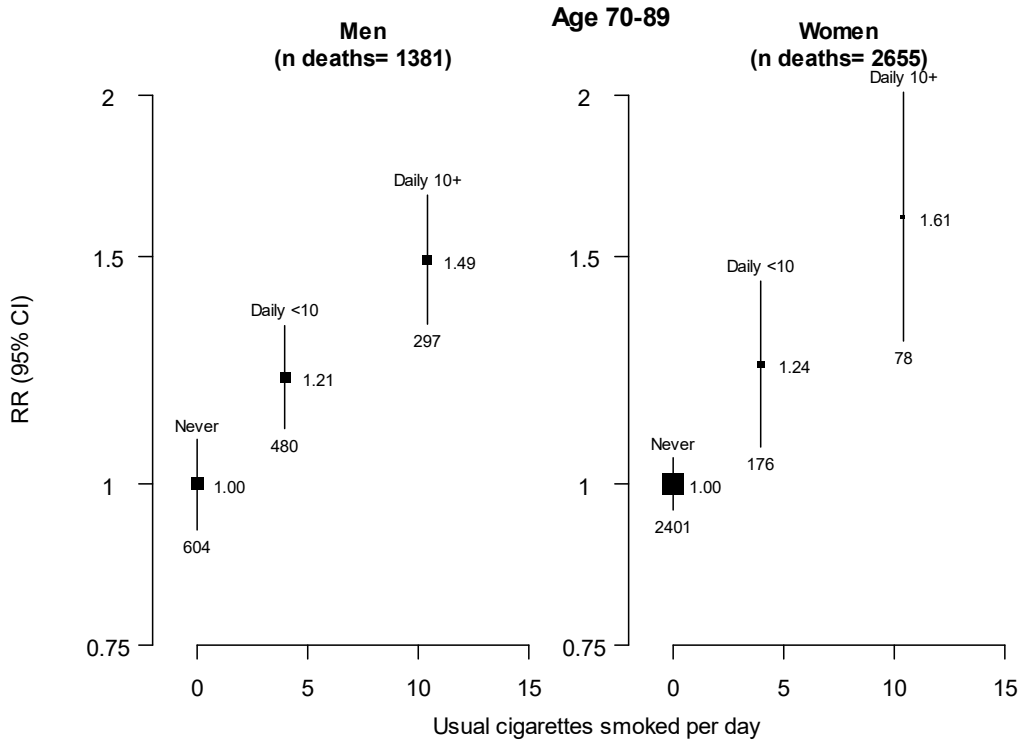
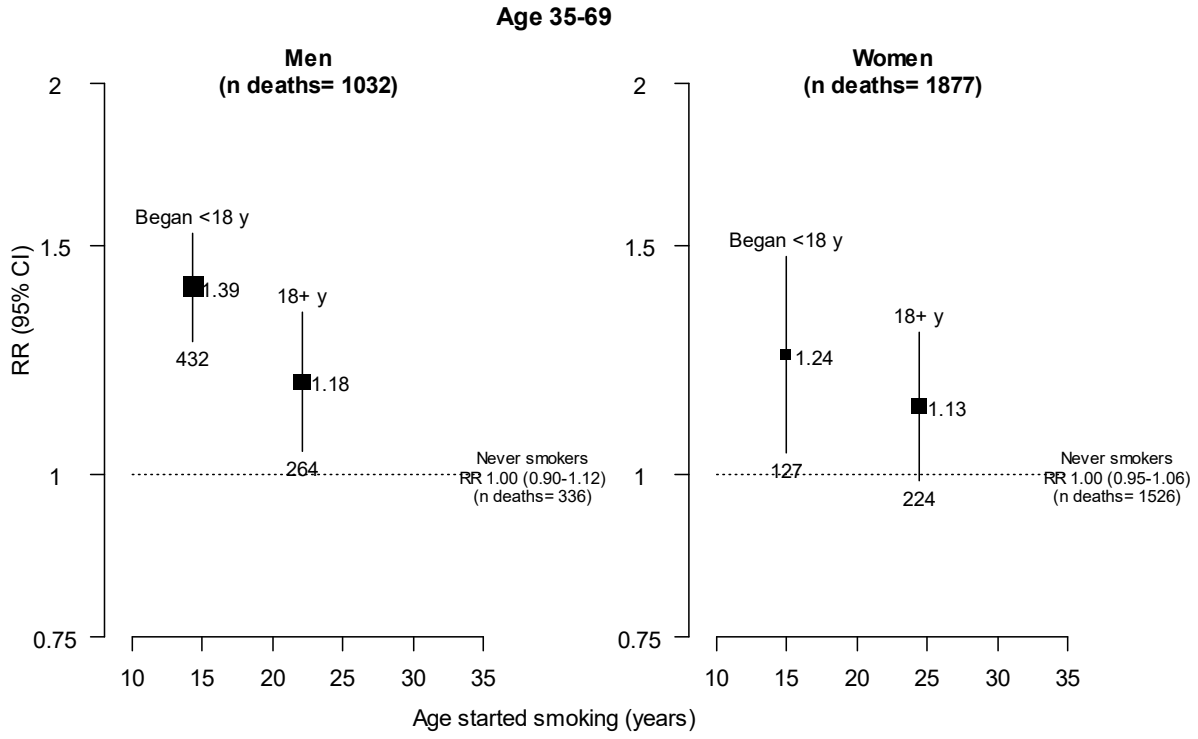


Figure 4.7: All-cause mortality rate ratio at ages 35-69 and 70-89, daily smoker vs never smoker, by age started smoking and sex

Participants with chronic disease at recruitment were excluded from these analyses.



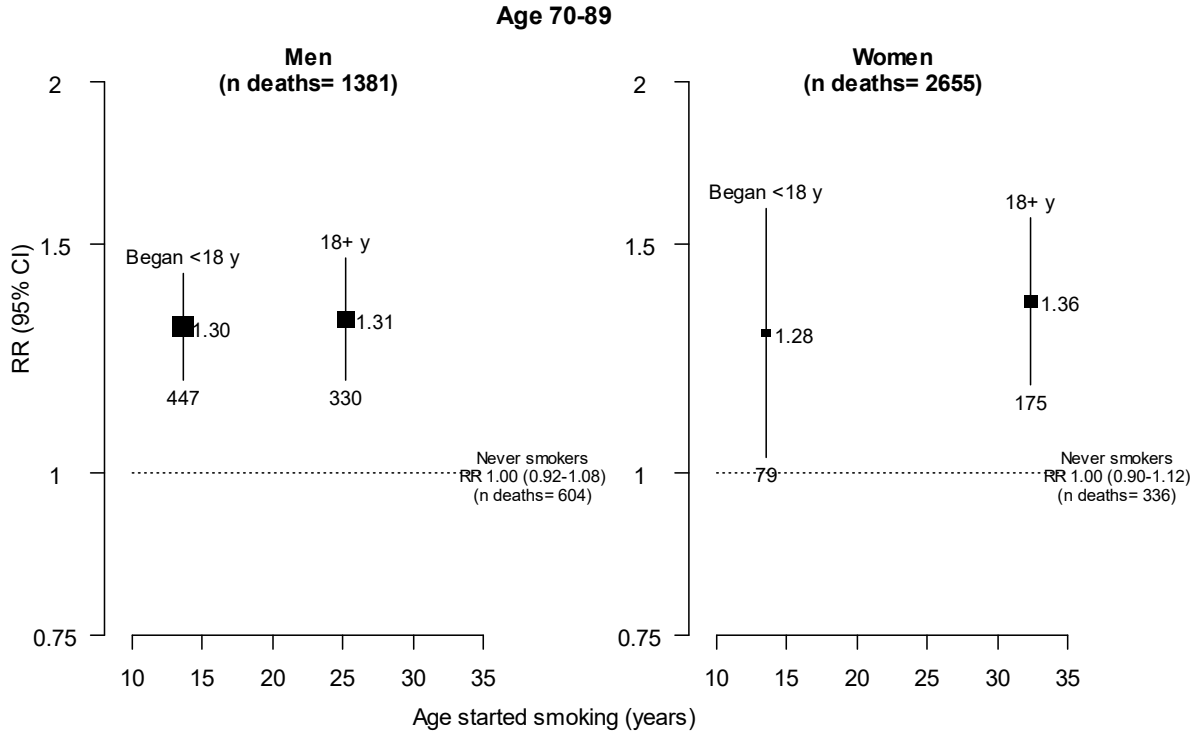
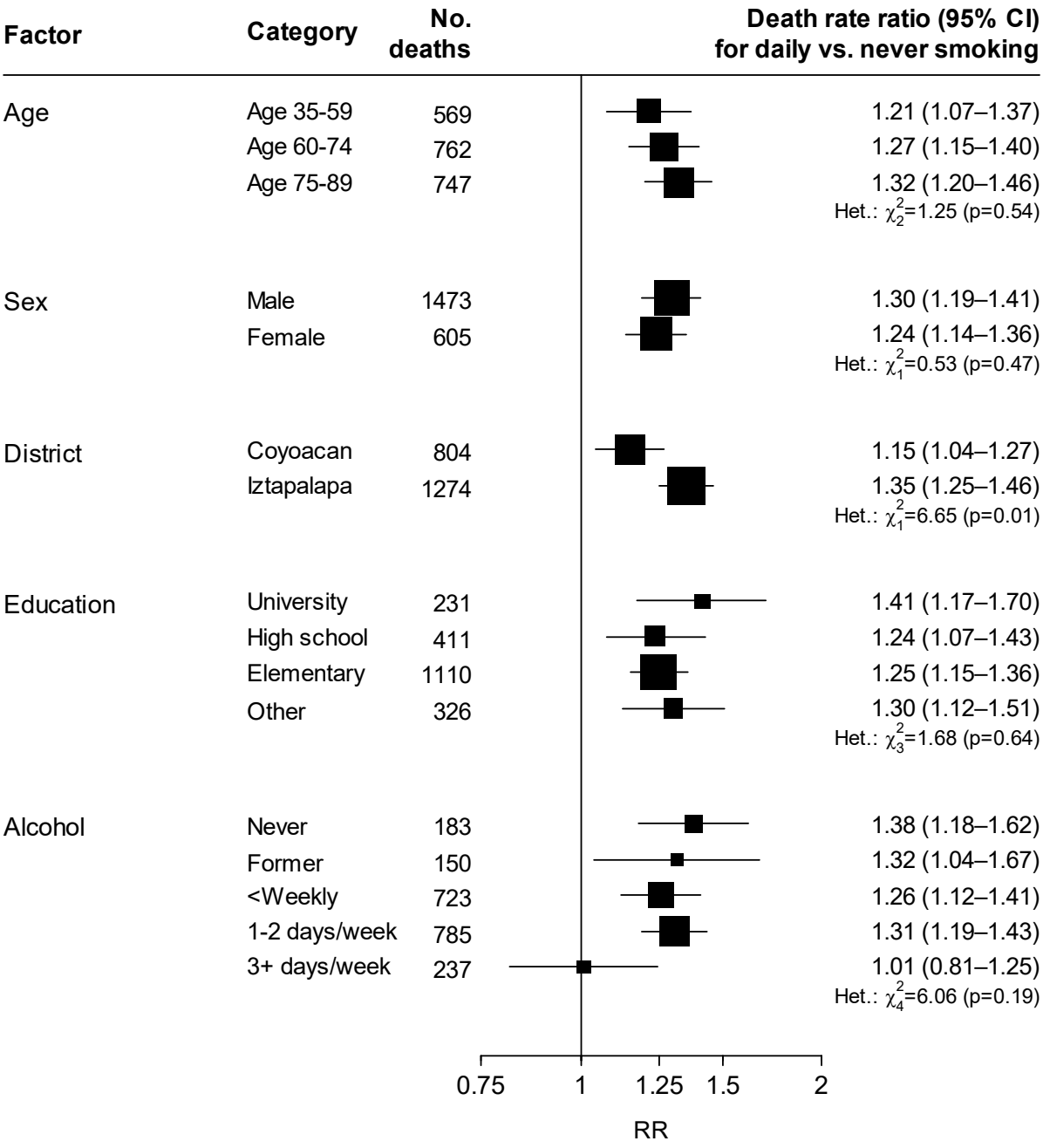


Figure 4.8: All-cause mortality rate ratio at ages 35-89, daily smoker vs never smoker, by levels of factors adjusted for in main analyses

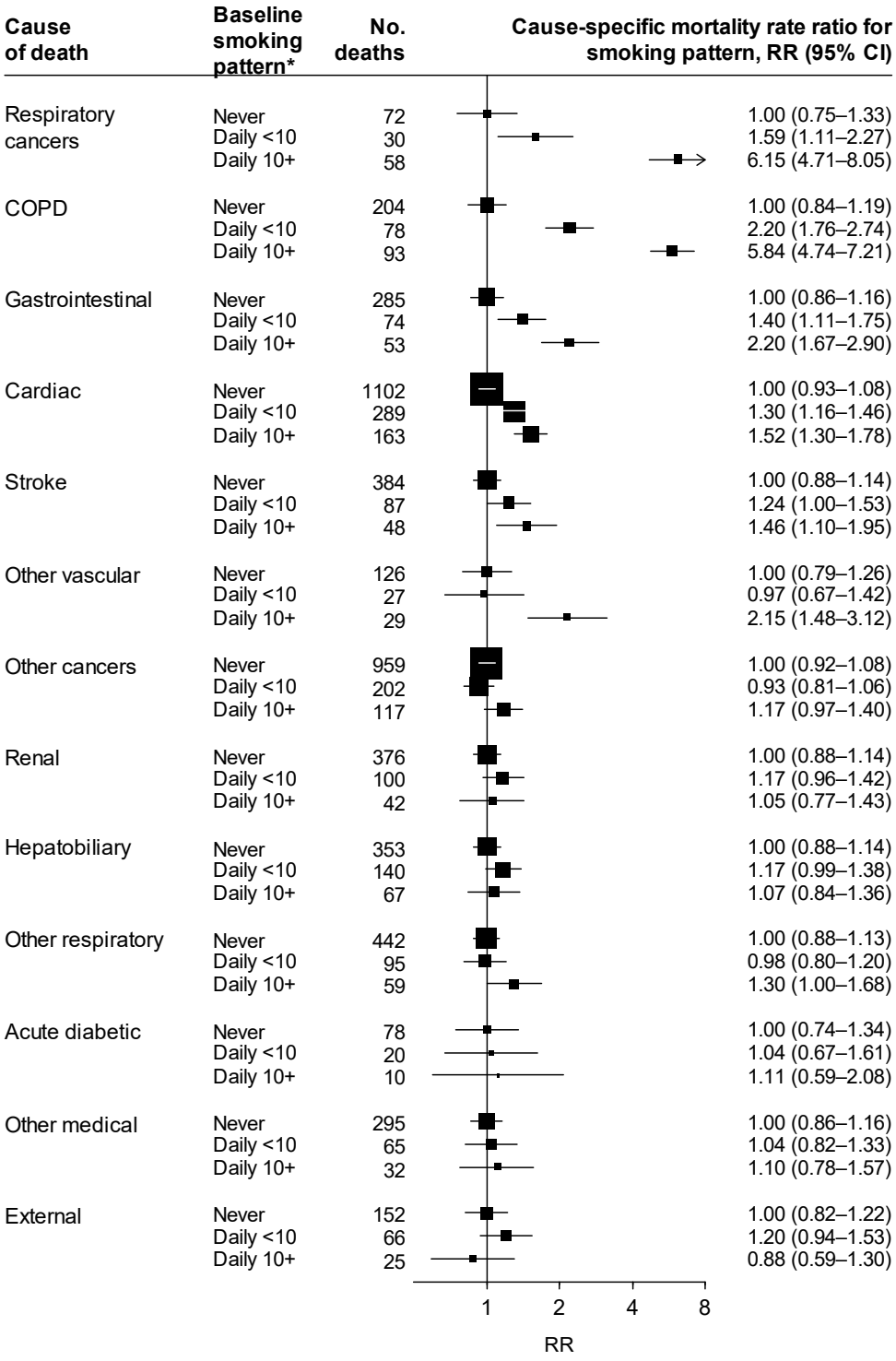
Participants with chronic disease at recruitment were excluded from these analyses.



RRs estimated separately by levels of each confounder shown, mutually adjusted for each other.
Number of deaths among daily smokers shown.

Figure 4.9: Cause-specific mortality at ages 35-89, daily smoker vs never smoker, by amount smoked

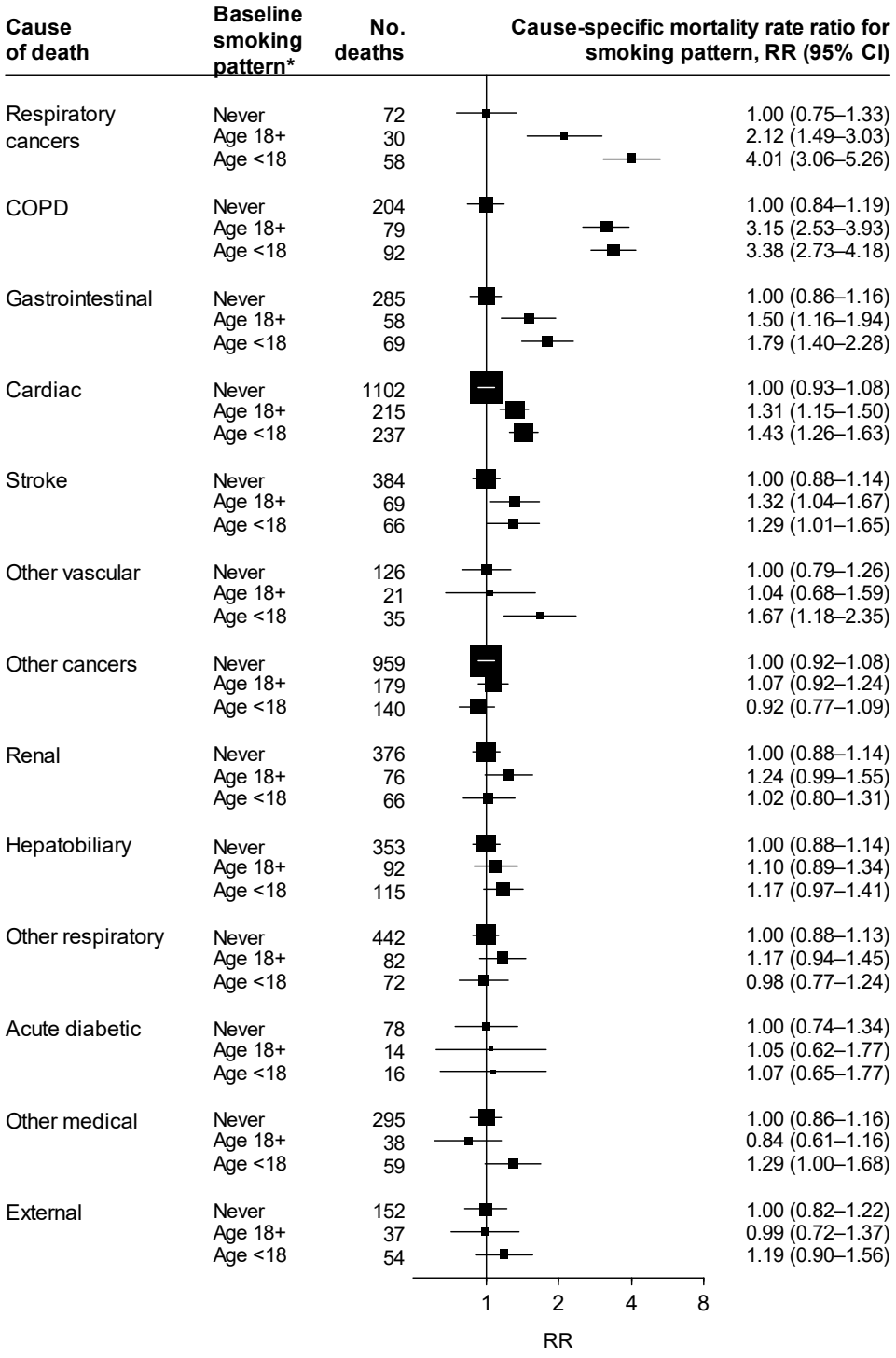
Participants with chronic disease at recruitment were excluded from these analyses.



*RRs associated with smoking 0, 4, and 16 cigarettes per day on average at recruitment, respectively.

Figure 4.10: Cause-specific mortality at ages 35-89, daily smoker vs never smoker, by age started smoking

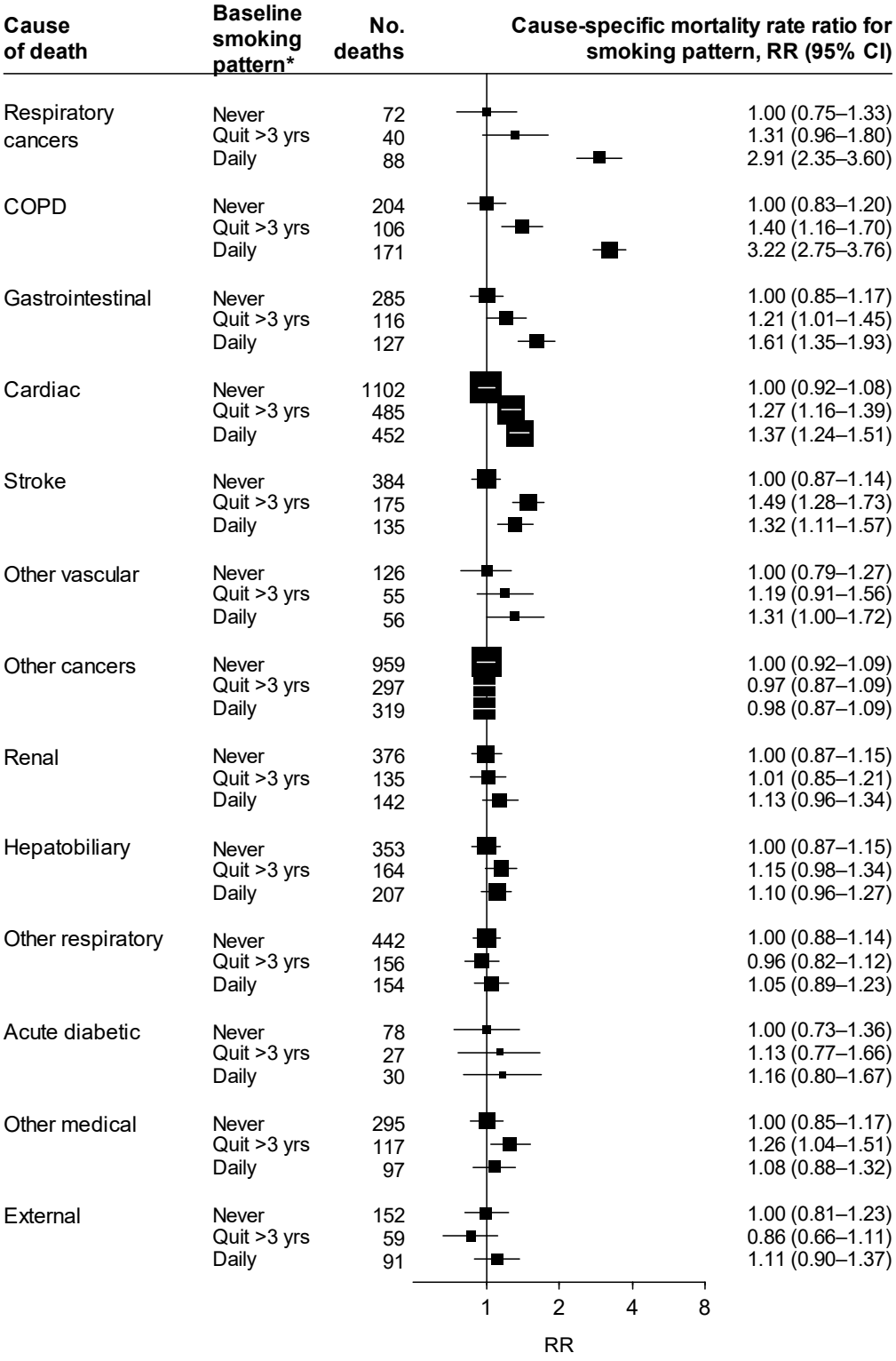
Participants with chronic disease at recruitment were excluded from these analyses.



*RRs associated with starting at age 24 or 14 years on average, respectively.

Figure 4.11: Cause-specific mortality at ages 35-89, never vs 'long-term' ex-daily and current daily smokers

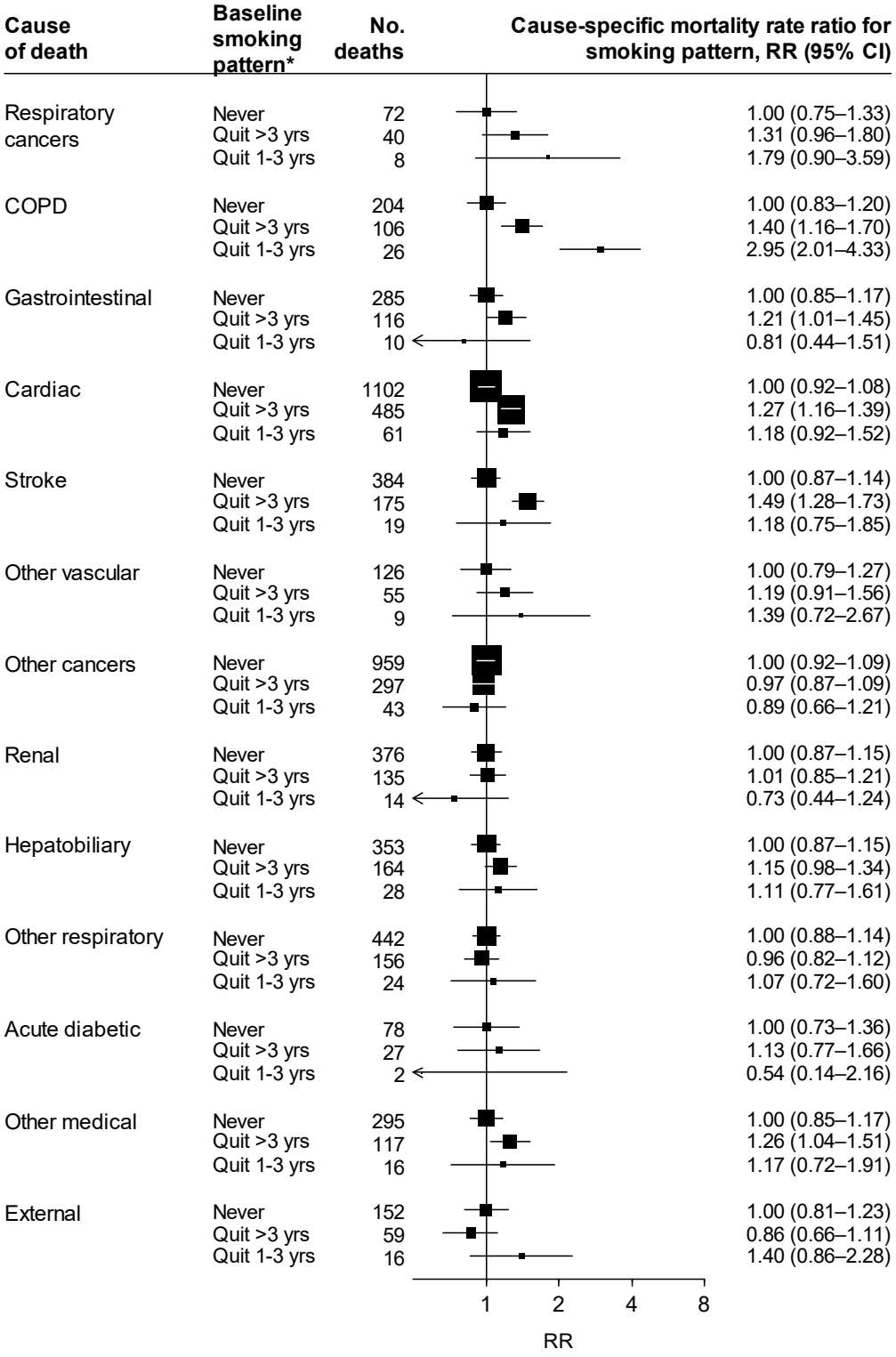
Participants with chronic disease at recruitment were excluded from these analyses.



*Precise number of years since quitting smoking was not reported.

Figure 4.12: Cause-specific mortality at ages 35-89, never vs ex-daily smokers, by quitting duration

Participants with chronic disease at recruitment were excluded from these analyses.



*Precise number of years since quitting smoking was not reported.

Chapter 5: Smoking and mortality in the US National Health Interview Survey

5.1 Chapter aims

The contemporary hazards of smoking and benefits of quitting in the United States have been previously reported using data from several large, prospective studies,[8,112] including the National Health Interview Survey.[4,98] However, findings reported in **Chapter 3** and **Chapter 4** of this thesis have raised new hypotheses that warrant a re-examination of US data; namely, whether the low-intensity and low-frequency smoking patterns and subsequently lower mortality risk found in Mexicans is also found in Mexican Americans in the US, and whether the hazards of smoking initiation at a young age found in Cuba are also found in the US. Previous studies of US populations have tended to report overall estimates adjusted for race/ethnicity (here referred to as 'race'),[4,8,112] and to focus on a narrow range of ages for smoking initiation,[98,99] both possibly due to limited power. As a result, neither the relevance of differential smoking patterns by race to mortality, nor the relevance of young smoking initiation (especially before age 10) to mortality, has been established directly and reliably in a contemporary US population.

The objective of this chapter is to report the contemporary hazards of smoking and benefits of quitting in the US National Health Interview Survey (NHIS), with a focus on the relevance of smoking patterns by race and age at smoking initiation to mortality. Specific aims are:

1) To assess the prevalence of smoking patterns across birth cohorts by race and sex

2) To estimate the relevance of daily smoking and smoking cessation to mortality in the United States (pooled across races) using data from the NHIS linked to the National Death Index, with an emphasis on age started smoking
3) To characterise the association between daily smoking and mortality by race
4) To characterise the association between smoking cessation and mortality by race
5) To estimate the smoking-attributable mortality by race in the United States
6) To interpret findings in the context of prior evidence of smoking in the United States and other populations globally

5.2 Baseline survey

5.2.1 Study design

The National Health Interview Survey (NHIS) is a cross-sectional household interview survey conducted annually in the United States. Conducted by the National Center for Health Statistics (NCHS), the NHIS serves as a primary source of information on the health of the US population. Although the NHIS has been conducted continuously for over 60 years, it has been restructured and updated periodically, which limits its long-term comparability. A major revision of the survey was conducted in 1997, and since then the NHIS has been standardised to allow for analysis of trends and pooling of data. Therefore, only data from surveys conducted since 1997 were included in this thesis, and from this point forward all discussion of NHIS will refer only to survey waves conducted since that time.

5.2.2 Study population

The NHIS is designed to be nationally representative of the civilian, noninstitutionalized population of the United States, and draws its sample from all 50 states and the District of Columbia.[113] A stratified, multistage sample design is used in order to ensure representative sampling of households, with minority respondents (particularly non-

Hispanic Black, Hispanic, and Asian participants) oversampled to improve the precision of estimates for these groups. However, for nationally-representative estimates, the relevant sample weights are calculated to ensure that estimates take into account the true national proportion of respondents according to sex and race/ethnicity. Response rates varied by year, but the annual response rate was typically around 70% of eligible sample households. In 2014, the latest survey wave for which mortality follow-up was conducted (described below), the household response rate was 74%, and among participating households, the response rate for adults was 80.5%.[114]

5.2.3 Data collection

Data were collected through personal household interviews conducted by interviewers using a laptop computer enabled with computer assisted personal interviewing (CAPI) software. Interviewers were trained and employed by the United States Census Bureau and followed standardised interview procedures. Selected households were mailed an advance letter describing the NHIS's purpose, time commitment, and voluntary nature prior to the initial visit. During the interviewer's visit, respondents were provided with another copy of the advance letter, and the interview proceeded only after verbal consent to participate was given. One household member of legal age to be considered an adult (depending on the state, and usually age 18 years) provided basic demographic and relationship information about all members of the household and the family or families living there. For each family in the household, a family member provided similar demographic and relationship information for all children and adult family members. One adult per family was randomly selected for the Sample Adult questionnaire, which collected information on a range of health behaviours (including smoking habits), medical history, healthcare access and utilization, and demographic information. The exposures of interest for this chapter are described in detail below.

Individual survey waves of the NHIS are released annually by the NCHS through its online portal (https://www.cdc.gov/nchs/nhis/data-questionnaires-documentation.htm).[113,114] The individual participant data files are released to the general public (with some limitations, described below) along with supporting documentation. I downloaded and cleaned the relevant modules of each annual survey and merged/appended them as necessary to create the final dataset used for prospective analyses. As similar procedures were used for each survey wave, surveys conducted in 1997-2014 were pooled for the main analyses in this chapter. For each annual wave of the survey, the NCHS' Research Ethics Review Board reviewed and approved the survey content and methods of data collection.

5.3 Resurvey

The NHIS is designed to be a cross-sectional survey, and participants are only re-contacted by telephone to complete unfinished interviews. No attempt was made to resurvey participants after the baseline survey. Therefore, only information provided at baseline was available (other than vital status, described below), and changes in smoking behaviour over the follow-up period could not be assessed.

5.4 Follow-up

Participants were followed-up to 1 January, 2016 through linkage to the National Death Index (NDI). The NDI is periodically linked to several NCHS surveys, including the NHIS, to allow for health outcomes to be studied. At the time of writing, the most up-to-date version of the NDI linkage (released February, 2019) was to NHIS survey waves from 1986-2014 (of which those from 1997-2014 were used). Linkage is conducted through probabilistic matching using a combination of Social Security Number (either in its entirety or final 4 digits), name, date of birth, sex, race, marital status, state of residence, and father's surname, as needed.[115] Sufficiently detailed information to reliably link

participants to the NDI was collected for nearly 96% of adult participants between 1997-2014, ranging from 92% complete in 2003 to 99% complete in 2010. The 4% of participants with insufficient information for reliable follow-up were excluded from all analyses. Causes of death from death certificate records were coded according to the International Classification of Diseases (ICD) 9th edition (for deaths during 1997-1998) and 10th edition (1999-2015).

There are two versions of mortality files available for linkage to the NHIS survey: public-use (de-identified and freely available online) and restricted-use (which are more detailed, must be accessed in person at secure research facilities, and are associated with access fees). To protect participant confidentiality, public-use mortality files include only a limited subset of variables and are limited to adult participants. Further, public-use mortality files are subject to data perturbation techniques to enhance participant anonymity. For this thesis, public-use versions of these linked mortality files were used due to access restrictions associated with the restricted-use files. Although the public-use mortality files are subject to perturbation techniques and less precise time scales (dates of baseline survey and censoring were rounded to the nearest quarter), this should have had little impact on exposure-outcome relationships.[116] Another limitation of the public-use mortality files is that the underlying cause of death is reported in only ten categories (given, with underlying ICD codes, in Table 5.1). Although the NHIS can additionally be linked to other health outcome datasets (including Medicare and Medicaid enrolment and claims records from the Centers for Medicare and Medicaid Services [CMS]), these linkages were considered outside the scope of this thesis.

5.5 Methods

5.5.1 Assessment of smoking

As in the Cuban and Mexican studies, smoking behaviour was self-reported by participants according to their responses to the standardised questionnaire. Participants were classified as ever-regular smokers if they answered “yes” to the question “Have you smoked at least 100 cigarettes in your ENTIRE LIFE?” Those who were ever-regular smokers then reported the age at which they first started to smoke fairly regularly (ages 6-94 years allowed; ages ≤6 years were assigned a value of 6). All ever-regular smokers then answered the question “Do you NOW smoke cigarettes every day, some days, or not at all?” Participants who reported now smoking not at all were asked how long it had been since they quit smoking cigarettes (as a number of days, months, or years). Those who currently smoked daily were asked how many cigarettes they smoked per day, and those who smoked only some days were asked on how many of the past 30 days they smoked a cigarette, and how many cigarettes they smoked on the average day they did smoke. Participants were also asked about use of tobacco products other than cigarettes, but because these products were reported as a single group (including cigars, pipes, hookahs, bidis, or cigarillos), the exposure of interest for these analyses was cigarette smoking.

Those who had reported not smoking at least 100 cigarettes in their lives were classified as never smokers, while those who had smoked at least 100 cigarettes but did not smoke at all at baseline were classified as ex-smokers. As in the Cuban study, for ex-smokers, the frequency and amount previously smoked was not recorded. Those who smoked some days were classified as non-daily smokers, and those who smoked every day were classified as daily smokers.

5.5.2 Assessment of race/ethnicity

To determine a respondent's race or ethnicity (henceforth referred to as 'race'), respondents were asked a series of deterministic questions. First, respondents were asked, "Do you consider yourself Hispanic/Latino?" If they answered "yes," they were then prompted with a list of eight choices to select the option that best represented their Hispanic origin or ancestry, including Puerto Rican, Cuban/Cuban American, Dominican (Republic), Mexican, Mexican American, Central or South American, Other Latin American, and Other Hispanic/Latino. Although the precise origin/ancestry options given varied slightly between years, Mexican and Mexican American were given separately as options for all years. Next, respondents were asked, "What race do you consider yourself to be?" Again, options varied slightly over survey years, but the primary options of interest for this thesis (White and Black/African American) were included for all survey years. Respondents who identified with more than one race were subsequently asked, "Which one of these groups, that is [options read aloud] would you say BEST represents your race?"

Respondents who self-identified as Mexican or Mexican American were considered together as 'Mexican American,' while those who self-identified as Hispanic/Latino but identified any origin or ancestry other than Mexican or Mexican American were considered 'other Hispanic.' (Although otherwise of interest for this dissertation, Cuban Americans were not analysed separately due to their limited sample size.) Those who did not identify as Hispanic or Latino were then categorised as either non-Hispanic White, non-Hispanic Black, or other race/ethnicity.

5.5.3 Selection of covariates

The process for selecting covariates in each set of prospective analyses in this thesis was described previously in **Chapter 3**. Although the same set of demographic,

socioeconomic, and behavioural factors were adjusted for as in **Chapter 3**, the exact categories used to adjust for confounding differed slightly. The specific categorisations are defined below.

5.5.4 Participant exclusions

An initial dataset of 552 837 individuals was constructed using data from survey waves in 1997-2014 comprised of those who were interviewed for the Adult module of the NHIS, after which exclusions were made sequentially. To increase the likelihood of restricting analyses to smokers who had formed long-term habits, and for consistency with the other cohorts in this thesis, those age <30 years at recruitment were excluded (n=106 306). Those age >84 years at recruitment were also excluded, as all participants aged ≥85 years were recorded as age 85 during NHIS data collection, leaving their true age unknown (n=13 284). Analyses were restricted to those who were eligible for mortality follow-up (ie, those who had sufficient identifying data to be followed for mortality, as determined by the NCHS), excluding 18 749 participants.

Those with missing values for smoking or the covariates of interest for the main analyses as described above were also excluded. This included 2706 participants with missing values for smoking frequency (ie, unable to classify as never, ex, non-daily, or daily smoker), 2281 with missing values for education, and 5526 with missing values for alcohol consumption. This resulted in an analytical dataset of 402 793 participants for the main analyses.

The potential impact of reverse causation, whereby prior disease may have led to behaviour change, was examined (primarily in analyses of smoking cessation) by further excluding those with chronic disease at recruitment. Chronic disease was defined as ever having been diagnosed with cancer, heart disease, stroke, diabetes, or chronic obstructive pulmonary disease (COPD). A further 80 443 participants with chronic

disease were excluded from these sensitivity analyses (excluding 37 167 never smokers, 29 705 former smokers, and 13 571 daily smokers).

5.5.5 Statistical analysis

As in **Chapter 3** and **Chapter 4**, cross-sectional analyses of the prevalence of smoking behaviours are unadjusted (and un-weighted, except where specified), and are presented by sex and birth cohort. For prospective analyses, Cox regression was used to calculate mortality rate ratios (RRs) and accompanying 95% CIs according to categories of smoking at recruitment. Participants were followed up until the date of death, the end of the risk period under consideration (age 90 for most analyses), or the end of the follow-up period (1 January 2016).

Prospective results are presented initially for pooled races (to provide a benchmark of 'overall' results), and subsequently for race-specific findings (to examine divergences from pooled results). Analyses by race were conducted at ages 30-89 (to improve the precision of estimates), while certain analyses of pooled race were conducted at ages 30-69 and 70-89 separately to enhance comparability to other findings in this thesis (eg, smoking from age <10 in Cuba). Because of limited power to examine cause-specific associations between smoking and mortality, the main analyses were limited to all-cause mortality. (However, limited cause-specific analyses for pooled results are presented in Appendix C). As in **Chapter 3** and **Chapter 4**, prospective analyses were stratified by age-at-risk (in 5-year age groups), sex, region, and further adjusted for formal education completed (<high school, high school, some university, or university), and frequency of alcohol consumption (never, former, <weekly, 1-<3 times/week, or ≥3 times/week). Similarly, the proportional hazards assumption was examined by estimating RRs in separate 5-year periods (ie, 0-5, >5-10, >10-15, and >15-20 years after recruitment). Analyses of pooled race further adjusted for race (Mexican American, other Hispanic,

non-Hispanic White, non-Hispanic Black, other). As in **Chapter 3** and **Chapter 4**, group-specific CI's are plotted for comparisons involving at least three categories of exposure, but conventional CI's are reported for any direct comparisons made between two groups (and are used in text to describe comparisons).

Following on findings in prior chapters of this thesis, cross-sectional analyses reported non-daily and daily smoking behaviours separately, but prospective analyses focused on the association between daily versus never smoking at recruitment and mortality. For analyses of pooled races, fine groupings were used to distinguish shape of association and maximise comparability to previous research. Categories of age started smoking were defined to align with the Cuban findings (eg, age<10, 10-14) and to align with US smoking policies (legal smoking ages are 18 or 21 years in most states, so categories included ages 15-17, 18-20, and ≥21 to determine the potential relevance of these policies to health). Analyses of amount smoked were categorised to take into account digit preference, and to create groups with roughly equal numbers of participants, resulting in the following categories of cigarettes/day: 1-9, 10, 11-19, 20, 21-30, or >30. For comparability to prior research, and comparability to other results in this thesis, analyses of ex-smokers examined both age at smoking cessation (age 15-24, 25-34, 35-44, 45-54 years) and duration of smoking cessation at recruitment (<5, 5-14, 15-24, 25-34, 35-39 years).

As race-specific analyses were comparably limited in power, these analyses of current and ex-smokers were conducted with broader categories. In part for consistency with prior analyses in this thesis, and in part to maximise power, analyses of daily smoking were subsequently dichotomised by amount smoked (<10 and ≥10 cigarettes/day) and age started smoking (<18 and ≥18 years). Similarly, race-specific RRs were calculated according to years since quitting prior to recruitment (<10, 10-19, and 20-39 years) and

age at quitting (15-34, 35-44, and 45-54 years). As no resurvey was conducted among study participants, unlike analyses in prior chapters of this thesis, the RRs associated with amount smoked were plotted at the mean cigarettes/day reported at recruitment (ie, did not take into account 'usual' amount smoked). Methods for calculating the race- and sex-specific population attributable fractions (PAFs) associated with smoking at ages 30-89 were similar to those used in **Chapter 3** and **Chapter 4**.

Estimates of the prevalence of adult smokers in the US were calculated using the most recent NHIS survey available (conducted in 2018) to determine the relevance of prospective findings to the contemporary US population as a whole. To produce nationally-representative estimates, survey weights taking into account the survey design were used (as provided by the NCHS). These calculations provide estimates of both the prevalence of different smoking patterns and the absolute number of smokers nationally that fall into each smoking category (eg, those who began smoking daily at ages <10, 10-14, etc.) and number of smokers by race.

5.6 Results

5.6.1 Participant characteristics by race

There were 402 793 participants overall aged 30-84 at recruitment, described in Table 5.2. Of these, 36 392 were Mexican American, 26 453 were 'other Hispanic', 261 683 were non-Hispanic White, 56 738 were non-Hispanic Black, and 21 527 were 'other race'. Overall, the average age was 52 years (SD 15), 44% of participants were male, 37% of participants resided in the Southern US and 23% resided in the Western US, approximately one quarter had graduated from university (26%), and 21% had a history of major chronic disease at recruitment. On average, Mexican Americans were several years younger (mean age 46 years, SD 13), were more likely to reside in the West (54%) and less likely to reside in the Northeast (2%), less likely to have completed high

school or higher education (48%), and more likely to have a history of diabetes (12%) but less likely to have a history of IHD (4%) than other participants. Other Hispanics were broadly similar to Mexican Americans, though they were more likely to live in the Northeast (32%) and to have completed high school or higher education (65%). In contrast, non-Hispanic Whites were older on average (mean age 53 years, SD 15), were well represented across regions (19% resided in the Northeast, the least-represented region), were more likely to have completed high school or higher education (88%), and were more likely to have been diagnosed with IHD (8%) or cancer (11%). Non-Hispanic Blacks were somewhat less likely to be male (39%), more likely to live in the South (57%) and to have been diagnosed with diabetes (15%), while participants of 'other race' were more likely to live in the West (48%) and more likely to have graduated from university (43%). These characteristics were very similar among men and women of the same race/ethnicity (Table 5.3).

5.6.2 Trends in smoking behaviour across birth cohorts by race

Among Mexican Americans, 40% of men and 18% of women were ever smokers. However, approximately two-thirds of both men and women had quit smoking by recruitment, and smoking cessation was positively associated with age (Figure 5.1). Among men, ever smoking prevalence was approximately 60% among those born before 1950, but less than 40% among those born in the 1960's and 1970's, and the majority of ever-smokers born before 1955 had quit by recruitment. Among women, ever smoking prevalence was slightly over 20% among those born in the 1950's or earlier, and somewhat lower among more recent birth cohorts, and the majority of those born before 1955 had quit by recruitment. Among both men and women, non-daily smoking was more common among younger participants (comprising more than one-third of current smokers born after 1960), and smoking ≥20 cigarettes/day was uncommon at all ages. These patterns were very similar for other Hispanics, although non-daily smoking

was less common, while smoking ≥20 cigarettes/day was somewhat more common (Figure 5.2).

Among non-Hispanic Whites, ever-smoking prevalence was approximately 70% among men born in the 1920's and approximately 45% among men born in the 1960's and 1970's, while it was 40% among women born in the 1920's and 50% among women born in the 1940's (and slightly lower in subsequent birth cohorts; Figure 5.3). Smoking cessation was strongly associated with age: among those born before 1940, 78% of male ever-smokers and 85% of female ever-smokers had quit smoking by the time of recruitment, while approximately 45% of male and female smokers born after 1960 had quit by recruitment. In contrast to Mexican Americans and other Hispanics, heavy smoking was common among non-Hispanic Whites, with two-thirds of male smokers and half of female smokers smoking ≥20 cigarettes/day. Non-daily smoking was uncommon at all ages. These patterns were generally similar among participants who identified as non-Hispanic Black or other race/ethnicity, though smoking ≥20 cigarettes/day was less common than among non-Hispanic Whites (Figures 5.4 and 5.5).

Figure 5.6 shows the mean number of cigarettes smoked/day and age started smoking among daily smokers by race and sex. Within each race/ethnicity, men smoked more than women (approximately 2-4 cigarettes/day on average), and men started smoking younger than women. However, although women of older birth cohorts (ie, born before 1940) tended to start smoking at older ages, women born after 1960 began smoking at approximately the same age as men. Consumption was lowest among Mexican Americans (9 cigarettes/day [SD 7] among women and 11 cigarettes/day [SD 8] among men), and highest among non-Hispanic Whites (17 cigarettes/day [SD 9] among women and 21 cigarettes/day [SD 11] among men). However, average age at smoking initiation was generally similar among men by race, ranging from 17 years (SD 5) among non-

Hispanic Whites to 18 years (SD 5) among Mexican Americans, other Hispanics, non-Hispanic Blacks, and 'other race' participants. Among women, on average, non-Hispanic Whites started smoking the youngest (age 18 years [SD 5]), and Mexican Americans started the oldest (age 20 years [SD 7]).

5.6.3 Characteristics of participants by race and smoking status at recruitment

Table 5.4 shows the characteristics of participants aged 30-84 at recruitment by race and smoking status. As described above, among current smokers, Mexican Americans were on average more likely to smoke less than daily (39%) and smoke fewer cigarettes/day among daily smokers (mean 11) than participants of other races/ethnicities, while non-Hispanic Whites were less likely to smoke less than daily (16%) and smoked more cigarettes/day (mean 19). Among ever-smokers, the prevalence of smoking cessation at recruitment was lowest among non-Hispanic Blacks (44%) and highest among non-Hispanic Whites (57%). However, as there was no resurvey conducted on these data, it is not known whether these baseline smoking patterns persisted among participants.

Among Mexican Americans, compared with never smokers, daily smokers were more likely to be male (61% and 37%, respectively), similarly likely to have graduated from high school (47% and 48%, respectively), more likely to consume alcohol at least weekly (36% and 15%, respectively), and more likely to be born in the US (53% and 38%, respectively). These patterns were generally similar among other Hispanics. Among non-Hispanic Whites, compared with never smokers, daily smokers were more likely to be male (49% and 39%, respectively), less likely to have graduated from high school (82% and 91%, respectively), and were more likely to consume alcohol at least weekly (37% and 29%, respectively). These patterns were again generally true for non-Hispanic Blacks and participants of other race/ethnicity. However, while >90% of non-Hispanic

White and non-Hispanic Black participants were born in the US overall, daily smokers of other/race ethnicity were more likely to be born in the US than never smokers: 62% and 29%, respectively).

5.6.4 Pooled results: Association between smoking (and quitting) and mortality

5.6.4.1 Participant characteristics by age started smoking

In order to produce estimates comparable to those from the Cuba Prospective Study, analyses of the relationship between age started smoking and mortality focused on death at ages 30-69 (although the relationship to mortality at ages 70-89 is also reported in Appendix C). The characteristics of the 326 456 participants aged 30-69 who were either never smokers, ex-smokers, or daily smokers at recruitment are shown in Table 5.5. Of the 63 101 daily smokers, 1225 (2%) began smoking regularly before age 10, and 12 111 (19%) began at ages 10-14. Compared with never smokers, daily smokers at recruitment were on average more likely to be male (50% and 40%), more likely to consume alcohol at least weekly (37% and 25%), less likely to have graduated from high school (78% and 85%), and more likely to be non-Hispanic White (71% and 57%). Compared with the average daily smoker, participants who began smoking at ages <10 and 10-14 were on average more likely to be male (68% and 55% versus 50%, respectively), more likely to smoke heavily (mean 21 [SD 13] and 20 [SD 11] cigarettes/day versus 17 [SD 10], respectively), and less likely to have graduated from high school (58% and 70% versus 78%, respectively).

5.6.4.2 Association between age started smoking and mortality

Over 3.5 million person-years of follow-up at ages 30-89 (mean 9 years, SD 5), there were 48 440 deaths: 3242 among Mexican Americans, 2177 among other Hispanics, 34 482 among non-Hispanic Whites, 6978 among non-Hispanic Blacks, and 1561 among participants of other race/ethnicity. Of these, 17 994 occurred at ages 30-69.

Compared with never smokers, smoking initiation at any age was associated with at least twice the risk of premature death (death at ages 30-69), but smoking initiation at younger ages was associated with even greater excess risk than smoking initiation at older ages. Compared with never smokers, the all-cause mortality RRs at ages 30-69 associated with smoking initiation at ages <10, 10-14, 15-17, 18-20, and ≥21 years were 3.78 (3.32-4.31), 2.94 (2.78-3.11), 2.45 (2.34-2.58), 2.21 (2.09-2.33), and 2.05 (1.93-2.19), respectively (Figure 5.7); these risks were associated with smoking initiation, on average, at ages 8, 13, 16, 19, and 26, respectively, compared with never starting to smoke. Starting to smoke at younger versus older ages was associated with a similar pattern of excess mortality risk at ages 30-69 for deaths from respiratory, cardiac, and cancer-related causes (Figure C1 in Appendix C). However, the observed excess mortality risk associated with starting to smoke at younger versus older ages was substantially diminished at ages 70-89 (Figure C2 in Appendix C). As there was no resurvey conducted among these participants, it is not known whether participants who began smoking at different ages were more or less likely to have changed the amount smoked per day, or quit smoking entirely, over the follow-up period.

5.6.4.3 Association between amount smoked and mortality

Compared with never smokers, smoking any number of cigarettes daily was associated with at least twice the risk of premature death (death at ages 30-69), but smoking a greater number of cigarettes/day was associated with even greater excess risk than smoking fewer cigarettes/day at recruitment. Compared with never smokers, the all-cause mortality RRs at ages 30-69 associated with smoking 1-9, 10, 11-19, 20, 21-30, and >30 cigarettes/day were 2.04 (1.89-2.19), 2.16 (2.02-2.31), 2.13 (1.97-2.31), 2.45 (2.33-2.57), 2.81 (2.62-3.02), and 3.55 (3.31-3.81), respectively (Figure 5.8). These risks were associated with smoking, on average, 5, 10, 15, 20, 29, and 42 cigarettes/day at recruitment, respectively. Although a dose-response relationship was observed between

amount smoked/day and all-cause mortality at ages 30-69, the relationship appeared to be generally linear only beyond >10 cigarettes/day at recruitment, above which approximately double the amount smoked was associated with less than double the risk. For example, compared with never smokers, the excess risk associated with smoking about 5, 10, 15, 20, 30, and 40 cigarettes/day was 104%, 116%, 113%, 145%, 181%, and 255%, respectively. These relationships were similar at ages 70-89, and for deaths from respiratory, cardiac, and cancer-related causes (Figures C3 and C4 in Appendix C). Again, the observed association between amount smoked and all-cause mortality was based on information reported at baseline, and in the absence of resurvey information, the 'usual' smoking behaviours of these participants are not known.

5.6.4.4 Association between smoking cessation and mortality

Among ex-smokers without chronic disease at recruitment who quit at ages 15-54, quitting by approximately age 40 avoided most excess risk of continued smoking. Compared with never smokers, the RR for all-cause mortality at ages 55-89 associated with quitting at ages 15-24, 25-34, 35-44, and 45-54 years was 0.95 (95% CI 0.86-1.05), 0.97 (0.91-1.04), 1.13 (1.07-1.20), and 1.36 (1.29-1.43), respectively, and was 2.88 (2.77-2.99) for continuing smokers (Figure 5.9). Similarly, the RR for all-cause mortality at ages 30-89 attenuated with quitting duration: quitting <5 years, 5-14 years, 15-24 years, 25-34 years, and 35-39 years prior to recruitment was associated with RRs of 1.82 (1.72-1.93), 1.52 (1.45-1.60), 1.17 (1.12-1.24), 1.07 (1.01-1.13), and 0.99 (0.90-1.10), compared with never smokers, and was 2.63 (2.54-2.72) for continuing smokers (Figure 5.10). These patterns were similar for mortality from respiratory causes, heart disease, and cancer (Figures C5 and C6 in Appendix C). Sensitivity analyses including participants with chronic disease at baseline resulted in somewhat higher RRs for all-cause mortality, especially among those quitting at older ages and those quitting more

recently, which may be evidence of the 'sick-quitter' effect (Figures C7 and C8 in Appendix C).

5.6.4.5 National estimates of smoking among US adults (NHIS 2018)

According to data from the latest survey wave of the NHIS (2018), which incorporates the sample weights provided by the NCHS to produce nationally-representative results for the population aged ≥18 years, there were an estimated 25 million daily smokers in the US in 2018 (Table 5.6). Of these, more than 5 million daily smokers began before the age of 15, of which an estimated 0.6 million began before age 10, and a further 4.6 million began at ages 10-14. Of the estimated 55 million ex-smokers in the US in 2018, 0.8 million had begun before age 10, and an additional 7.6 million had begun at ages 10-14, but reported having quit by the time of recruitment. The estimated number of adult smokers in the United States by race (and the unweighted number of participants used to produce these estimates) can be found in Tables C1 and C2 in Appendix C.

5.6.5 Association between smoking and mortality by race

5.6.5.1 Overall

Compared with never smokers, the all-cause mortality RR for current daily smokers ('daily smokers') was 1.39 (95% CI 1.24-1.54) among Mexican Americans, 2.45 (2.20-2.72) among other Hispanics, 2.74 (2.61-2.89) among non-Hispanic Whites, 2.29 (2.19-2.41) among non-Hispanic Blacks, and 2.08 (1.84-2.35) among participants of other race/ethnicity (Figure 5.11; $p_{heterogeneity}<.001$). As the RRs were little changed after excluding those with chronic disease at recruitment (Figure 5.12), all subsequent analyses of current daily versus never smokers included those with chronic disease at recruitment. Compared with never smokers, the all-cause mortality RRs were higher among Mexican Americans born in the US than among those born outside of the US: 1.61 (1.38-1.88) and 1.09 (0.92-1.29), respectively ($p_{heterogeneity}<.001$; 1052 and 1332

deaths, respectively). However, due to the limited precision of these estimates, subsequent analyses of Mexican Americans considered these two groups together.

Although there was insufficient power to conduct the analyses by race, there was little evidence that the overall all-cause mortality RRs comparing daily and never smokers varied over the follow-up period; the RRs at 0-5, >5-10, >10-15, and >15-20 years after recruitment were 2.53 (2.43-2.64), 2.43 (2.33-2.54), 2.51 (2.37-2.65), and 2.65 (2.39-2.94), respectively.

5.6.5.2 By age started and amount smoked

The excess risk of all-cause mortality at ages 30-89 was determined by both the age at smoking initiation and the amount smoked, with greater excess risk associated with starting younger and smoking more across race groups. Compared with never smokers, the all-cause mortality RRs associated with starting to smoke at ages <18 and ≥18 years were, respectively, 1.44 (1.25-1.66) and 1.34 (1.15-1.56) among Mexican Americans, 2.71 (2.37-3.09) and 2.17 (1.86-2.54) among other Hispanics, 3.02 (2.86-3.19) and 2.46 (2.32-2.61) among non-Hispanic Whites, 2.49 (2.30-2.70) and 2.13 (1.96-2.31) among non-Hispanic Blacks, and 2.16 (1.83-2.56) and 2.02 (1.72-2.39) among participants of other race/ethnicity (Figure 5.13). Similarly, compared with never smokers, the all-cause mortality RRs associated with smoking <10 and ≥10 cigarettes/day were, respectively, 1.28 (1.09-1.50) and 1.48 (1.29-1.69) among Mexican Americans, 1.91 (1.56-2.33) and 2.69 (2.39-3.02) among other Hispanics, 2.46 (2.26-2.68) and 2.78 (2.64-2.93) among non-Hispanic Whites, 1.99 (1.79-2.21) and 2.42 (2.25-2.60) among non-Hispanic Blacks, and 1.66 (1.27-2.17) and 2.21 (1.94-2.53) among participants of other race/ethnicity (Figure 5.14).

5.6.6 Association between smoking cessation and mortality by race

The excess risk associated with ever smoking and mortality at ages 30-89 was considerably lower in those who had quit smoking than in those who had continued smoking at recruitment, with large reductions in excess risk among quitters for each race/ethnicity considered. The excess risk among ever smokers depended strongly on years since quitting smoking at recruitment and age at smoking cessation. To determine the relevance of the 'sick-quitter' effect, analyses including and excluding those reporting chronic disease at recruitment were reported separately. Among all participants (including those with chronic disease at recruitment), compared with never smokers, the all-cause mortality RRs associated with smoking cessation <10 years, 10-19 years, or 20-39 years before recruitment, respectively, were 1.38 (1.22-1.56), 1.08 (0.93-1.25), and 1.00 (0.88-1.13) among Mexican Americans, 1.69 (1.45-1.97), 1.46 (1.25-1.71), and 1.10 (0.96-1.27) among other Hispanics, 2.14 (2.02-2.27), 1.60 (1.51-1.69), and 1.25 (1.18-1.32) among non-Hispanic Whites, 1.87 (1.72-2.04), 1.28 (1.16-1.41), and 1.06 (0.97-1.16) among non-Hispanic Blacks, and 1.56 (1.31-1.86), 1.24 (1.02-1.52), and 1.00 (0.84-1.18) among participants of other race/ethnicity (Figure 5.15). Similarly, compared with never smokers, the all-cause mortality RRs at ages 55-89 associated with smoking cessation at ages 15-34, 35-44, or 45-54, respectively, were 1.06 (0.87-1.28), 1.20 (1.00-1.43), and 1.39 (1.18-1.63) among Mexican Americans, 1.13 (0.89-1.45), 1.21 (0.98-1.48), and 1.25 (1.06-1.48) among other Hispanics, 1.00 (0.93-1.07), 1.20 (1.11-1.29), and 1.48 (1.38-1.59) among non-Hispanic Whites, 1.07 (0.93-1.22), 1.15 (1.02-1.30), and 1.39 (1.25-1.55) among non-Hispanic Blacks, and 1.11 (0.87-1.43), 1.19 (0.94-1.50), and 1.41 (1.13-1.74) among participants of other race/ethnicity (Figure 5.16). Excluding those with chronic disease at recruitment produced similar results, although the excess risk associated with quitting more recently or at older ages was somewhat lower following exclusion of those with chronic disease, suggesting the

possibility that some of these quitters may have quit due to illness already caused by smoking (Figures 5.17 and 5.18).

Within each racial group, smoking cessation 10-19 years prior to recruitment was associated with a two-thirds (or more) reduction in the excess risk associated with continued smoking, and smoking cessation ≥20 years prior to recruitment was associated with a 90% (or more) reduction in excess risk compared with continued smoking at recruitment, avoiding nearly all of the excess risk associated with smoking. Similarly, smoking cessation before age ~40 was associated with little of the excess risk associated with smoking, and smoking cessation before age ~30 avoided nearly all of the excess risk associated with smoking.

5.6.7 Population attributable fraction of mortality associated with ever smoking by race

If the association between smoking and mortality at ages 30-89 was chiefly causal, then smoking caused an estimated 12% of male deaths and 4% of female deaths among Mexican Americans, 21% of male deaths and 13% of female deaths among other Hispanics, 31% of male deaths and 26% of female deaths among non-Hispanic Whites, 24% of male deaths and 16% of female deaths among non-Hispanic Blacks, and 18% of male deaths and 11% of female deaths among participants of other race/ethnicity (Table 5.7). Despite broadly similar ever-smoking prevalences across races within each sex, variation in the PAFs across races was driven by substantial differences in the RRs associated with ever smoking. Compared with never-smoking men, ever smoking among men was associated with a 28% increase in risk among Mexican Americans and a 69% increase in risk among non-Hispanic Whites; this 2.5-fold higher excess risk was associated with an approximately 2.5-fold larger PAF. Compared with never-smoking women, ever smoking among women was associated with a 21% increase in risk among Mexican Americans and a 94% increase in risk among non-Hispanic Whites; this 4.5-fold

higher excess risk was associated with an approximately 6.5-fold higher PAF due to higher ever-smoking prevalence among non-Hispanic White women. However, there is considerable uncertainty around the RRs that enter into these estimates. For example, the excess mortality risk associated with ever versus never smoking among Mexican-American women was 21% (95% CI 6%-37%).

5.7 Discussion

5.7.1 Summary of findings

Of the ~400 000 adults recruited into the US National Health Interview Surveys in 1997-2014, more than half of men and one-third of women smoked. However, Mexican Americans had the lowest ever-smoking prevalence of the racial/ethnic groups studied (28%, ignoring non-daily smokers) and the largest proportion of ever-smokers who had quit by recruitment (64%), in contrast to the ever-smoking prevalence of non-Hispanic Whites (49%) and the relatively smaller proportion of ever-smoking non-Hispanic Blacks who had quit (51%). Mexican Americans were also more likely to smoke non-daily than any other group, and smoked the fewest cigarettes/day (11 on average, compared with 19 among non-Hispanic Whites). However, as information about smoking behaviours was collected at a single point in time, it is not known whether these baseline smoking patterns persisted among participants over the follow-up period.

The excess risk of death at ages 30-89 associated with current daily smoking compared with never smoking was also lower among Mexican Americans (39%) than among other Hispanics (145%), non-Hispanic Whites (174%), non-Hispanic Blacks (129%), and participants of other race/ethnicity (108%). The relatively lower excess risk associated with smoking among Mexican Americans was also observed when stratifying by age started smoking (age <18 and age ≥18 years) and cigarettes smoked/day (<10 and ≥10) as reported at recruitment. However, the proportional excess risk avoided among those

who quit smoking in comparison to those who were daily smokers at recruitment was similar across races, with those who quit smoking ≥10 years prior to recruitment or before age ~40 avoiding most of the excess risk associated with continued smoking at recruitment. Despite lower risks associated with smoking among Mexican Americans than those observed in other racial groups, smoking nevertheless caused an estimated 12% of male and 4% of female deaths at ages 30-89.

In pooled analyses considering participants of different races jointly, additionally adjusting for race, smoking initiation after age 20 approximately doubled all-cause mortality rates at ages 30-69, and smoking initiation before age 15 approximately tripled mortality rates at these ages, with particularly high mortality rates among those who started before age 10. According to weighted estimates from the 2018 NHIS, approximately 5 million of the 25 million daily smokers in the US had started smoking regularly before age 15, of whom 0.6 million had begun before age 10. If the association between smoking and death at ages 30-69 was mostly causal, then about two-thirds of those who started smoking before age 15 were killed by their habit. These results suggest that adult smokers in the US who began smoking regularly before age 15 would likely benefit greatly from smoking cessation, and those who do not quit will be at particularly high risk of death at ages 30-69.

5.7.2 Strengths and limitations

A major strength of this study is its size, which made it possible to estimate directly and reliably the hazards of smoking and benefits of quitting by race in the US, one of the first such studies to report findings separately for Mexican Americans. Furthermore, this is the first study globally (alongside the Cuba Prospective Study, described in **Chapter 3**) to demonstrate reliably the hazards of smoking from before the age of 10.

This study had several other important strengths. There was enough follow-up time to allow for sufficient accrual of events to examine race-specific mortality reliably, despite the relatively low prevalence and intensity of daily smoking, particularly among Mexican Americans. The nationwide linkage to the National Death Index, which employed robust methods to accurately capture deaths among participants over the follow-up period,[115] limited loss to follow-up despite potential movement of participants across the country. The highly standardised questionnaire across survey years ensured that similar information was collected from all participants, and again collected sufficiently detailed information to distinguish between daily and non-daily smokers, as well as age at smoking cessation, which are critical methodological details in the Mexican-American population in which non-daily smoking and smoking cessation were common. This also made it possible to examine the potential effect of chronic disease prior to recruitment on the relationship between smoking and mortality, as in **Chapter 3** and **Chapter 4**. Finally, the NHIS surveys are designed to be nationally representative, which enhanced generalisability and ensured large numbers of participants of different ages, sexes, and races were recruited. Detailed identifying information was collected for nearly all participants, which allowed for comprehensive follow-up of the participants (reducing the likelihood of bias due to differential follow-up). The NHIS survey design also allowed for weighted prevalence estimates for the US as a whole, which was particularly useful for estimating the prevalence of childhood smoking initiation among adults at the national level, in addition to the number of daily smokers by race.

There were also several limitations. One limitation of this study was that deaths occurring abroad are not captured in US vital registration data, and therefore the occurrence of deaths among Mexican Americans (in particular) who had returned to Mexico cannot be ruled out. However, there is no evidence that rates of migration would be different between smokers and non-smokers, and crude death rates from this study

are similar between foreign-born and US-born Mexican Americans. Furthermore, given that Mexican Americans and other Hispanics were on average several years younger than participants of other races/ethnicities at recruitment, their somewhat lower crude death rates suggest that this bias would likely be minimal.

An additional limitation is that these data relied on self-reported smoking behaviours (and other characteristics). Relying on self-reported information may have increased the imprecision of the underlying information (particularly for smoking), and may have increased the likelihood of under-reporting of smoking (and potential over-reporting of quitting). However, prior evidence from US and other populations suggests that the validity of self-reported smoking behaviours (among both self-reported never smokers and current smokers) are reasonably high, both as a binary exposure (ie, current or never) and as a continuous exposure (ie, number of cigarettes smoked per day).[36,117-119] Unlike the studies reported in **Chapter 3** and **Chapter 4**, no resurvey was conducted among participants in the US NHIS surveys, which limited the characterisation of smoking habits to information collected at a single point in time. Given that a number of participants were likely to have quit smoking over the course of follow-up, and some participants who reported having quit may have again taken up smoking, both the hazards of smoking and benefits of quitting may be underestimated (and the extent to which this occurred may vary by race).

The relevance of amount smoked is also more difficult to interpret, as, if smoking patterns from other populations also apply in this population, 'usual' amount smoked would be expected to tend towards the mean (ie, at resurvey, the average amount smoked would increase at lower baseline consumption levels and decrease at higher baseline consumption levels). Likely as a result, the slope of the association between amount smoked and mortality risk is shallower than would be expected from 'usual'

smoking patterns observed in other populations. Finally, use of these publicly-available data resulted in two further limitations: first, that these data were subject to some perturbation in order to reduce the likelihood of anonymised participant data being re-identified, and second, again to protect participant anonymity, causes of death were released as one of ten broad pre-specified causes. However, results for all-cause and cause-specific mortality should be very similar to those resulting from the true (unmodified) underlying data,[116] and cause-specific mortality was examined only as a sensitivity analysis of pooled data due to the limited numbers of deaths by race for relevant causes. Therefore, in practice, use of the public-use NHIS file should have had little impact on these findings.

5.7.3 Findings in the context of prior evidence

The cross-sectional smoking patterns observed in this population by race are similar to those previously documented. For example, Rodriquez *et al.* (2016) have previously reported the prevalence of smoking in the US NHIS surveys between 1997-2014, showing that participants with a Mexican background were especially likely to smoke <10 cigarettes/day compared with smokers of Puerto Rican, Cuban, or other Latino backgrounds (but similar to those of Central or South American backgrounds, though the sample size was much smaller).[120] The authors also reported that Latinos smoked fewer cigarettes/day than non-Latinos across birth cohorts (a difference of 5-7 cigarettes/day for all birth cohorts examined). Similarly, a study using data from the Hispanic Community Health Study/Study of Latinos conducted in four large cities in the US reported that Mexican Americans were more likely to smoke non-daily, and less likely to smoke ≥20 cigarettes/day, than participants of other Hispanic backgrounds.[121] Using data from the US National Health and Nutrition Examination Survey waves between 1999-2010, Jain (2014) reported mean cigarettes/day of 10.4 (95% CI 9.2-11.6) among

Mexican Americans, 11.4 (10.9-11.9) among non-Hispanic Blacks, 18.6 (18.1-19.1) among non-Hispanic Whites, and 13.7 (12.6-14.8) among all other participants.[122]

The relevance of current smoking to mortality among non-Hispanic Whites observed in the US NHIS population (RR for current versus never smokers: 2.74) was similar to that of previously-reported evidence from pooled contemporary US cohorts (RR for current versus never smokers: 2.76 among women and 2.80 among men)[8]. However, the relationship between smoking and mortality among Mexican Americans was appreciably weaker. Most prior evidence on the association between smoking and mortality among Mexican Americans relied on small studies or indirect estimates. Reynales-Shigematsu *et al.* (2017) indirectly estimated the tobacco-attributable mortality among Mexican-American men to be 11% in ~2002, 8% in ~2007, and 6% in ~2012, and among Mexican-American women to be 6% in ~2002, 5% in ~2007, and 4% in ~2012.[111] These indirect estimates are somewhat lower (for men, but not for women) than the direct estimates from this study of participants recruited around 2002, which suggested that approximately 12% of male deaths and 4% of female deaths were caused by smoking in ~2010. Using NHIS data from 1990-2004, Fenelon (2013) directly estimated the difference in life expectancy by race according to smoking status and reported that Mexican Americans had a lower prevalence of medium or heavy current smoking, which contributed to the 1-2 year life expectancy advantage of Mexican Americans in comparison to non-Hispanic Whites. Fenelon estimated that the smoking attributable fractions for US-born Mexican-American men and women were 26% and 15%, respectively, and for foreign-born Mexican Americans were 16% for men and 10% for women.[123] However, given the considerably lower smoking prevalence and smoking intensity among Mexican Americans, it is implausible that the smoking attributable fractions of Mexican Americans would be similar to those of non-Hispanic Whites, as reported by the author. A small study previously discussed in **Chapter 4**, conducted

among 2629 Mexican Americans and 1136 non-Hispanic Whites, reported all-cause mortality RRs associated with current smoking that were broadly compatible with those reported here, though the estimates lacked precision (age- and sex-adjusted RR 1.6, 95% CI 1.1-2.5).[83]

As described in **Chapter 3** in the context of childhood smoking in the Cuba Prospective Study, prospective evidence on smoking initiation before age 15 is limited, and has not yet been produced on smoking initiation before age 10. A study previously discussed in **Chapter 3**, using US NHIS data from 1997-2005, found limited evidence of elevated risk associated with smoking initiation before age 13 compared with smoking initiation at age ≥21 years.[98] A smaller study conducted by Huxley *et al.* (2012) among 14 200 participants from the Atherosclerosis Risk in Communities Study followed between 1987-2007 reported increased risk of cardiovascular disease among those who began smoking at ages ≤12 compared with those who began at ages ≥22 years (RR 2.7, 95% CI 1.9-3.9).[99] Despite limited prospective evidence (particularly from large-scale studies), a number of cross-sectional studies have reported smoking initiation before the age of 15. Cross-sectional evidence from the Global Youth Tobacco Surveys[124] and Global School-based Student Health Surveys[125] have documented smoking in late childhood/early adolescence (ie, ages 12-15) in countries around the world. However, whether these children and adolescents who experimented with tobacco became long-term smokers remains unclear, and evidence among adults reporting smoking initiation before age 15, and especially before age 10, remains limited. The prevalence and number of adult smokers who began smoking in childhood are reported in greater detail in **Chapter 6** using novel analyses of data from the Global Adult Tobacco Surveys (GATS) in 30 low- and middle-income countries.

5.8 Conclusion

This study found that Mexican Americans smoked relatively few cigarettes/day compared with other Hispanics, non-Hispanic Whites, non-Hispanic Blacks, and participants of other race/ethnicity. Compared with never smokers, the excess risk of all-cause mortality at ages 30-89 associated with daily smoking among Mexican Americans (39%) was consequently low in comparison to participants of other races: 145% among other Hispanics, 174% among non-Hispanic Whites, 129% among non-Hispanic Blacks, and 108% among participants of other race/ethnicity. Nevertheless, smoking caused an estimated 12% of deaths among Mexican-American men and 4% of deaths among Mexican-American women. However, smoking cessation by age ~40 avoided most of the excess risk associated with continued smoking among all participants, regardless of race. In analyses pooling participants of different races/ethnicities, smoking initiation before age 15 was associated with three times the all-cause mortality risk at ages 30-69 of never smoking, with especially high risk among those who began smoking before age 10. The estimated 5 million adult smokers in the US who began smoking before age 15, of whom 0.6 million began before age 10, are at especially high risk of premature death if they do not quit.

Table 5.1: ICD-10 codes for 10 cause-specific mortality endpoints (as provided by the National Center for Health Statistics, Maryland, USA)

Endpoint	ICD-10 codes
Cardiac	I00-I09, I11, I13, I20-I51
Cancer	C00-C97
Respiratory (chronic lower)	J40-J47
Cerebrovascular	I60-I69
Diabetes mellitus	E10-E14
Alzheimer's disease	G30
Influenza and pneumonia	J09-J18
Nephritis, nephrotic syndrome, and nephrosis	N00-N07, N17-N19, N25-N27
External (unintentional injuries)	V01-X59, Y85-Y86
All other causes	(residual of those above)
All-cause	A00-Y89

Table 5.2: Baseline characteristics of the US (NHIS) cohort participants at ages 30-84, by race/ethnicity

	NHIS (US) cohort participants, by race/ethnicity					
	Mexican-American	**Other Hispanic**	**Non-Hisp. White**	**Non-Hisp. Black**	**Other race**	**Overall**
Participants	36 392	26 453	261 683	56 738	21 527	402 793
Age, years	46 (13)	49 (14)	53 (15)	51 (14)	49 (14)	52 (15)
Male (%)	45%	42%	45%	39%	45%	44%
Region‡						
Northeast	2%	32%	19%	16%	16%	18%
Midwest	9%	5%	28%	18%	14%	22%
South	35%	39%	34%	57%	22%	37%
West	54%	23%	20%	9%	48%	23%
Education attained						
University	8%	17%	30%	17%	43%	26%
Some college	18%	23%	29%	30%	24%	27%
High school	22%	25%	29%	30%	20%	28%
<High school	52%	35%	12%	23%	13%	18%
History of chronic disease†						
Diabetes	12%	11%	9%	15%	10%	10%
Ischaemic heart disease	4%	5%	8%	7%	5%	7%
Stroke	2%	2%	3%	5%	3%	3%
Cancer	3%	4%	11%	5%	5%	9%
Emphysema	1%	1%	3%	1%	1%	2%
Any	16%	16%	22%	22%	17%	21%

Unweighted mean (SD) or percent shown for cohort participants.
‡ U.S. census regions.
† Self-reported history of diabetes, ischemic heart disease, stroke, cancer (any site), or emphysema.

Table 5.3: Baseline characteristics of the US (NHIS) cohort participants at ages 30-84, by sex and race/ethnicity

	Characteristics among MEN						Characteristics among WOMEN					
	Mexican-American	Other Hispanic	Non-Hisp. White	Non-Hisp. Black	Other race	All men	Mexican-American	Other Hispanic	Non-Hisp. White	Non-Hisp. Black	Other race	All women
Participants	16 451	10 999	117 485	22 302	9779	177 016	19 941	15 454	144 198	34 436	11 748	225 777
Age, years	46 (13)	48 (14)	53 (14)	51 (14)	49 (14)	51 (14)	47 (13)	49 (14)	54 (15)	51 (14)	50 (14)	52 (15)
Male (%)	100%	100%	100%	100%	100%	100%	0%	0%	0%	0%	0%	0%
Region‡												
Northeast	2%	30%	19%	15%	16%	17%	1%	34%	19%	17%	16%	18%
Midwest	10%	6%	28%	18%	14%	23%	8%	5%	27%	18%	13%	22%
South	35%	40%	33%	58%	23%	36%	36%	38%	34%	57%	22%	37%
West	53%	24%	20%	10%	47%	24%	54%	23%	19%	8%	49%	23%
Education attained												
University	8%	18%	32%	17%	47%	28%	8%	16%	28%	17%	39%	24%
Some college	17%	23%	28%	28%	23%	26%	19%	23%	30%	31%	24%	28%
High school	23%	25%	28%	32%	19%	28%	22%	25%	30%	29%	22%	28%
<High school	52%	33%	12%	23%	11%	18%	52%	35%	12%	23%	15%	19%
History of chronic disease†												
Diabetes	12%	11%	9%	14%	11%	10%	12%	11%	8%	15%	10%	10%
Ischaemic heart disease	5%	6%	10%	7%	7%	9%	4%	4%	6%	6%	4%	5%
Stroke	2%	2%	3%	5%	3%	3%	2%	2%	3%	5%	3%	3%
Cancer	2%	4%	10%	5%	4%	8%	4%	5%	12%	5%	6%	10%
Emphysema	1%	1%	3%	2%	2%	3%	1%	1%	2%	1%	1%	2%
Any	15%	15%	21%	22%	16%	20%	17%	16%	22%	22%	17%	21%

Unweighted mean (SD) or percent shown for cohort participants.
‡ U.S. census regions.
† Self-reported history of diabetes, ischemic heart disease, stroke, cancer (any site), or emphysema.

Table 5.4 (Part 1 of 2): Baseline characteristics of the US (NHIS) cohort participants at ages 30-84, by race/ethnicity and smoking status

	Mexican American				Other Hispanic				Non-Hisp. White			
	Never	**Ex**	**< Daily**	**Daily**	**Never**	**Ex**	**< Daily**	**Daily**	**Never**	**Ex**	**< Daily**	**Daily**
Participants	24 687	6133	2192	3380	17 287	4771	1183	3212	128 553	76 328	8843	47 959
Age, years	45 (13)	52 (14)	43 (11)	47 (12)	48 (14)	54 (14)	44 (11)	46 (12)	53 (15)	58 (14)	48 (13)	49 (12)
Male (%)	37%	64%	65%	61%	35%	56%	53%	51%	39%	52%	47%	49%
Smoking behaviour												
Mean cigarettes/day	-	-	-	10.5 (7.8)	-	-	-	12.7 (8.9)	-	-	-	18.8 (10.0)
Age started smoking (years)	-	18.2 (5.9)	19.7 (7.2)	18.5 (6.1)	-	18.2 (5.6)	19.6 (6.5)	18.1 (5.9)	-	17.7 (4.7)	19.4 (6.2)	17.6 (5.0)
Other characteristics												
High school graduate	48%	50%	49%	47%	65%	67%	69%	63%	91%	87%	90%	82%
Consume alcohol weekly	15%	27%	42%	36%	15%	28%	35%	32%	29%	38%	46%	37%
Born in the US	38%	50%	45%	53%	33%	44%	47%	55%	95%	96%	95%	97%
Region‡												
Northeast	2%	2%	2%	2%	33%	30%	34%	36%	19%	21%	20%	17%
Midwest	9%	9%	10%	10%	5%	6%	6%	6%	28%	26%	29%	29%
South	35%	35%	40%	37%	40%	37%	31%	37%	33%	32%	31%	37%
West	54%	55%	48%	51%	22%	28%	29%	22%	21%	20%	20%	17%

Unweighted mean (SD) or percent shown for cohort participants.
‡ U.S. census regions.

Table 5.4 (Part 2 of 2): Baseline characteristics of the US (NHIS) cohort participants at ages 30-84, by race/ethnicity and smoking status

	Non-Hisp. Black				Other Race			
	Never	**Ex**	**< Daily**	**Daily**	**Never**	**Ex**	**< Daily**	**Daily**
Participants	32 923	10 528	3182	10 105	14 085	3910	766	2766
Age, years	49 (14)	58 (13)	49 (12)	49 (12)	49 (14)	54 (14)	45 (12)	47 (12)
Male (%)	33%	49%	47%	48%	38%	62%	60%	59%
Smoking behaviour								
Mean cigarettes/day	-	-	-	12.8 (8.0)	-	-	-	14.1 (9.0)
Age started smoking (years)	-	18.5 (5.0)	20.8 (6.8)	18.8 (5.5)	-	18.6 (5.1)	20.3 (6.6)	18.5 (5.7)
Other characteristics								
High school graduate	81%	72%	74%	70%	88%	86%	85%	81%
Consume alcohol weekly	15%	21%	36%	38%	14%	29%	38%	34%
Born in the US	87%	95%	96%	97%	29%	53%	59%	62%
Region‡								
Northeast	16%	16%	16%	15%	17%	13%	12%	13%
Midwest	16%	19%	20%	21%	13%	14%	16%	17%
South	59%	55%	54%	55%	21%	22%	23%	28%
West	8%	10%	10%	9%	49%	51%	49%	42%

Unweighted mean (SD) or percent shown for cohort participants.
‡ U.S. census regions.

Table 5.5: Pooled US (NHIS) baseline characteristics by smoking habits, ages 30-69

	Never smoker	Ex-smoker	Daily smoker	Daily smokers, by age began regular smoking					Overall†
				< 10	10-14	15-17	18-20	> 20	
Cohort participants (1997-2014)	186 040	77 315	63 101	1225	12 111	21 381	16 132	12 252	326 456
Age, years	46 (11)	52 (11)	47 (10)	49 (11)	47 (10)	46 (10)	47 (10)	48 (10)	48 (11)
Male (%)	40%	53%	50%	68%	55%	50%	49%	44%	45%
Smoking behaviour									
Age started smoking (years)§	-	18 (4)	18 (5)	8 (1)	13 (1)	16 (1)	19 (1)	26 (5)	18 (5)
Cigarettes/day at baseline	-	-	17 (10)	21 (13)	20 (11)	18 (10)	16 (9)	15 (9)	17 (10)
Alcohol consumption									
Consume alcohol weekly	25%	37%	37%	36%	37%	38%	38%	34%	30%
Education									
High school graduate	85%	86%	78%	58%	70%	76%	84%	84%	84%
Race/ethnicity									
Non-Hisp. White	57%	73%	71%	72%	75%	74%	70%	62%	64%
Non-Hisp. Black	16%	10%	15%	12%	11%	14%	16%	21%	14%
Hispanic	21%	12%	10%	11%	10%	9%	10%	12%	17%
Other	7%	4%	4%	5%	4%	4%	5%	5%	6%
Region‡									
Northeast	17%	19%	17%	14%	18%	18%	17%	15%	18%
Midwest	21%	23%	26%	24%	25%	26%	27%	24%	22%
South	37%	34%	39%	41%	38%	39%	38%	41%	37%
West	25%	24%	18%	20%	19%	17%	18%	20%	23%

Unweighted mean (SD) or percent shown for cohort participants.
† Includes all cohort participants, except for age started smoking (limited to current daily and ex-smokers) and cigarettes/day at baseline (limited to current daily smokers).
§ Those who reported smoking from younger than age 6 were assigned age 6 during NHIS data collection.
‡ U.S. census regions.

Table 5.6: Estimated number of adult smokers (age≥18) in the United States in 2018 by age started smoking

Smoking group	Unweighted n	Weighted n
Daily, started:		
Age<10	69	620,000*
Age 10-14	517	4,640,000
Age 15-17	971	8,870,000
Age 18-20	729	6,750,000
Age >20	504	4,230,000
Non-daily	924	8,660,000
Ex	6,283	54,990,000
Never	15,292	159,570,000

Estimated from the US NHIS 2018 survey (the most recently-available survey wave), rounded to the nearest 10,000

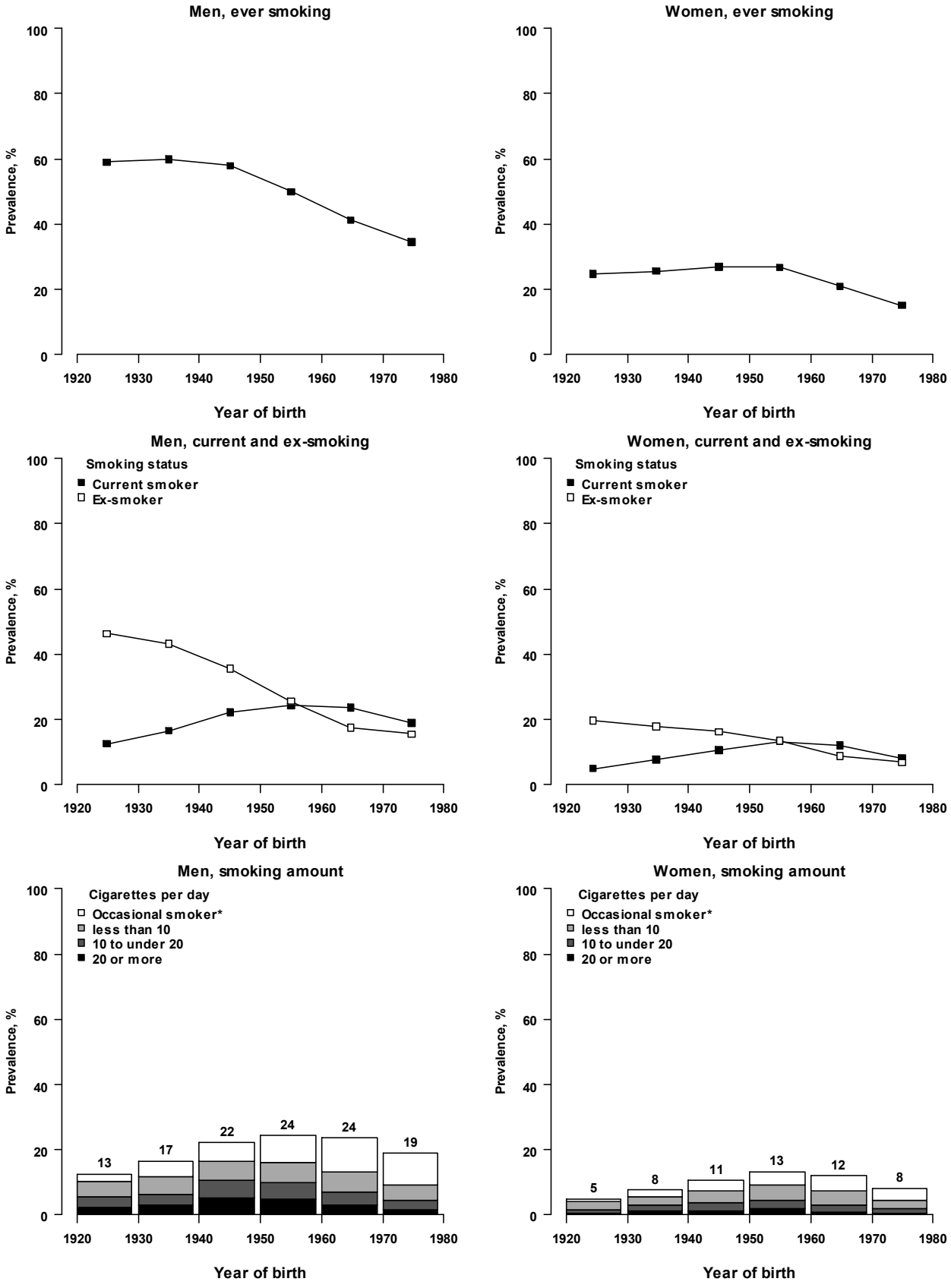
*Weighted estimate 95% CI 440,000-790,000

Table 5.7: Population attributable fraction (PAF) for ever versus never smoking at ages 30-89, by race and sex

	Male			Female		
Race/ethnicity	No. of deaths smokers/non-smokers	RR (95% CI)	PAF (%)	No. of deaths smokers/non-smokers	RR (95% CI)	PAF (%)
Mexican American	842/704	1.28 (1.15-1.42)	12	350/1155	1.21 (1.06-1.37)	4
Other Hispanic	644/417	1.52 (1.33-1.73)	21	354/680	1.62 (1.41-1.85)	13
Non-Hisp. White	12302/4037	1.69 (1.63-1.75)	31	9272/7823	1.94 (1.88-2.00)	26
Non-Hisp. Black	2135/1011	1.54 (1.43-1.67)	24	1690/2045	1.57 (1.46-1.68)	16
Other race	463/290	1.41 (1.20-1.66)	18	235/501	1.52 (1.28-1.82)	11

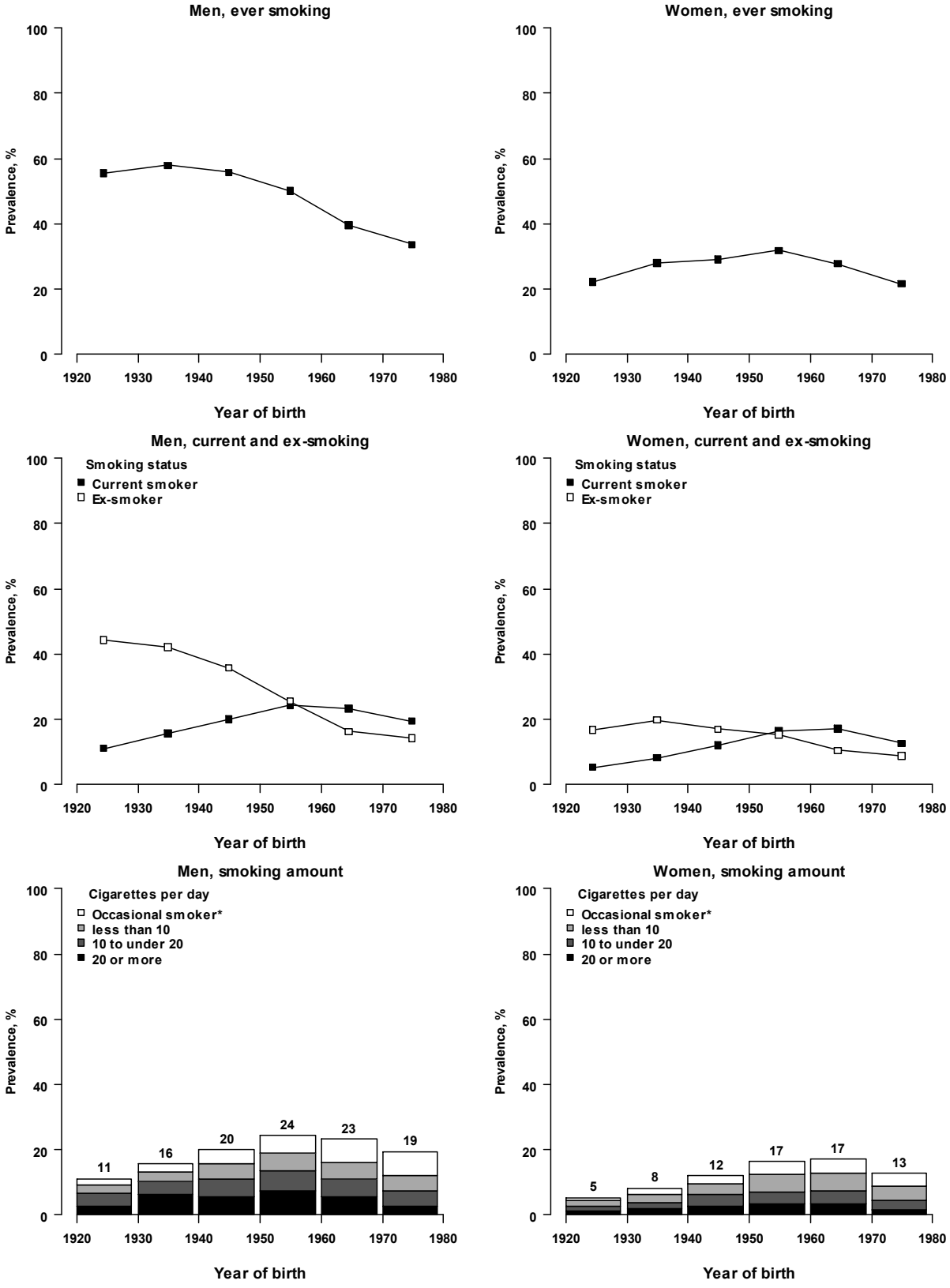
Calculated for all medical causes of death

Figure 5.1: Baseline smoking prevalence among Mexican Americans in the US NHIS



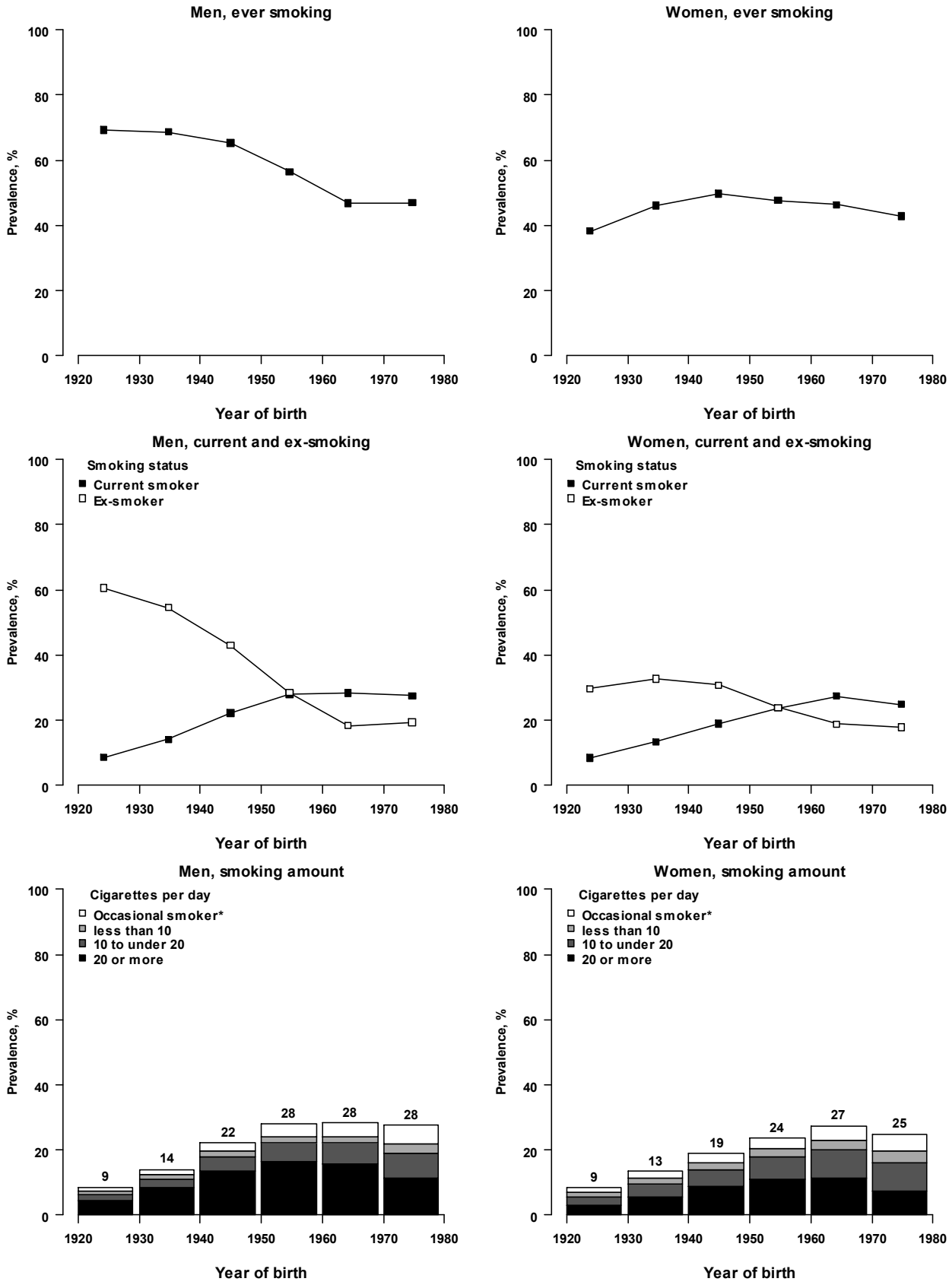
Number above bar in plots 5 and 6 is percent current smokers
*: Non-daily smokers

Figure 5.2: Baseline smoking prevalence among other Hispanics in the US NHIS



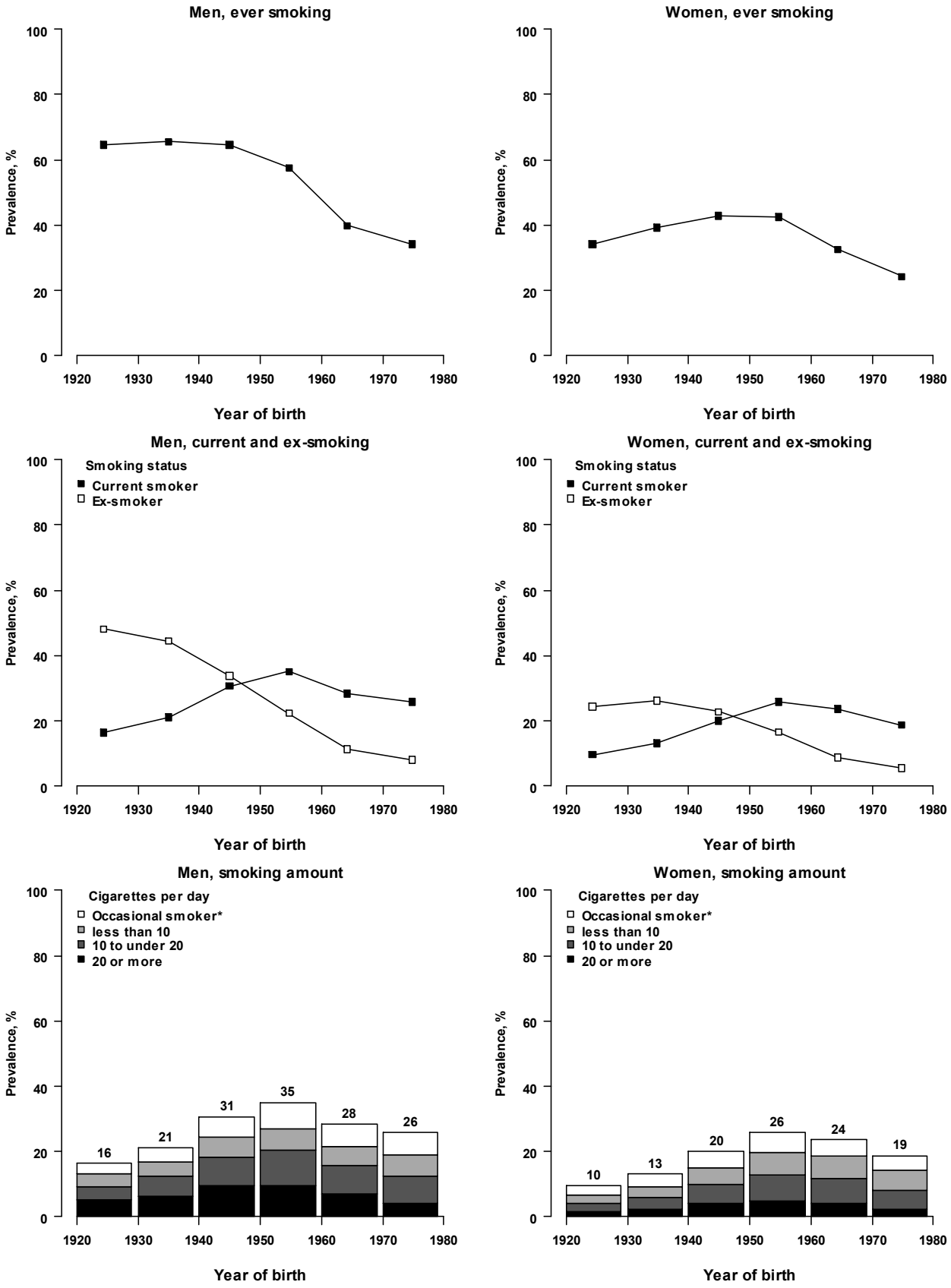
Number above bar in plots 5 and 6 is percent current smokers
*: Non-daily smokers

Figure 5.3: Baseline smoking prevalence among non-Hispanic Whites in the US NHIS



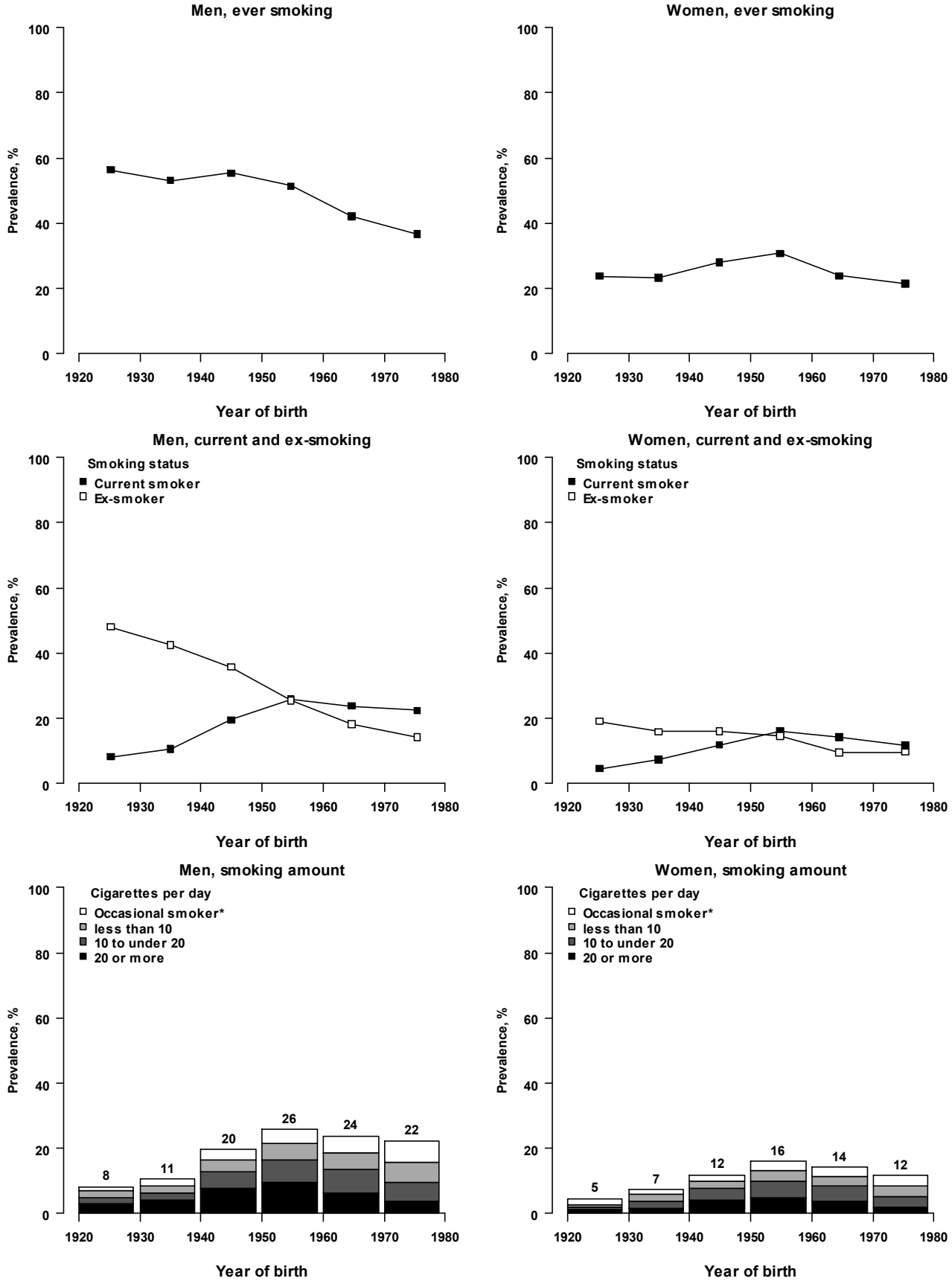
Number above bar in plots 5 and 6 is percent current smokers
*: Non-daily smokers

Figure 5.4: Baseline smoking prevalence among non-Hispanic Blacks in the US NHIS



Number above bar in plots 5 and 6 is percent current smokers
*: Non-daily smokers

Figure 5.5: Baseline smoking prevalence among other race/ethnicity in the US NHIS



Number above bar in plots 5 and 6 is percent current smokers
*: Non-daily smokers

Figure 5.6: Mean cigarettes/day and age started smoking by race and sex among 68 853 daily smokers in the US NHIS prospective study (1997-2014)

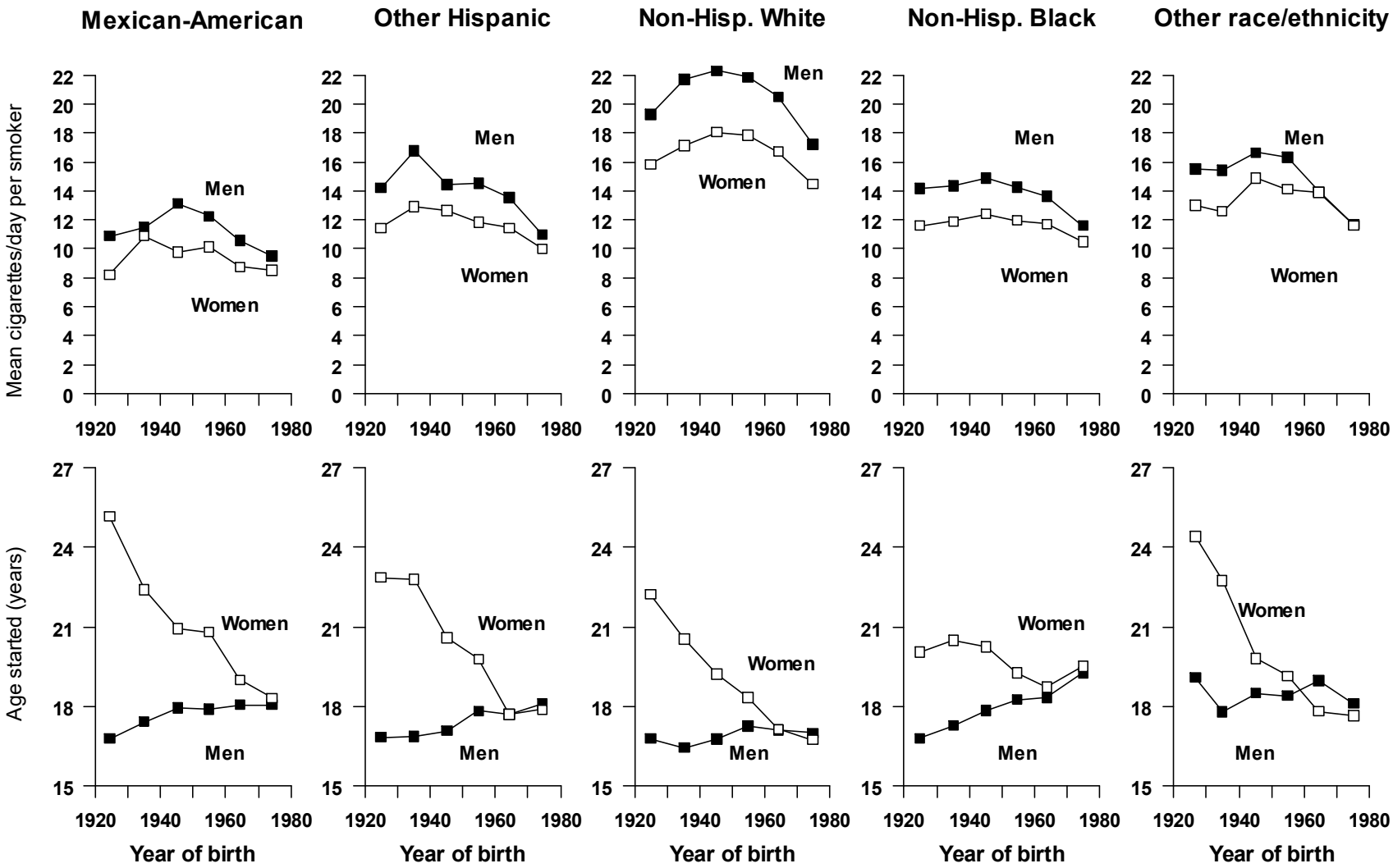


Figure 5.7: All-cause mortality rate ratio at ages 30-69, daily vs never smokers, by age started smoking

Analyses conducted among all participants (pooled across race), further adjusting for race. Analyses include those with illness at recruitment.

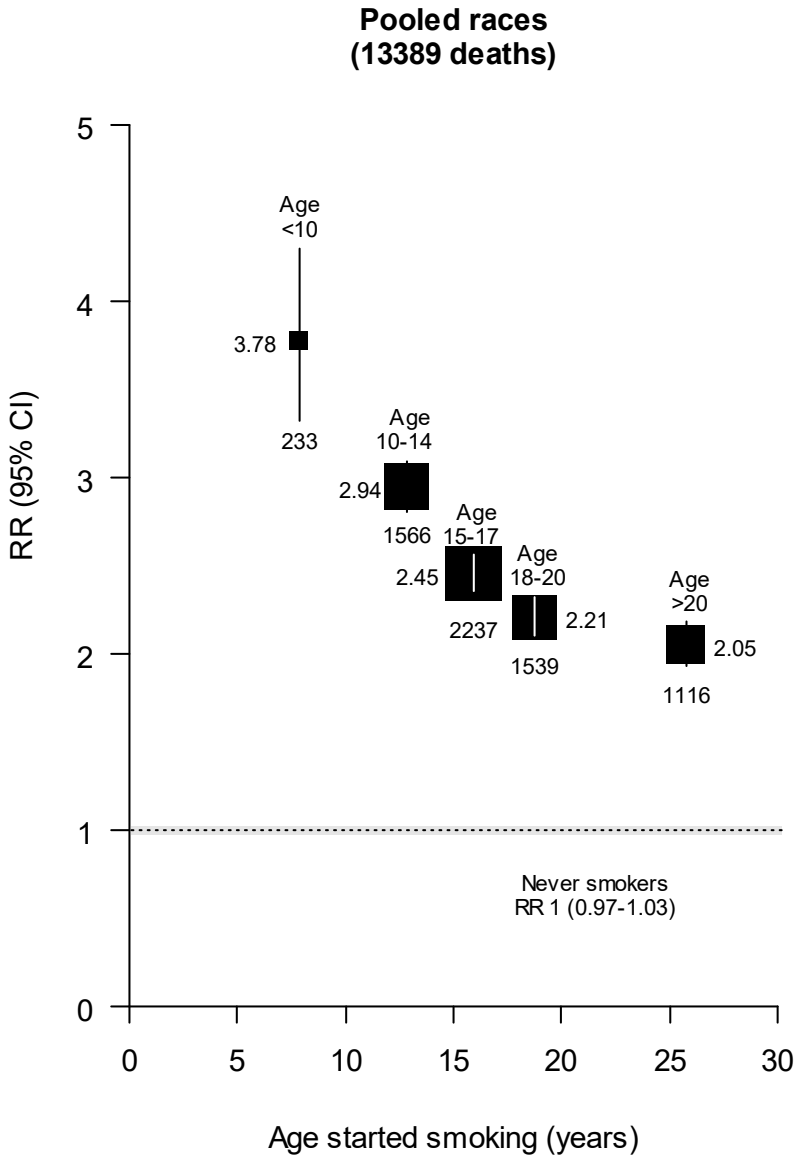


Figure 5.8: All-cause mortality rate ratio at ages 30-69, daily vs never smokers, by amount smoked

Analyses conducted among all participants (pooled across race), further adjusting for race. Analyses include those with illness at recruitment.

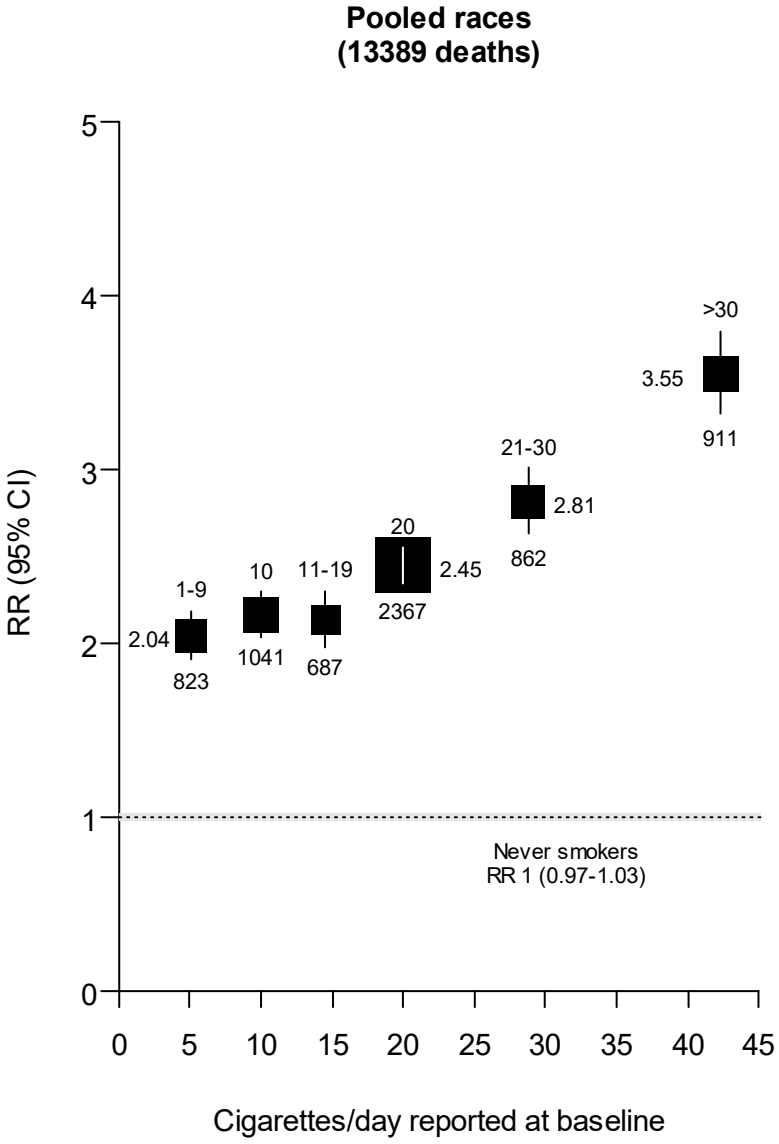


Figure 5.9: All-cause mortality rate ratio by age quit smoking among those quitting at ages 15-54, compared to never and current smokers at baseline, age at risk 55-89

Analyses conducted among all participants (pooled across race), further adjusting for race. Analyses exclude those with illness at recruitment.

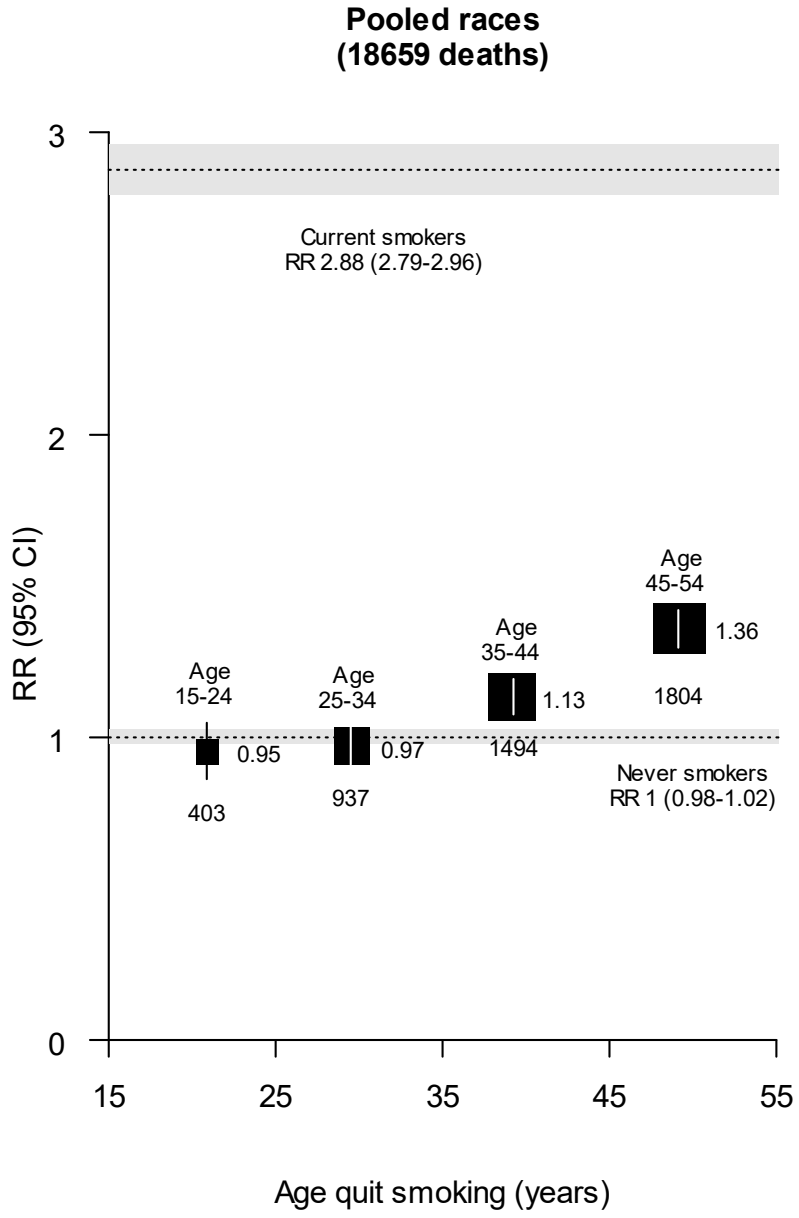


Figure 5.10: All-cause mortality rate ratio by years since quitting smoking, compared to never and current smokers at baseline, age at risk 30-89

Analyses conducted among all participants (pooled across race), further adjusting for race. Analyses exclude those with illness at recruitment.

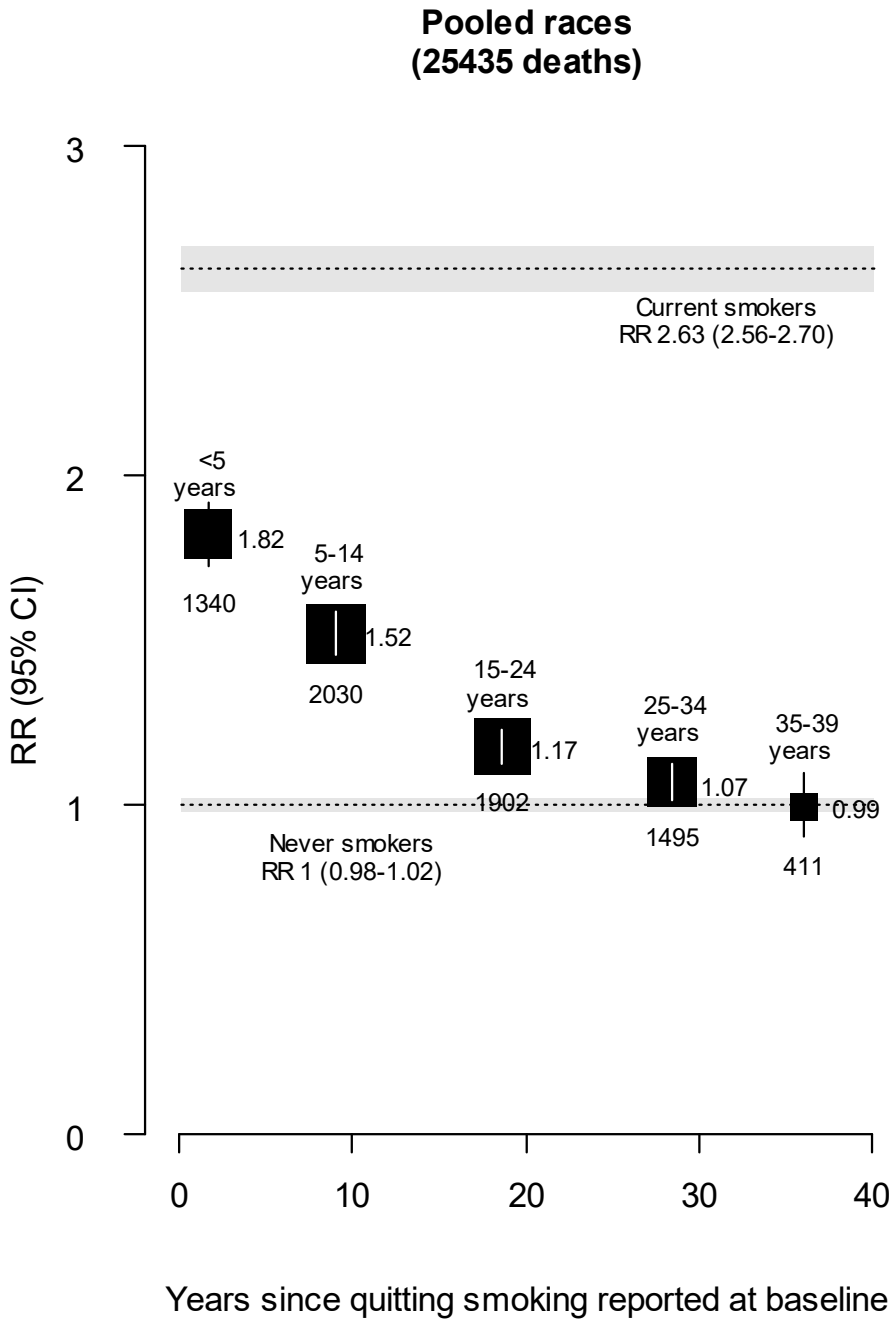


Figure 5.11: All-cause mortality rate ratio at ages 30-89, current daily smoker vs never smoker, by race

Participants with chronic disease at recruitment were included in these analyses.

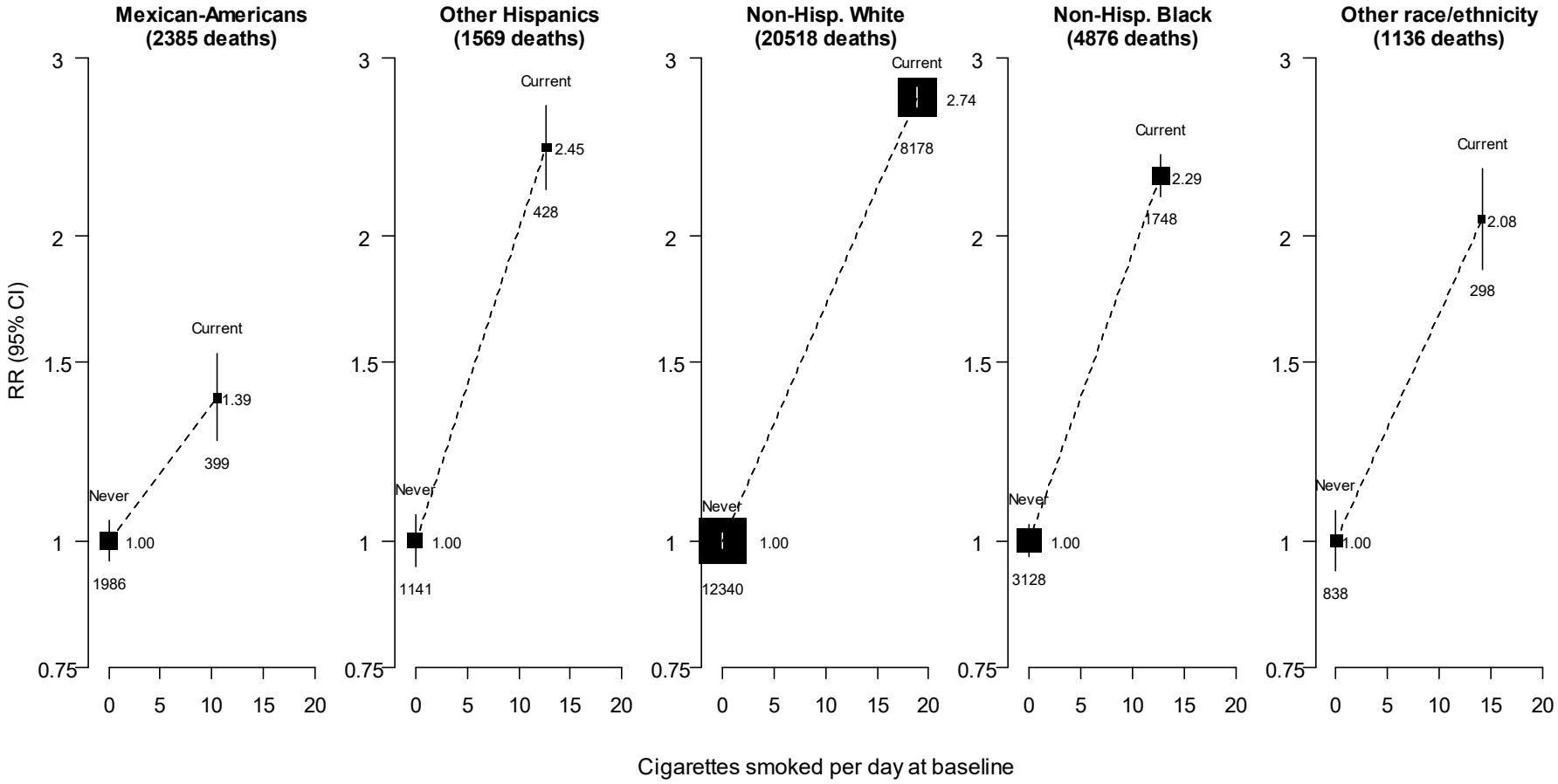


Figure 5.12: All-cause mortality rate ratio at ages 30-89, current daily smoker vs never smoker, by race

Participants with chronic disease at recruitment were excluded from these analyses.

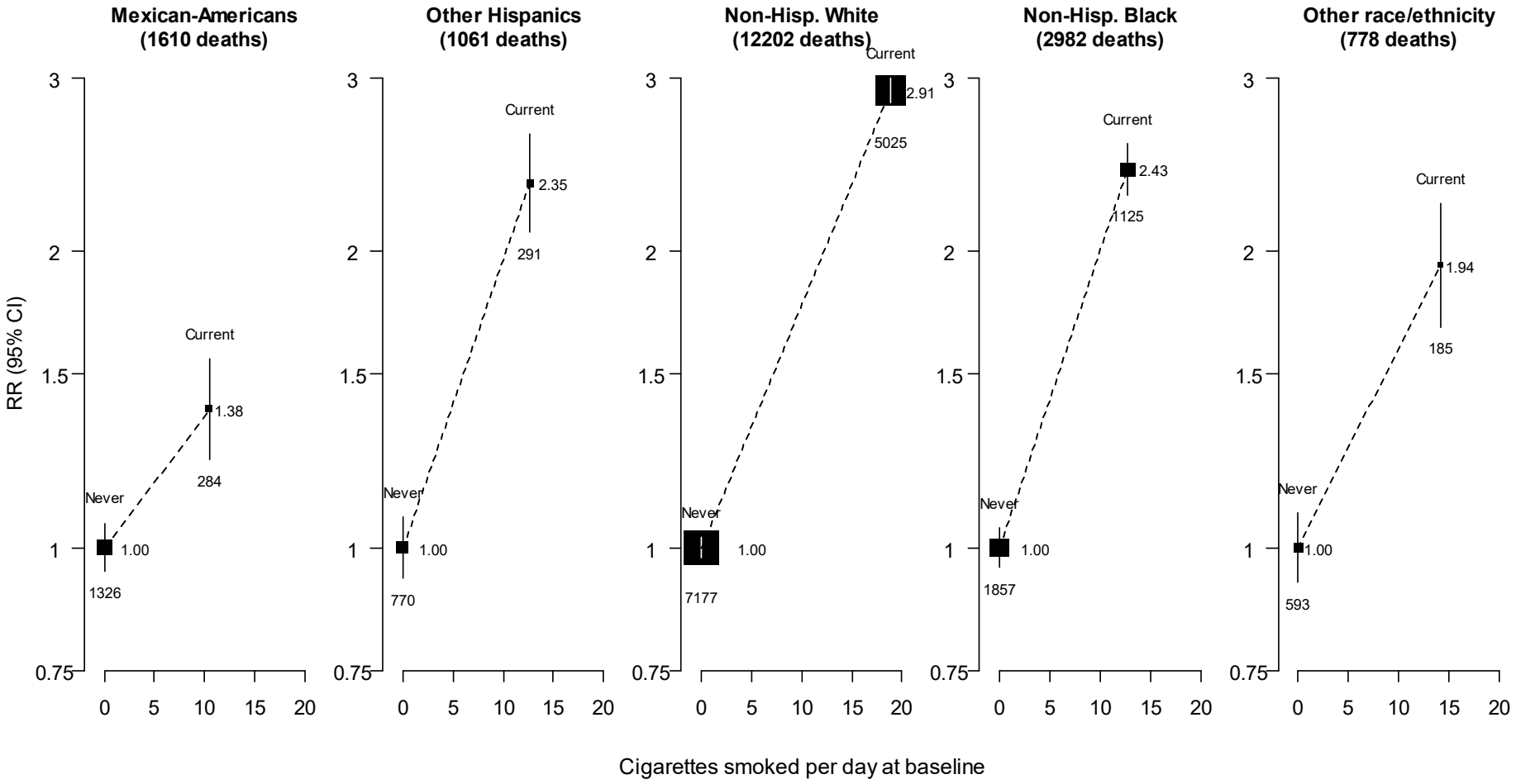


Figure 5.13: All-cause mortality rate ratio at ages 30-89, current daily smoker vs never smoker, by age started smoking and race

Participants with chronic disease at recruitment were included in these analyses.

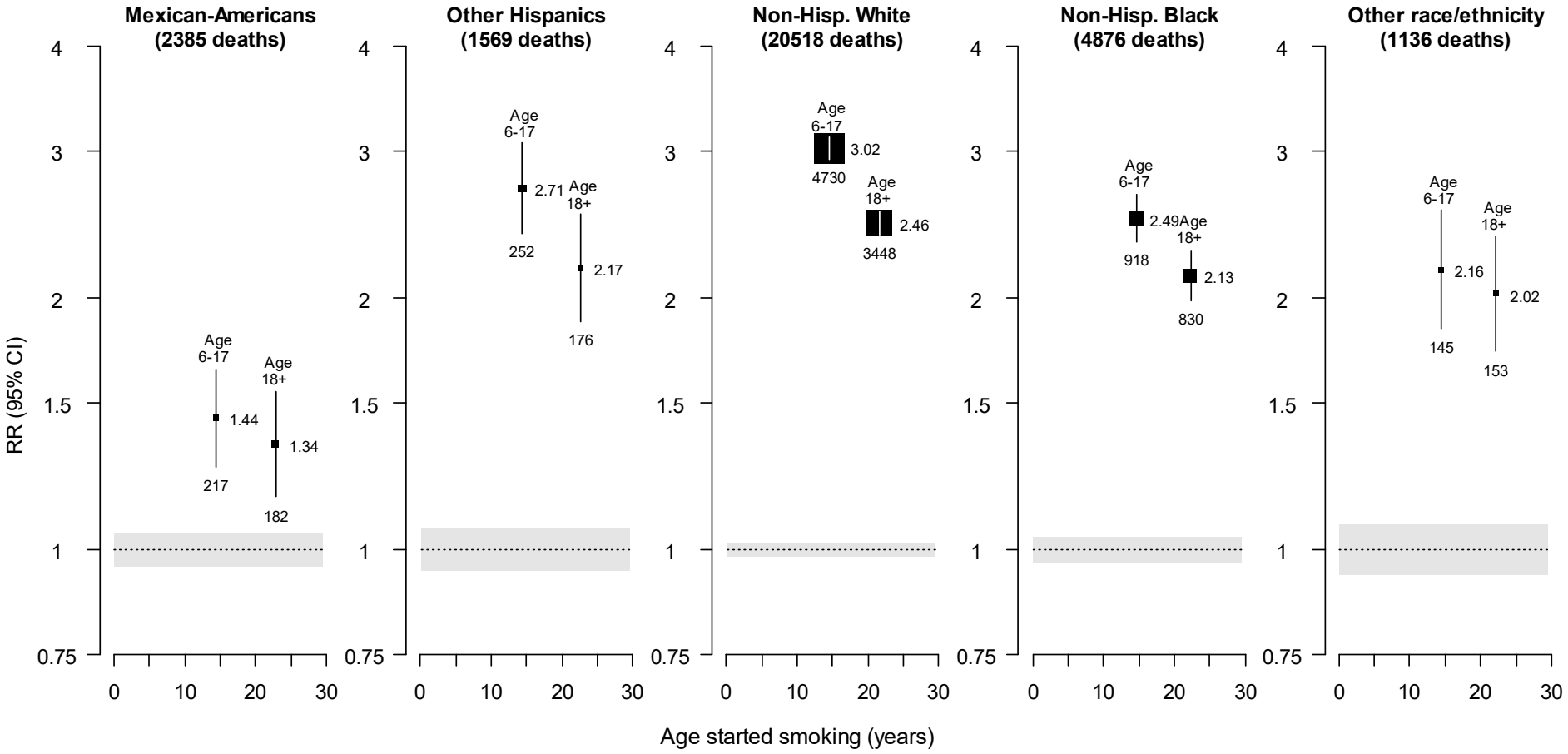


Figure 5.14: All-cause mortality rate ratio at ages 30-89, current daily smoker vs never smoker, by amount smoked and race

Participants with chronic disease at recruitment were included in these analyses.

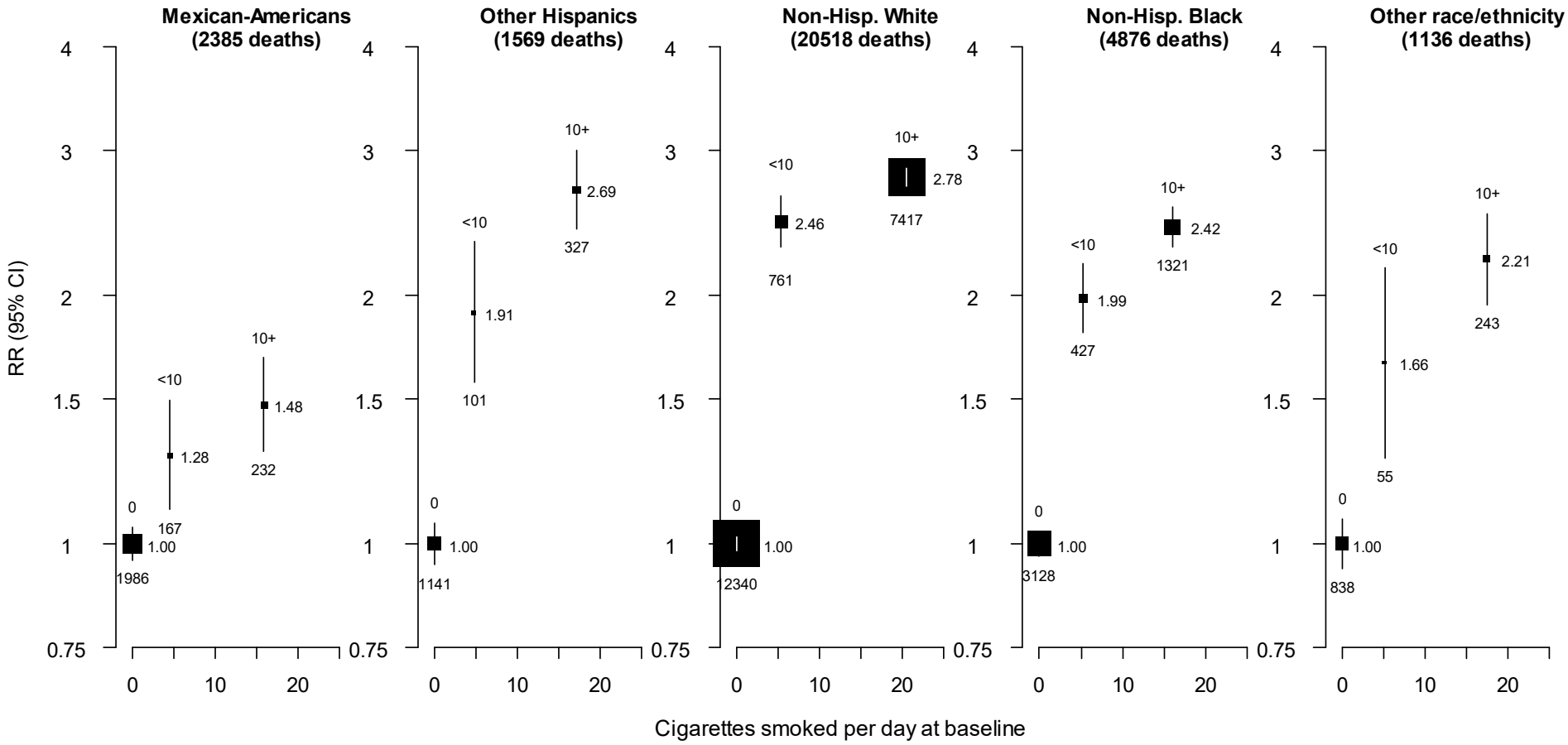


Figure 5.15: All-cause mortality rate ratio by years since quitting smoking and race, compared to never and current smokers at baseline, age at risk 30-89

Participants with chronic disease at recruitment were included in these analyses.

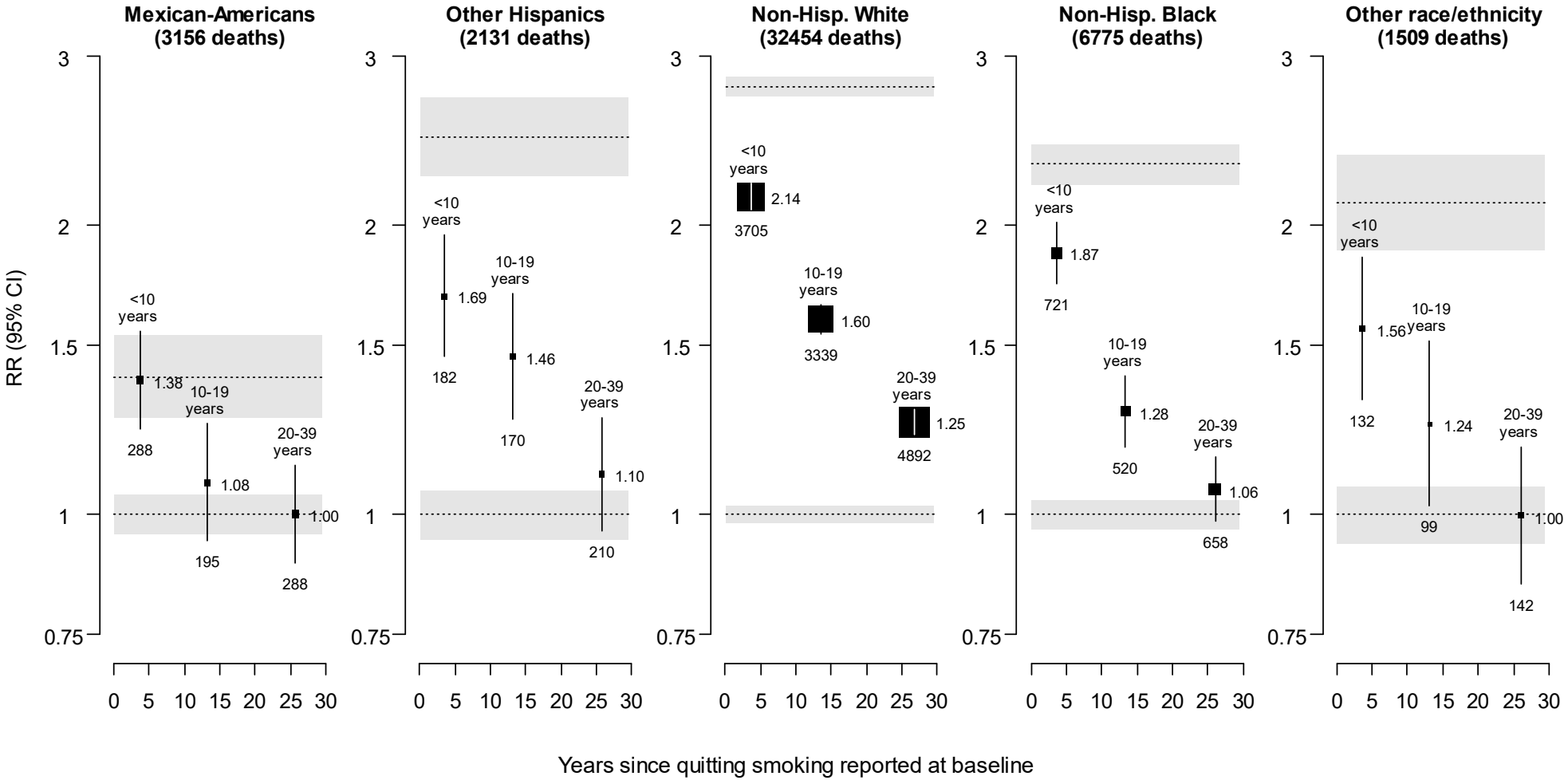


Figure 5.16: All-cause mortality rate ratio by age quit smoking and race among those quitting at ages 15-54, compared to never and current smokers at baseline, age at risk 55-89

Participants with chronic disease at recruitment were included in these analyses.

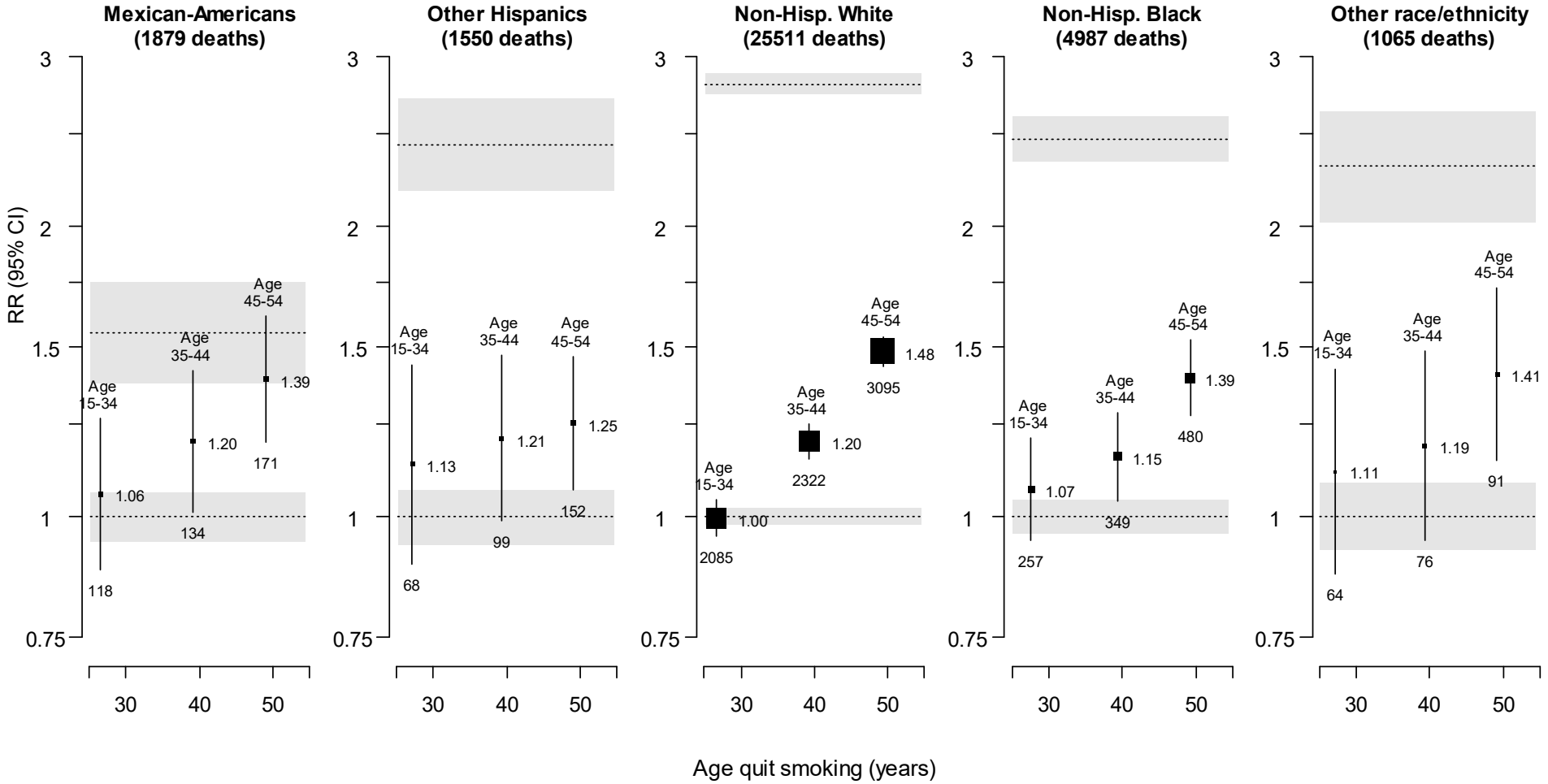


Figure 5.17: All-cause mortality rate ratio by years since quitting smoking and race, compared to never and current smokers at baseline, age at risk 30-89

Participants with chronic disease at recruitment were excluded from these analyses.

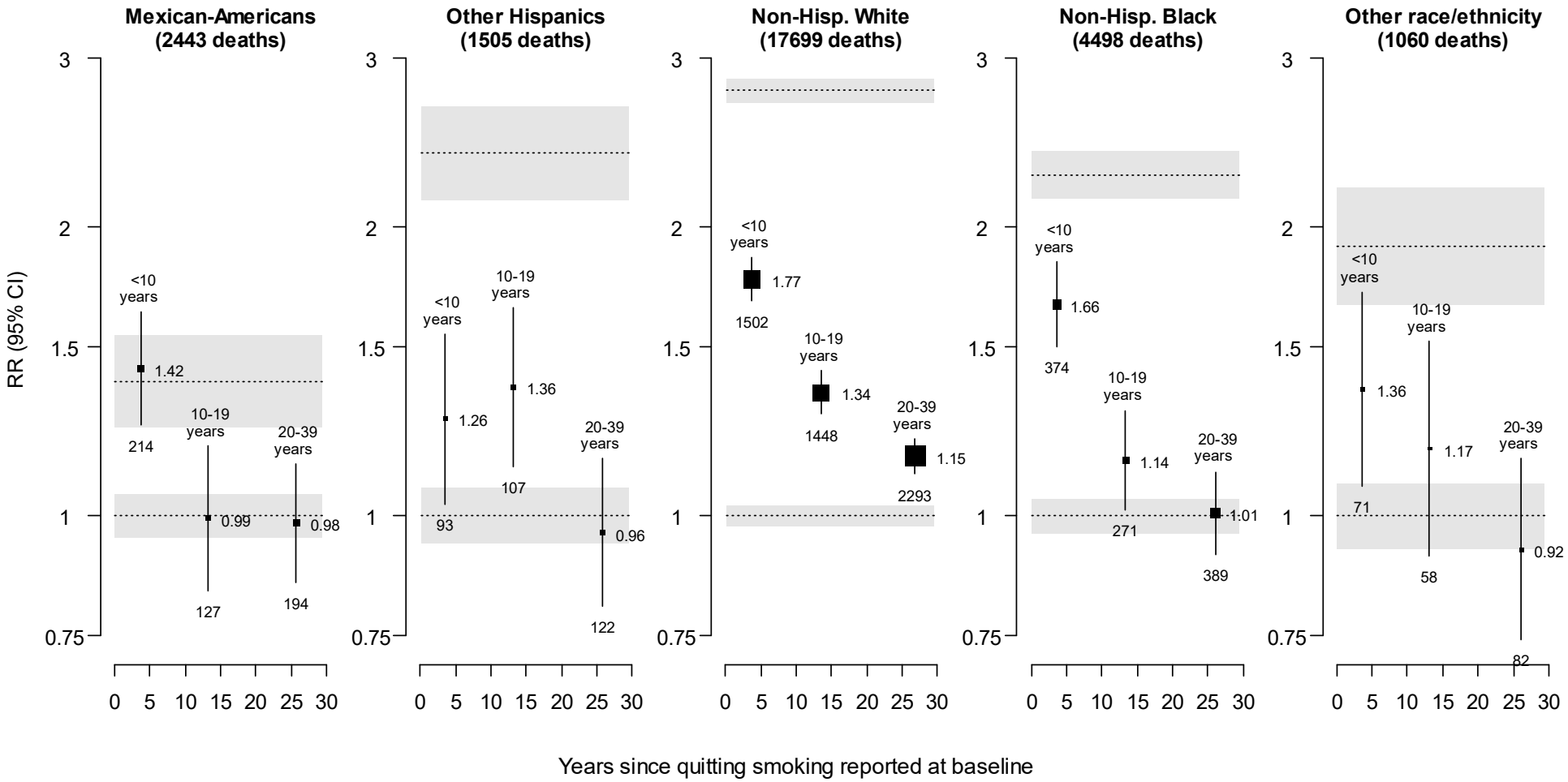
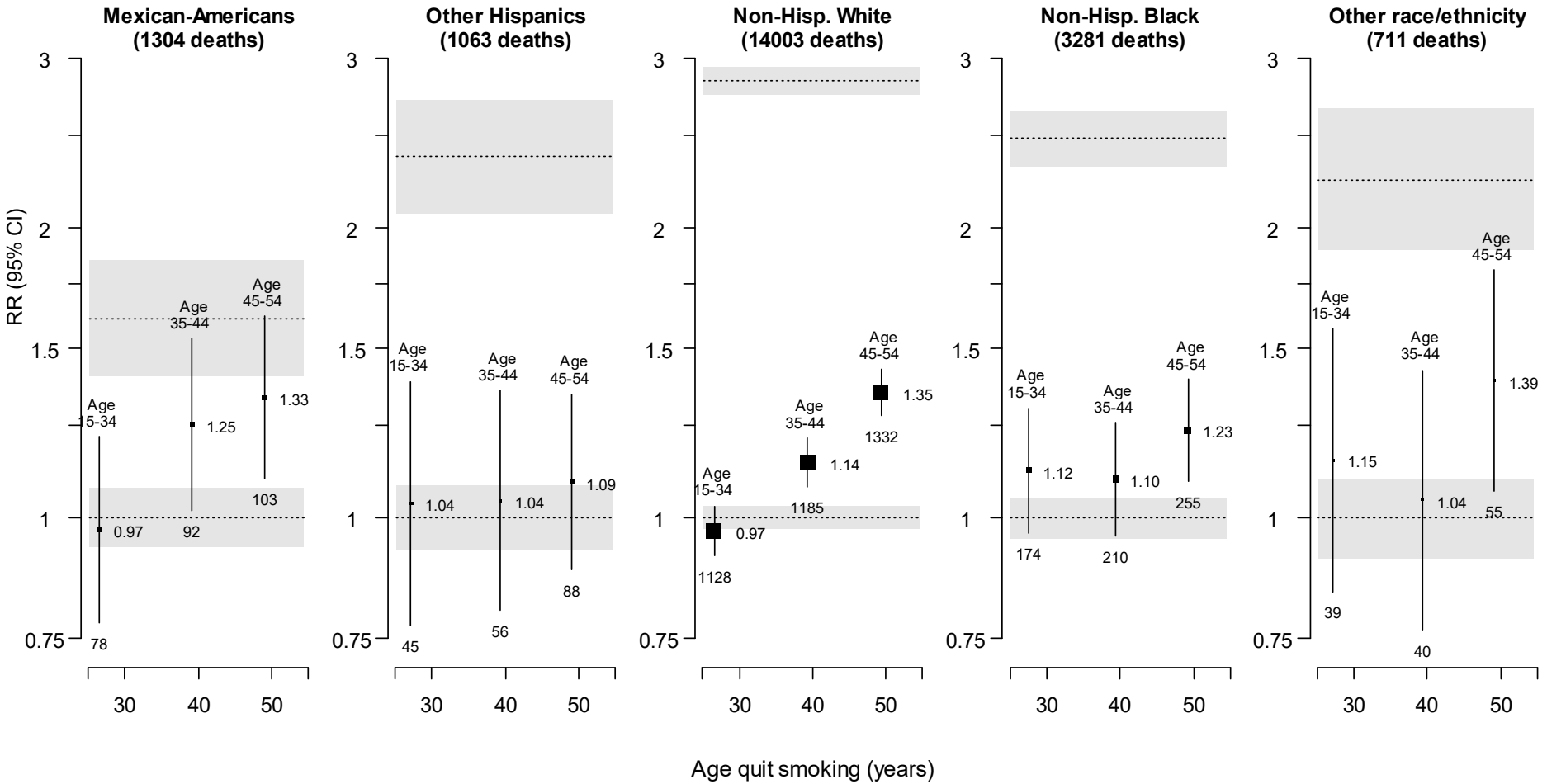


Figure 5.18: All-cause mortality rate ratio by age quit smoking and race among those quitting at ages 15-54, compared to never and current smokers at baseline, age at risk 55-89

Participants with chronic disease at recruitment were excluded from these analyses.



Chapter 6: Discussion

6.1 Thesis overview

A systematic review of the literature established that, at the time of writing, there have been only 5 prospective studies of smoking and cause-specific mortality in a Latin American population involving at least 1000 participants (and the largest of these involved <10 000 participants; **Chapter 2**). The main aims of this thesis, then, were to assess cross-sectional smoking patterns across birth cohorts, to characterise prospective associations between smoking (and smoking cessation) and mortality, and to estimate the population attributable fraction of smoking in relation to mortality in large prospective studies from Cuba, Mexico, and the United States. This was accomplished, first, in the Cuba Prospective Study of 140 000 adults (**Chapter 3**); second, in the Mexico City Prospective study of 150 000 adults (**Chapter 4**); and third, in the US National Health Interview Survey (NHIS) prospective study of 400 000 adults (**Chapter 5**). Key findings from each of these prospective studies were discussed in the context of the wider literature, with strengths and limitations of each study independently presented. This chapter summarises the key findings of these chapters and discusses their implications in the context of each other, and in a broader global context. The implications of these findings for other low- or middle-income populations are discussed alongside cross-sectional estimates from the Global Adult Tobacco Surveys in 30 countries around the world (including several from Latin America). Finally, the main strengths and limitations of the overall work presented in this thesis, along with areas for future research, are discussed.

6.2 Summary of main findings across studies

6.2.1 Trends in smoking across birth cohorts and sex

In the Cuban and Mexican cohorts, approximately half of men and one-quarter of women reported currently smoking at recruitment. However, in the Cuban cohort, only one in eight ever-smokers had quit smoking by recruitment, whereas in the Mexican cohort, approximately half of ever-smokers had quit by recruitment. At the time of resurvey in the respective cohorts, smoking cessation remained uncommon among surviving Cuban participants, but a significant proportion of Mexican participants who had been current smokers at recruitment had quit by the resurvey.

In Cuba, the smoking prevalence and amount smoked were similar among younger and older participants, and the average age at smoking initiation was slightly higher among younger male birth cohorts and slightly lower among younger female birth cohorts. Overall, approximately one-third of current smokers at recruitment began smoking before the age of 15, and 4% began before the age of 10. Smokers predominantly smoked cigarettes, though a small number of (mostly older, male) smokers exclusively smoked cigars.

In Mexico, current smoking prevalence was positively associated with year of birth (ie, was more common among younger participants), while smoking cessation was inversely associated with year of birth (ie, was less common among younger participants). Non-daily smoking was more common among younger participants, and heavy smoking (ie, ≥10 cigarettes/day) was relatively uncommon among all participants, regardless of age or sex. Overall, one-third of current smokers smoked less than daily, and the average number of cigarettes smoked/day at recruitment was 9 among men (SD 8) and 6 among women (SD 6).

In the US cohort, smoking patterns varied by race. However, across all races, ever smoking prevalence was lower among more recent birth cohorts of men, but for women ever smoking prevalence was generally similar across birth cohorts. Also consistent across race, smoking cessation was inversely associated with year of birth, with the large majority of ever-smokers born before 1940 having quit by the time of recruitment among both men and women. As observed in the Mexican cohort, Mexican-American current smokers smoked relatively few cigarettes/day (mean 11 cigarettes/day, SD 8), and non-daily smoking was common (particularly among younger participants). In contrast, among non-Hispanic white smokers, non-daily smoking was uncommon, and on average male smokers smoked nearly a pack of cigarettes/day (mean 19 cigarettes/day, SD 10). Across races, among men, the average age at smoking initiation was similar (or slightly higher) among more recent birth cohorts, while among women, the average age at smoking initiation was lower among more recent birth cohorts.

6.2.2 Prospective evidence on the hazards of smoking

Overall, the mortality risk associated with smoking a given number of cigarettes/day was similar between the Cuban and Mexican cohorts (although the current versus never relative risks were lower among Mexicans, given their lower amount smoked). Among Cuban participants, smoking an average of ~10 cigarettes/day during follow-up compared with never smoking was associated with a 39% higher in risk of death at ages 30-69. Among Mexican participants without chronic disease at recruitment, smoking ~10 cigarettes/day compared with never smoking was associated with a 52% higher in risk of death at ages 35-69.

These risks were similar to those of Mexican Americans in the United States: smoking ~10 cigarettes/day *at recruitment* (amount smoked over follow-up was not known) compared with never smoking was associated with a 39% increase in risk of death at

ages 30-89. However, these risks were substantially lower than the risks of current smoking observed among other Hispanics, non-Hispanic Whites, non-Hispanic Blacks, and other races in the United States, which ranged from an increase in all-cause mortality of 108% among other race/ethnicity to 174% among non-Hispanic Whites (though these differences were likely driven in part by higher cigarette consumption at recruitment). Consequently, the estimated proportion of deaths attributable to smoking in these populations ranged from 4% among Mexican and Mexican-American women to 31% among non-Hispanic White men in the US (summarised in Table 6.1).

Although the relative risks were lower among Cubans, Mexicans, and Mexican Americans compared with other US races/ethnicities, similar dose-response patterns were observed with amount smoked and age started smoking within a given population. For example, although the relative risks for current versus never smoking were much lower among Cubans than among US participants, among smokers, the excess mortality risk at ages 30-69 associated with smoking initiation before age 15 years was approximately twice that of smoking initiation after age 15 in both cohorts. Similarly, among smokers, the excess mortality risk associated with smoking increased with increasing number of cigarettes smoked/day in the different populations.

Smoking behaviour was strikingly similar among Mexicans and Mexican Americans, and led to similarly low mortality risk associated with smoking (compared to never smokers, daily smoking was associated with an all-cause mortality increase of 28% among Mexicans and 39% among Mexican Americans). The PAF estimates of smoking-attributable mortality were similar among Mexican and Mexican-American men (10% and 12%, respectively) and women (4% in each population), which was due in part to the low average consumption in these populations (approximately 8 cigarettes/day in Mexicans and 10 cigarettes/day in Mexican Americans, compared with approximately 19

cigarettes/day in non-Hispanic Whites). Furthermore, it remains unclear whether those who smoked at recruitment were truly lifelong smokers at recruitment, or whether a substantial number of smokers smoked intermittently (which would further reduce the association between smoking at recruitment and mortality). Further research is needed to elucidate the potential relevance of acculturation of Mexican Americans to smoking, particularly given the appreciably different smoking patterns (and mortality risk) observed in Mexican Americans compared with other Hispanic populations in the United States.

There was no strong evidence that the relationship between current daily smoking and all-cause mortality varied over the follow-up period in these prospective studies (though there was some evidence that the RRs may have increased at greater follow-up periods in the Cuban cohort). However, it is possible that these stable RR estimates mask diverging trends, given that greater duration of smoking among continuing smokers may be expected to increase the mortality RR over the follow-up period, but more prevalent smoking cessation over the follow-up period among those who smoked at recruitment may be expected to decrease the mortality RR.

6.2.3 Prospective evidence on the benefits of quitting

Across populations and races, there was clear evidence that smoking cessation was associated with substantially lower mortality risk than continued smoking, and that both age at smoking cessation and duration of cessation were associated with lower mortality risk in a dose-dependent manner. Among Cubans and US participants of all races, quitting by age ~40 was associated with little of the excess risk of death associated with current smoking at recruitment (with even lower risk observed among those who quit at even younger ages). Similarly, among those who had quit for at least 15 years prior to recruitment (in Cuba), or 20 years prior (in the US), there was little to no excess risk of mortality compared with those who had never smoked. Although information on smoking

cessation was recorded less precisely in the Mexican cohort, substantial reductions in mortality risk were observed among those who had quit smoking in comparison to those who were current smokers at recruitment. Among Mexicans without chronic disease at recruitment, smoking cessation >3 years prior to recruitment was associated with little of the excess mortality risk associated with current smoking, with particularly large reductions in mortality risk observed for several smoking-related causes of death (including COPD and respiratory cancers).

6.3 Consistency with published literature

The relevance of current smoking to mortality among non-Hispanic White men and women observed in the US NHIS population (RR for current versus never smokers: 2.74) was similar to that of previously-reported evidence from pooled contemporary US cohorts (RR for current versus never smokers: 2.76 among women and 2.80 among men)[8]. However, the association between smoking and mortality among the Mexican, Mexican American, and Cuban populations studied was appreciably weaker. While smoking patterns varied across each of these cohorts, for a given pattern (eg, smoking <10 cigarettes/day), the mortality RRs for current versus never smoking were generally similar among Cubans, Mexicans, and Mexican Americans. However, the current versus never smoker RRs were, in some cases, substantially lower than those seen in high-income populations previously studied (eg, the US overall as reported here and elsewhere,[4] UK,[35] and Australia[5]), but were generally similar to those reported in studies in non-Western populations (eg, China,[9] India,[62] Bangladesh,[63] Thailand,[126] Iran,[127] and South Africa[128]). Nevertheless, the observation that smoking cessation before age ~40 was associated with little of the excess risk associated with continued smoking was similar to findings from both high-income and lower- and middle-income countries (where studies with a sufficiently large number of quitters have been conducted). However, reliable, direct estimates of the association between smoking (and smoking

cessation) and mortality have not yet been established in most non-Western populations. Given that indirect estimates of smoking-attributable mortality from the GBD Study are similar to those directly estimated in Cuban and Mexican populations in this thesis, the GBD Study may provide the best estimate of the population impact of smoking on mortality in countries in which direct evidence has not yet been produced.

Prior evidence on the prevalence of behaviours such as smoking regularly from childhood or relatively low levels of cigarette consumption (eg, <10 cigarettes/day) in diverse populations globally is currently limited. As a result, it is difficult to place the smoking patterns observed in these populations into broader context. Therefore, to determine the potential relevance of these findings to other populations, with a focus on non-Western populations, analyses were carried out using the Global Adult Tobacco Surveys (GATS). These nationally representative, standardised household surveys have been conducted in 30 (predominantly) low- and middle-income countries at least once in 2008-2016. While findings from each of these national cross-sectional surveys have previously been published as country-level reports, and a number of peer-reviewed publications have reported findings from these underlying data,[106,129,130] these data have not been used to systematically report a number of indicators of interest in relation to this thesis (eg, the number of adult daily smokers who began smoking age<10). Findings of primary interest to this thesis are described in the following section, and a more detailed description of the GATS surveys, as well as the methodology used for analyses, can also be found in Appendix D. These cross-sectional estimates of smoking patterns serve primarily to underscore the potential relevance of the findings from Cuban and Mexican populations reported in this thesis, particularly in terms of early age at smoking initiation and low levels of cigarette consumption, to other diverse populations globally.

6.4 Findings in global context: Evidence from the cross-sectional Global Adult Tobacco Surveys in 30 low- and middle-income countries

Among the 30 countries in which GATS surveys were conducted, there were 366 179 participants representing an estimated 3.3 billion adults, of whom an estimated 600 million smoked daily (and a further 116 million smoked less than daily) at the time of survey (Table 6.2). Overall, daily smoking prevalence ranged from 3.2% in Ethiopia to 36.6% in Greece, and the mean number of (manufactured or hand-rolled) cigarettes smoked/day among smokers ranged from 6 cigarettes/day in India and Uganda to 19 cigarettes/day in Egypt. Among men, daily smoking prevalence ranged from 4.4% in Panama to 57% in Indonesia (Figure 6.1). Among women, smoking prevalence ranged from 0.3% in Nigeria (and <1% in Senegal, Egypt, Cameroon, Kenya, Malaysia, and Vietnam) to 24% in Greece (Figure 6.2).

Among current cigarette smokers, non-daily smoking was generally uncommon, with two major exceptions: Panama and Mexico. Among men, non-daily smokers made up the majority of current smokers in Panama and Mexico, and more than one-third of current smokers in Costa Rica, but comprised fewer than one-quarter of current smokers in the remaining 27 countries (Figure 6.3). Non-daily smoking was somewhat more common among female smokers, but most smokers nevertheless smoked daily. Among women, non-daily smokers outnumbered daily smokers only in Mexico and Panama, and nearly half of current smokers smoked less than daily in Qatar (Figure 6.4). However, in half of countries, fewer than one in five female smokers smoked less than daily.

There was a wide range of daily smoking consumption behaviours between countries. Among male daily smokers, ‘light’ smoking prevalence (<10 cigarettes/day) ranged from <10% of smokers in Greece (4%), Egypt (6%), Ukraine (8%), and Turkey (9%), to more than two-thirds of smokers in Mexico (67%), Nigeria (67%), Uganda (74%), and India (77%; Figure 6.5). Among female daily smokers, ‘light’ smoking prevalence ranged from

<20% of smokers in Poland (13%), Greece (13%), and Romania (19%), to more than 70% in Cameroon (72%), Mexico (72%), India (72%), Indonesia (73%), Bangladesh (90%), Uganda (91%), and Kenya (96%; Figure 6.6). In the Mexico City Prospective Study, 63% of male daily smokers and 77% of female daily smokers smoked <10 cigarettes/day, which was common in several other countries (though on the lower end of consumption among these 30 countries). Therefore, the finding that even 'light' smoking was associated with increased mortality in this population may be relevant to smokers in these and other populations globally. The level of cigarette consumption was relatively low in several populations, and may lead to lower smoker versus never smoker RRs (as in the Mexican and Mexican American populations). This finding underscores the importance of reporting specific smoking patterns beyond broad groups such as 'current', 'ex', and 'never' smokers if RRs are to be understood in their proper context.

The prevalence of smoking initiation in childhood (age <15) also varied substantially between countries. Among male daily smokers, the prevalence of smoking initiation at ages 5-14 was highest in the participating Latin American countries: 31% in Costa Rica, 26% in Mexico, 25% in Argentina, 24% in Brazil, 23% in Uruguay, and 19% in Panama (Figure 6.7). In comparison, the prevalence of smoking from childhood among men was lowest in China (3%), Qatar (5%), Greece (6%), and Kazakhstan (6%). However, smoking initiation in early childhood (ie, before age 10) was uncommon among male daily smokers, with prevalence ranging from 3% (in Uganda, Ukraine, Ethiopia, Costa Rica, Brazil, and Russia), to <1% in Qatar, Greece, Argentina, China, and Malaysia. (Smoking initiation before age 15 among women is not described here due to limited precision of estimates, but is shown in Figure 6.8. Trends in childhood smoking initiation among men and women are not shown for similar reasons.) The prevalence of childhood smoking initiation in the Cuba Prospective Study (34% at ages 5-14; 4% at ages 5-9) is only slightly higher than the prevalence of smoking from childhood in several of these 30

LMICs, suggesting that the hazards of smoking from childhood may have important implications for a number of countries globally (particularly several other Latin American countries). Furthermore, the finding that smoking from a young age was common in many of these countries (particularly age<15), which would be expected to lead to substantially elevated risk of mortality in adulthood, suggests that policies aimed at reducing smoking initiation before adulthood may be particularly important in reducing smoking attributable mortality at the population level in many countries.

Overall, an estimated 52 million adult smokers in these 30 LMICs began smoking regularly before age 15, of whom an estimated 6 million began smoking regularly before age 10 (Figures 6.9 and 6.10). In each of the following six countries, at least an estimated 3 million adult smokers had begun smoking regularly at ages 10-14: India (7.1 million), China (6.1 million), Indonesia (5.7 million), Russia (4.7 million), Brazil (4.3 million), and Bangladesh (3.0 million). In each of the following five countries, at least an estimated 0.25 million adult smokers had begun smoking regularly at ages 5-9: China (1.4 million), India (1.1 million), Russia (0.7 million), Brazil (0.6 million), and Indonesia (0.3 million). Therefore, if the excess risk associated with beginning to smoke regularly at young ages (ie, age<15, as found in Cuba and the US) is generalisable to other populations globally, these 52 million adult smokers (and likely many in other countries) will be at particularly high excess risk of premature mortality if they do not quit smoking.

Among the 30 countries in which the GATS surveys have been conducted, smoking cessation was not yet common in most countries. Within countries, rates of smoking cessation were generally similar among men and women (Figures 6.11 and 6.12), though cessation was somewhat more common among women overall. Among ever-smoking men, fewer than one in five ever-regular smokers had quit by the time of survey in eight countries: Pakistan (13%), Malaysia (14%), Indonesia (14%), China (16%),

Kazakhstan (17%), India (17%), Egypt (18%), and Bangladesh (20%; Figure 6.11). However, the average age of GATS participants was relatively young (39 years overall, ranging from 31 in Ethiopia to 46 in Romania), and cessation may be higher at older ages. Of the seven countries with the highest proportion of male ever-smokers who had quit, six were in Latin America (Costa Rica [63%], Senegal [56%], Brazil [50%], Panama [48%], Argentina [48%], Uruguay [48%], and Mexico [47%]). Notably, time to first smoking, a strong predictor of nicotine dependence[131] and ability to quit smoking[132], was particularly long (ie, smoking within 30 minutes after waking was uncommon) among both Mexican men and women (Figures 6.13 and 6.14). The relevance of tobacco control policies, cultural aspects of smoking, and genetic drivers of nicotine addiction to these unique smoking behaviours in Mexico merits further research.

These results indicate that the proportion of ever-smokers who had quit in the Cuba Prospective Study (1 in 8) and in the Mexico City Prospective Study (1 in 2) occupy both ends of the quitting spectrum in these LMICs. The relatively high rates of quitting among some Latin American countries may be driven by strong tobacco control policies, while the low rates of quitting in many other countries, particularly those in Asia, may be due to comparatively weaker tobacco control policies (and enforcement of those policies). Unlike preventing future generations from smoking (the population health benefits of which will not be observed for several decades), smoking cessation has the potential to reduce premature mortality in populations almost immediately and into the next several decades. Until these countries reach (or exceed) the rates of cessation among ever smokers seen in high-income countries like the US (59% in 2016)[133], the UK (61% in 2018)[134], or Australia (61% in 2016)[135], the threat of smoking to the health of these populations will continue to grow in the coming years.

6.5 Strengths and limitations of the thesis

6.5.1 Strengths

The main strengths and limitations of this research have been presented in the discussion section of each relevant chapter, but for the purposes of comparison are summarised here. A central strength of this thesis is the size of the studies included, which allowed for direct, reliable estimates of the relationship between smoking and mortality in Cuba (>140 000 participants), Mexico (>150 000 participants), and the United States (>400 000 participants). The size of these studies allowed for reliable estimates not just overall, but by factors like age, sex, and race, identifying potential divergences in these associations in major sub-populations. These are the first reliable, direct estimates of the relationship between smoking and mortality in Latin American populations, and are among the few large studies globally to assess 'usual' smoking behaviours (done in the Cuban and Mexican cohorts). As these studies were conducted prospectively, before onset of disease, these findings were not subject to recall bias, and the potential effects of reverse causality could be tested through sensitivity analyses. Each of these studies was conducted in a population with comprehensive vital registration systems, where nearly all deaths are medically certified. As a result, this thesis presented findings of relevance to the populations of Cuba, Mexico, and the US, as well as several novel findings with potentially global implications.

In **Chapter 3**, I presented the first reliable evidence worldwide of the association between smoking regularly from childhood (ie, age<10) and premature mortality in adulthood. In **Chapter 4**, I contributed to the limited existing evidence on the relevance of low levels of cigarette consumption to mortality, presenting evidence that regularly smoking even a few cigarettes/day was associated with elevated mortality risk. In **Chapter 5**, I presented evidence that smoking regularly from childhood had also occurred in the United States, and was associated with a similar pattern of excess risk

as that observed in the Cuban cohort. These findings are especially relevant to the estimated 5 million US adults who began smoking regularly before age 15, of whom an estimated 0.6 million began before age 10. Furthermore, I presented evidence that Mexican Americans had smoking patterns similar to those of Mexicans, and had subsequently lower risk of mortality than US smokers of other races (although even 'light' smoking was associated with excess risk of death). Finally, in this chapter (**Chapter 6**), I presented evidence from the Global Adult Tobacco Surveys in 30 low- and middle-income countries highlighting the potential relevance of the findings from this thesis to other populations globally (and particularly the estimated >50 million smokers in these 30 countries who began smoking before age 15, of whom >6 million began smoking before age 10).

6.5.2 Limitations

Although study procedures and analyses were standardised to the extent possible, factors that may limit the comparisons between these studies nevertheless persist. For example, the smoking-mortality associations observed in those aged ≥70 in Cuba were implausible, as evidenced by the lack of a dose-response relationship between amount smoked and mortality risk, and little excess mortality risk associated with smoking despite similar age at smoking initiation and amount smoked compared with younger participants. Similarly, the high prevalence (and poor control) of diabetes in Mexico may limit comparisons of smoking and all-cause mortality in this population to those of Cuba and the US, as it resulted in a number of deaths due to causes unrelated to smoking (and may have impacted on smoking and other behaviours). In the US, despite the large sample size, race-specific estimates were insufficiently precise at ages 30-69 (ie, 'premature mortality'), and therefore analyses were extended to ages 30-89. US cause-specific analyses were further limited by pre-specified, broad causes of death, limiting

the potential for comparison of cause-specific mortality by smoking patterns of interest (such as amount smoked or age started smoking).

Furthermore, there are potential limitations to the generalisability of these findings. The Cuban and Mexican cohorts were not designed to be nationally representative, and a population sampled solely from two urban districts of Mexico City in particular may differ in a number of ways from the broader Mexican population.[111] What's more, those who volunteer to participate in prospective studies are often healthier than the general population, further limiting the extent to which these populations are representative.[136,137] However, as findings from the British Doctors Study and other research efforts have clarified, prospective studies need not be representative to provide insights into exposure-disease relationships that may be just as relevant to the general population. Moreover, in nationally representative surveys in Cuba and Mexico conducted around the median follow-up time in each study, the smoking prevalence and amount smoked were reasonably similar to the prevalence and consumption observed in the Cuban and Mexican study populations (and the US cohort was based on a nationally-representative survey). Even so, these potential differences between the study population and the general population, and subsequent potential impacts on exposure-disease relationships, merit consideration.

In each of these studies, follow-up was conducted passively (ie, through routine linkage to death registries). As a result, it is possible that some participants would be lost to follow-up, and that some deaths may occur without being recorded. However, emigration is relatively uncommon in these populations, and given the complete vital registration systems in these settings, unrecorded deaths should be rare. Each study also relied solely on self-reported information about smoking, which may have resulted in information bias. (Although studies of smoking biomarkers, such as cotinine, could

theoretically be possible in the Mexican study using stored blood samples, this was not possible at the time this research was conducted.) However, resurvey information among surviving participants in the Cuban and Mexican cohorts was used to correct (at least in part) this error. Finally, although analyses in each of these populations adjusted for major likely confounders, including age, sex, education, area, and alcohol consumption, residual confounding cannot be ruled out.

6.6 Implications for future research

Although the topic of this thesis was smoking and cause-specific mortality in three populations in the Americas, the conduct of this research generated several research questions of both immediate and broader relevance. These areas of future research pertain primarily to the exposures and outcomes studied in this thesis, and to their implications for future research not only for smoking, but for prospective studies of a range of exposures and health outcomes. These implications for future research are summarised below.

6.6.1 Assessment of exposure

Although the findings from this thesis and other historical and contemporary studies underscore the utility of self-reported information for estimating smoking-related morbidity and mortality, advances in technology and methodology have opened new avenues for determining smoking exposure. Although cotinine, a metabolite of nicotine, is commonly used as a biomarker for environmental (or 'passive') smoke exposure,[138] cotinine and a number of other blood-, urine-, and saliva-based biomarkers may offer insight into both active and passive smoking exposure.[117,139,140] Although the accuracy of these biomarkers may depend on factors such as race,[141,142] they may prove particularly informative among populations in which smoking is taboo (such as pregnant women),[143] or as a means of better quantifying this exposure for which digit preference is

common.[144] The use of such biomarkers for studying smoking may also aid in unravelling the specific mechanisms by which smoking-related diseases develop or progress, particularly when used in concert with other biomarkers of human health.[145,146]

Although randomised, controlled trials assigning individuals to smoke cigarettes would be unethical due to the indisputable evidence that smoking is harmful, studies mimicking this design may prove useful in providing further causal evidence of the relationship between smoking and disease. For example, by using single nucleotide polymorphisms (SNPs) associated with smoking propensity as instrumental variables (ie, using the Mendelian randomisation approach), it is now possible to examine smoking-disease relationships while avoiding (or at least reducing) confounding. This study design (in combination with genome-wide association studies [GWAS] that have already provided useful insight into the genetic determinants of smoking behaviours) has helped identify the specific pathways by which smoking may cause harm, and provided new insights into the relationship between smoking and (ill) health.[90,147,148] Given the particularly unusual smoking behaviours among Mexicans – highly prevalent non-daily smoking, low levels of cigarette consumption, and high rates of cessation – the Mexico City Prospective Study may provide valuable insight into the interplay between genetics, environment, and lifestyle when it comes to smoking and smoking cessation. Further investigation among Mexican Americans may shed additional light on this topic, and I have begun exploratory analyses of the relevance of nicotine metabolites to smoking behaviours (both active and passive) and health outcomes by race using the US National Health and Nutrition Examination Survey (NHANES).

Finally, further research into the rapidly-evolving tobacco products marketed as alternatives to traditional cigarettes is an increasingly urgent public health priority. As cigarette smoking prevalence continues to decline in the US and a number of other

countries,[1] e-cigarette usage (particularly among children and adolescents) has swiftly increased. In the US between the years 2011 and 2018, e-cigarette use among middle school students (ie, aged approximately 12-14 years) rose from 0.6% to 4.9%, and use among high school students (ie, aged approximately 14-18 years) rose from 1.5% to 20.8%.[149] Although concerns had already been mounting that e-cigarette use could lead children and adolescents to smoking regular cigarettes,[150,151] causing harm in later life, increasing evidence suggests that use of these products themselves may prove fatal.[152,153] Although the sale of e-cigarettes has been banned in several countries,[154] including Mexico,[155] evidence suggests that use of these products is nevertheless spreading across the globe.[156,157] The relevance of these alternative tobacco products to morbidity and mortality, as well as to quitting (or reducing the consumption of) cigarettes, requires further research. At present, the prevalence, correlates, and risks of these products are little-known even in high-income countries, let alone in diverse populations globally.

6.6.2 Assessment of outcomes

The prospective analyses performed in this thesis relied on follow-up through linkage to national mortality registries. Future research will benefit from the continued accrual of fatal events, which would enhance the precision of estimates and allow for analyses to be performed in additional sub-groups (and their potential interactions, eg, the combined effects of smoking and alcohol). Although nearly all deaths are medically certified in each of the populations studied, the causes of death may nevertheless, in some cases, be mis-certified (particularly in resource-limited settings, and at older ages).[91] Therefore, these studies could be strengthened by linkage to non-fatal event records (eg, electronic medical records), which could enhance study precision, provide further evidence on potential causes of death (and mediating factors), and allow for the study of disabling diseases that may not lead to death (or may not be the primary cause). For example,

there is some evidence that smoking is associated with incident diabetes in certain populations,[158,159] and unravelling this potential relationship could prove especially useful in Mexico, where diabetes is a major driver of premature mortality.[103,104] At present, linkage to electronic health records in Mexico City, and a resurvey of all surviving participants (which will capture non-fatal health events), are both in the pilot stages.

As electronic health records become more common outside of Western populations, and as vital registration systems continue to improve globally, follow-up for health outcomes should continue to improve the conduct of such studies in diverse populations globally. Furthermore, advances in technology may transform methods of data collection for both surveys and follow-up of participants (eg, by mobile phone or Internet).[160,161] Such changes have the potential to enhance the scope of follow-up for outcomes, and to allow studies to react to changes in the health landscape that inevitably occur over the long periods of follow-up essential to prospective studies (eg, e-cigarettes did not exist when the Mexican and Cuban cohorts were established ~20 years ago).[162] Furthermore, these advances may reduce loss to follow-up among less settled populations, and improve participation rates among certain groups.[163] Leveraging such methodological and technological advances may appreciably improve the scope of follow-up in prospective studies in diverse populations globally.

6.6.3 Prospective study of lifestyle risk factors in diverse populations globally

The use of prospective studies to directly estimate the relevance of risk factors like smoking to death and disability (particularly before old age) remains a global health priority. Smoking, blood pressure, blood lipids, and diabetes remain leading causes of death globally. The main need in addressing these risks is effective political action, but this could be aided by ongoing epidemiological research. Disease control priorities

should be shaped by reliable evidence to identify the most pressing threats to population health, as well as the most promising solutions.

As prospective studies may take several years to set up, and several years more to accrue a sufficient number of events to provide reliable results, case-control studies of these risk factors may prove useful in providing rapid results in populations where direct evidence is limited (ideally planned for and set up concurrently with prospective studies). This strategy has provided particularly important information on smoking in India[62] and on alcohol use in Russia[164], among other examples. From a population health perspective, such studies should be prioritised in large countries where they may have the biggest impact on population health (and particularly where limited local or regional information currently exists), such as Indonesia, Brazil, and Nigeria. However, as findings on smoking and mortality from Cuba have made clear, even in smaller countries, this evidence may prove to be relevant on a much wider scale.

Although large, prospective studies have been used to identify the major causes of death and disability in countries like the US and UK for generations, such studies have yet to be conducted in most populations globally. As the findings from this thesis make clear, these studies can provide important new insights even on 'known' risk factors like smoking, in addition to the other leading risk factors for mortality globally. In the past few years alone, large, prospective studies of diabetes in Mexico,[103] alcohol in China,[165] and body mass index in India[166] have generated new insights on the relevance of these risk factors to premature mortality, underscoring where extrapolations from Western populations fall short. As technology continues to transform how and where such studies can reasonably be conducted, large, prospective studies in diverse populations should help guide global health priorities and action around the world.

Sir Richard Doll observed in 1971, “The geographical distribution of cancer has been studied for so many years that it may well be thought that the subject can have very little more to teach us. That, however, I believe to be wrong.”[167] Nearly 50 years later, the philosophy of this sentiment appears to be true not only for cancer, but for many causes of death and their risk factors. As Doll and others clarified over decades of work on the British Doctors Study, the rapid emergence of lung cancer observed in the UK (but not in many other countries at the time) was caused by smoking. This finding, which resulted in clear public health messages and informed policies to combat smoking, led to widespread smoking cessation and large reductions in lung cancer rates in the UK. Hopefully, prospective studies from diverse populations globally may lead to similar success stories, and to healthier populations around the globe.

6.7 Conclusions

In large, prospective studies of adults in Cuba, Mexico, and the United States, tobacco smoking was strongly associated with increased risk of mortality overall and from several causes. These associations varied by population, and depended strongly on the age at smoking initiation and amount smoked. There was strong evidence that smoking initiation before age 15, and especially before age 10, was associated with particularly high excess risk of premature death. Although the excess risk of death associated with smoking was lower among those who smoked fewer cigarettes per day than among those who smoked more, regularly smoking even a few cigarettes per day was associated with elevated mortality risk. Although these associations were generally similar among men and women, higher smoking prevalence among men led to a greater estimated proportion of smoking-attributable deaths among men than among women. The associations between current smoking and mortality observed in Cuba and Mexico were similar to those observed in other low- and middle-income populations previously studied, but were weaker than those of high-income populations like the UK or US. As

observed in a number of populations, however, the excess risk of premature death among those who quit was substantially lower than that of continuing smokers, and those who quit before age 40 avoided most of the excess risk associated with smoking. These findings contribute to the body of evidence that smoking kills, but quitting works, in diverse populations globally. To avoid 1 billion tobacco deaths this century, prevention of smoking *and* widespread cessation must occur in populations around the world.

Table 6.1: Population attributable fractions for mortality from medical causes associated with ever-regular smoking in Cuba, Mexico, and the United States

Cohort	PAF* (Men, %)	PAF (Women, %)
Cuba	27	19
Mexico	10	4
United States		
Mexican American	12	4
Other Hispanic	21	13
Non-Hisp. White	31	26
Non-Hisp. Black	24	16
Other	18	11

*Calculated as ever- versus never-smoking, adjusted for age-at-risk, education, area, and alcohol consumption. In Cuba, restricted to deaths at ages 30-69, in Mexico restricted to deaths at ages 35-89, and in the United States restricted to deaths at ages 30-89.

Table 6.2: Global Adult Tobacco Survey respondent characteristics by country

Country	Survey year	Mean age (years)	Female (%)	Smoke daily (%)	Mean cigs/day	Unweighted sample size	Estimated population (millions)
Argentina	2012	41	52.4	17.1	15	6645	27.6
Bangladesh	2009	36	50.3	20.9	9	9629	95.4
Brazil	2008	40	52.1	15.1	14	39 425	143.0
Cameroon	2013	34	51.7	4.6	9	5271	12.0
China	2010	42	49.1	24.1	16	13 354	1070.0
Costa Rica	2015	40	49.7	5.8	13	8607	3.7
Egypt	2009	35	49.0	18.5	19	20 924	49.7
Ethiopia	2016	31	50.1	3.2	14	10150	68.4
Greece	2013	45	50.9	36.6	20	4359	9.3
India	2010	36	48.3	10.7	6	69 296	796.0
Indonesia	2011	38	50.1	29.2	13	8305	172.0
Kazakhstan	2014	39	52.8	19.1	15	4425	12.6
Kenya	2014	34	51.2	6.0	9	4408	21.9
Malaysia	2011	38	48.5	20.9	13	4250	20.5
Mexico	2015	39	52.1	7.6	8	14 664	87.6
Nigeria	2012	34	50.0	2.9	8	9765	81.7
Pakistan	2014	35	48.8	11.5	14	7831	126.0
Panama	2013	40	50.1	2.8	14	16 962	2.7
Philippines	2009	36	50.1	22.5	11	9701	61.3
Poland	2010	44	52.3	27.0	17	7840	32.3
Qatar	2013	35	47.5	9.5	17	8398	0.4
Romania	2011	46	51.8	24.3	17	4517	18.2
Russia	2009	44	54.7	33.8	17	11 406	112.0
Senegal	2015	35	51.5	4.9	9	4347	7.8
Thailand	2011	42	51.4	21.5	12	20 606	54.2
Turkey	2012	39	50.8	23.8	18	9851	54.5
Uganda	2013	34	52.8	4.8	6	8508	16.7
Ukraine	2010	44	54.6	25.5	17	8158	40.0
Uruguay	2009	44	52.6	20.4	15	5581	2.5
Vietnam	2015	40	51.5	19.2	14	8996	69.3
Total	NA	39	49.8	18.4	15	366 179	3269.2

NA=not applicable. All characteristics are estimated using survey weights, with the exception of unweighted sample size

Figure 6.1: Prevalence of smoking frequency among men in GATS surveys, by country

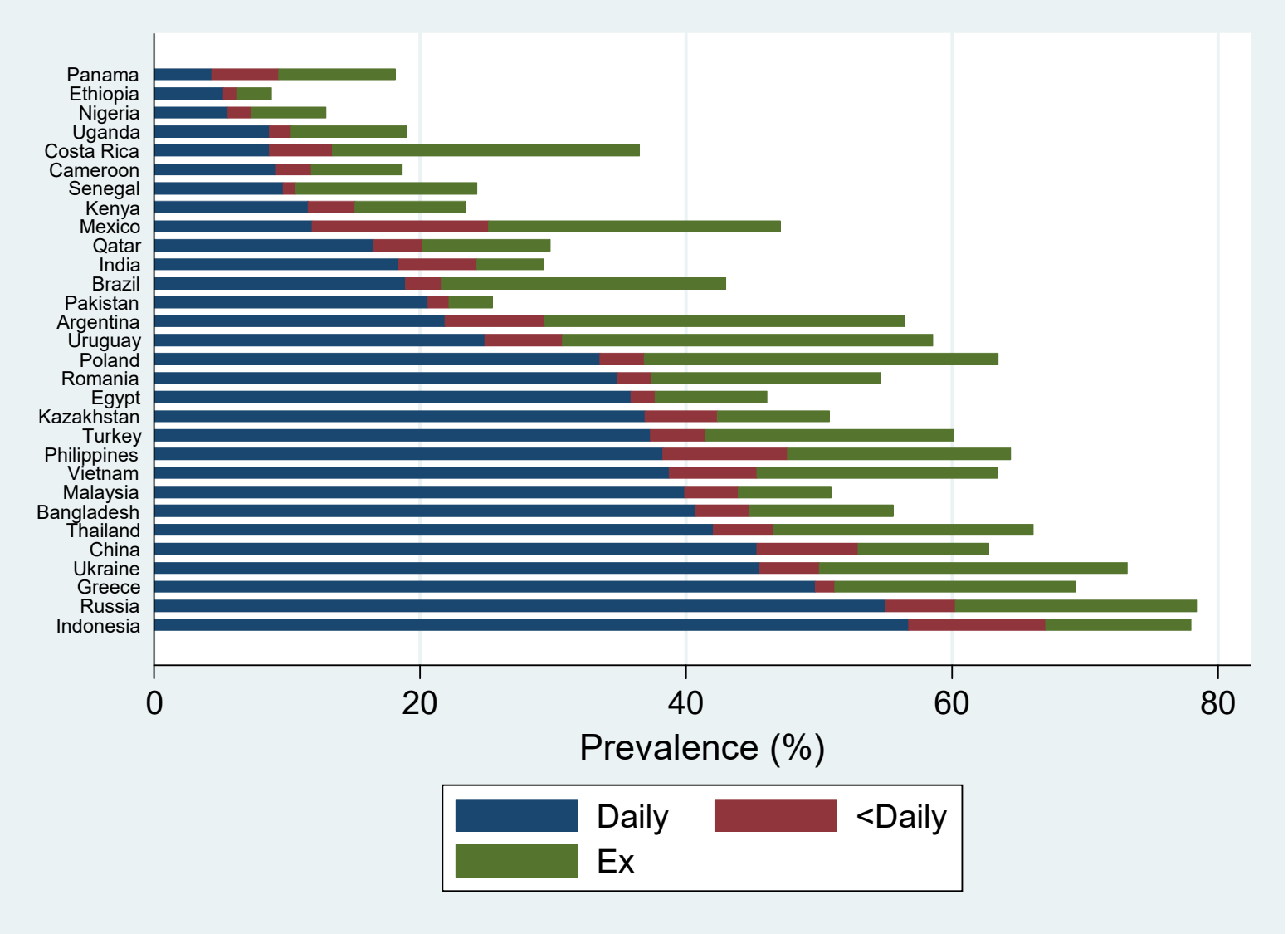


Figure 6.2: Prevalence of smoking frequency among women in GATS surveys, by country

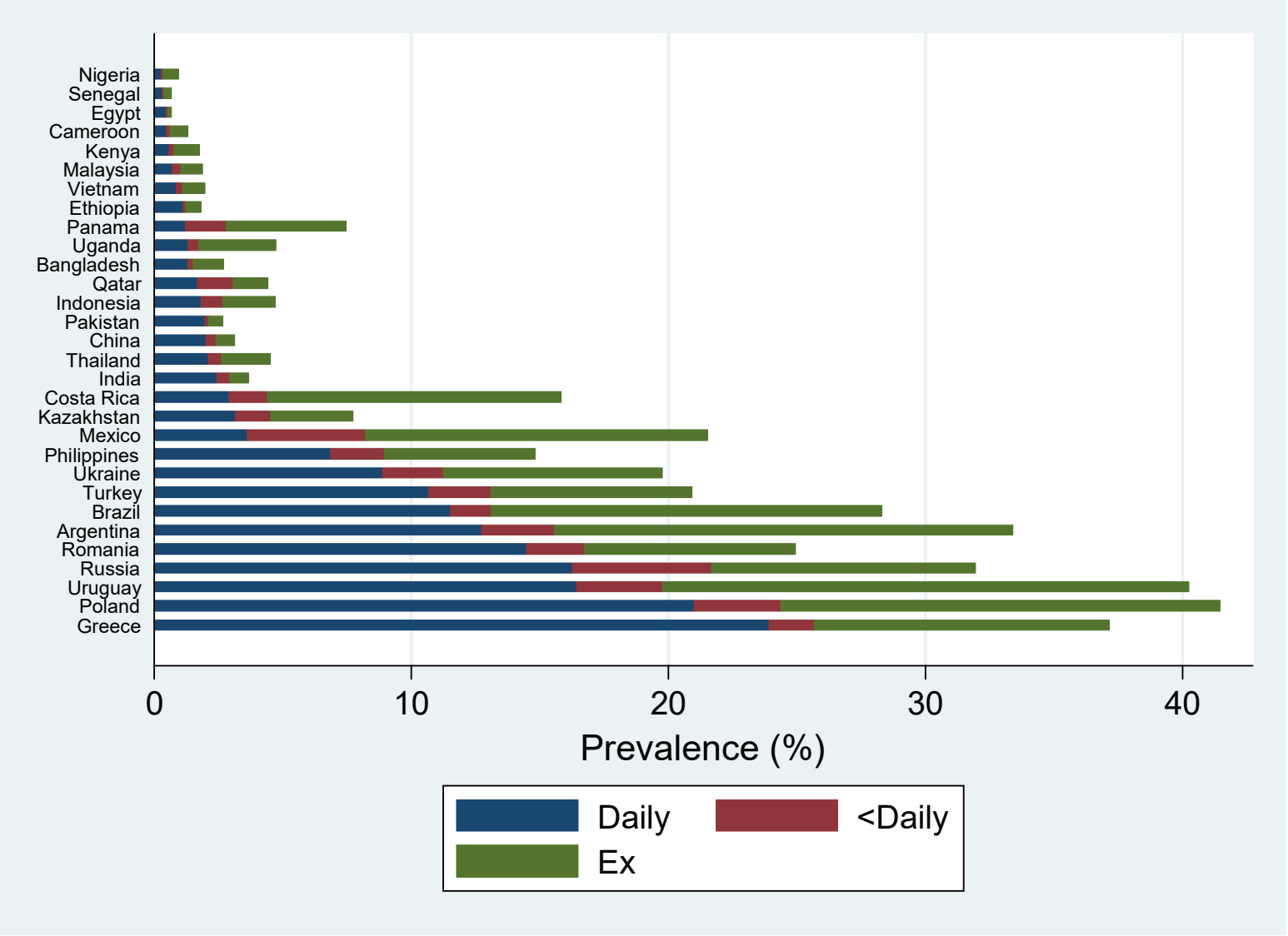


Figure 6.3: Proportion of male current smokers who smoke daily in GATS surveys, by country

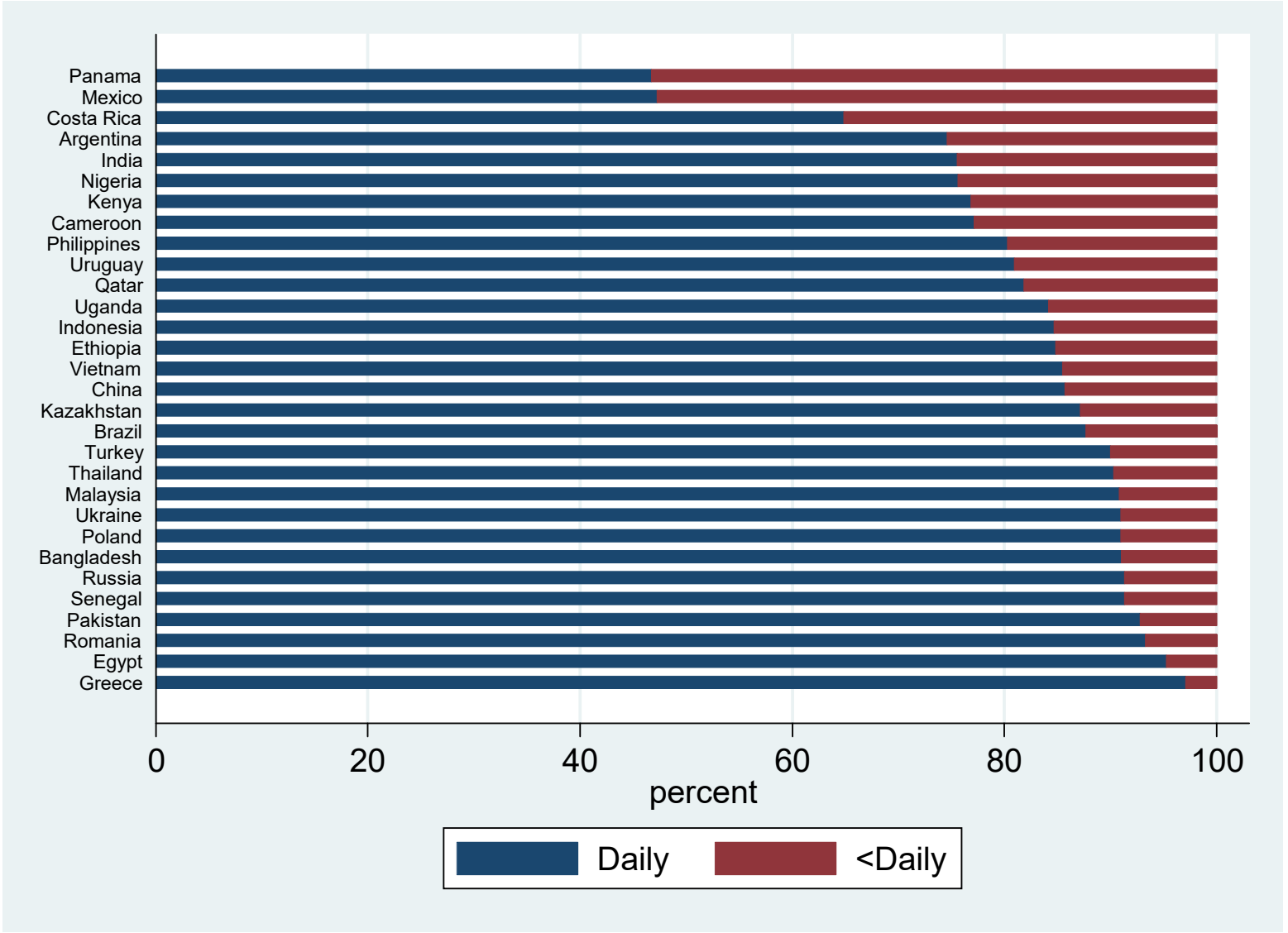


Figure 6.4: Proportion of female current smokers who smoke daily in GATS surveys, by country

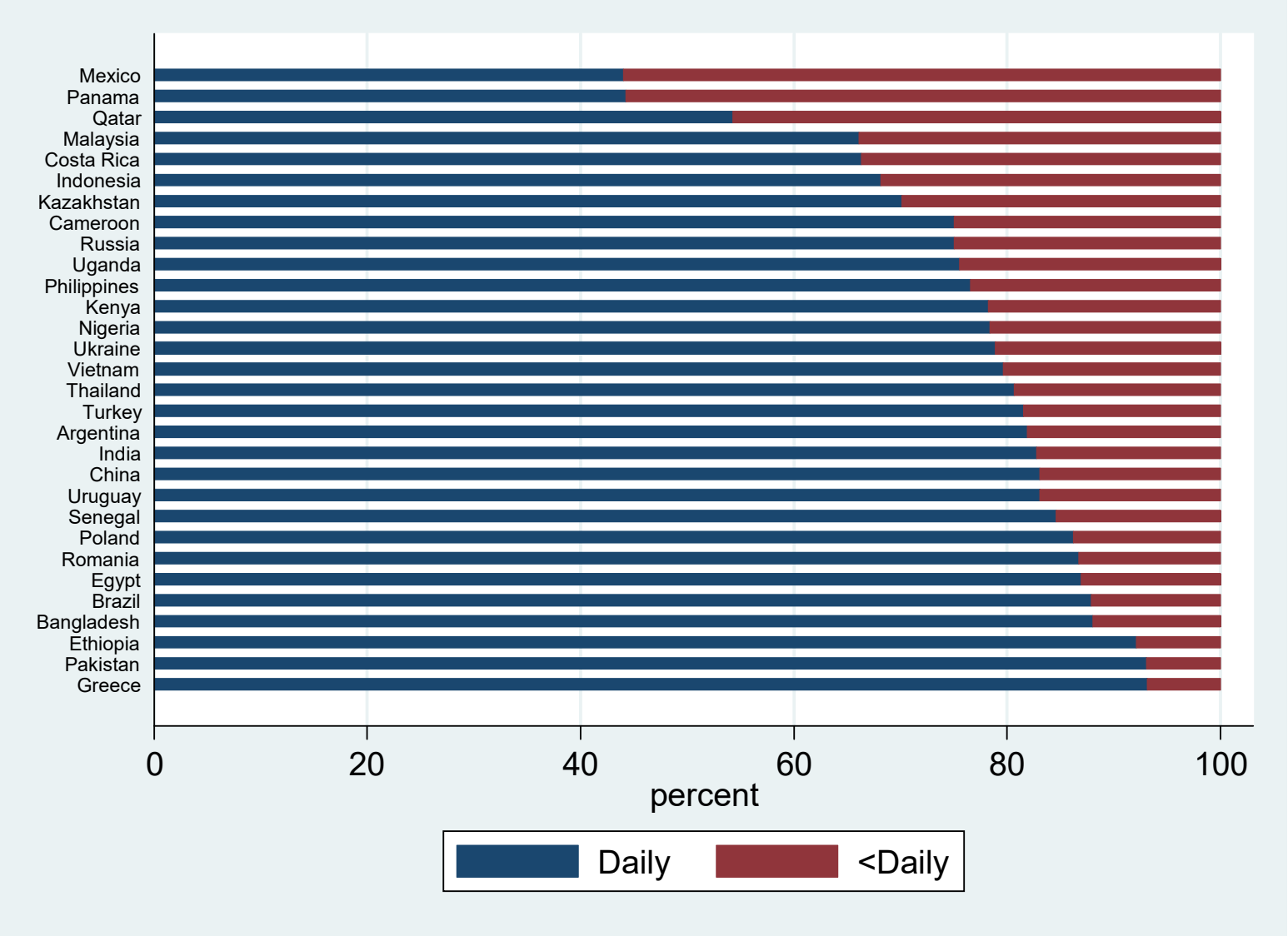


Figure 6.5: Cigarettes smoked per day among male cigarette smokers in GATS surveys, by country

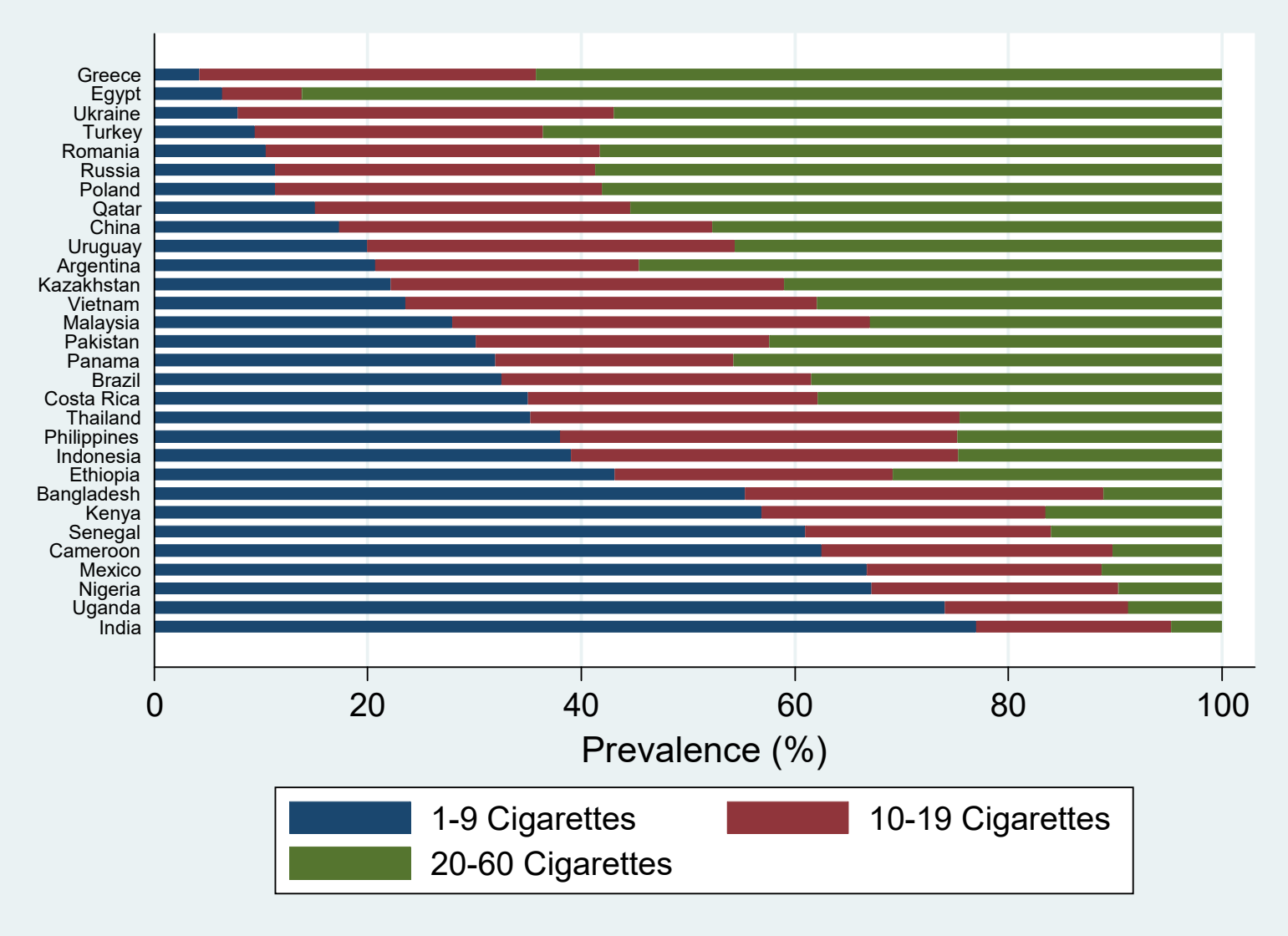


Figure 6.6: Cigarettes smoked per day among female cigarette smokers in GATS surveys, by country

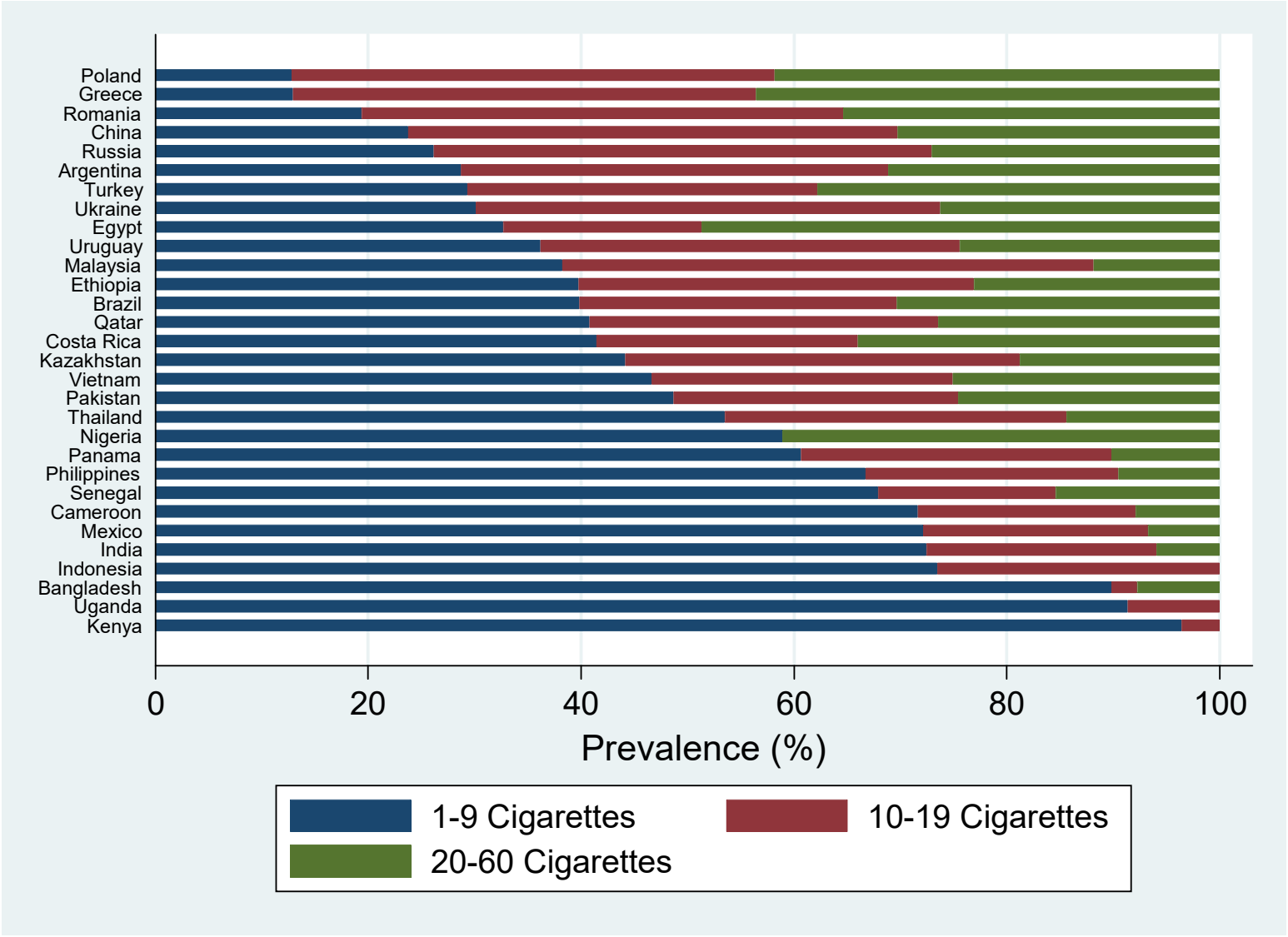


Figure 6.7: Age at smoking initiation among male cigarette smokers in GATS surveys, by country

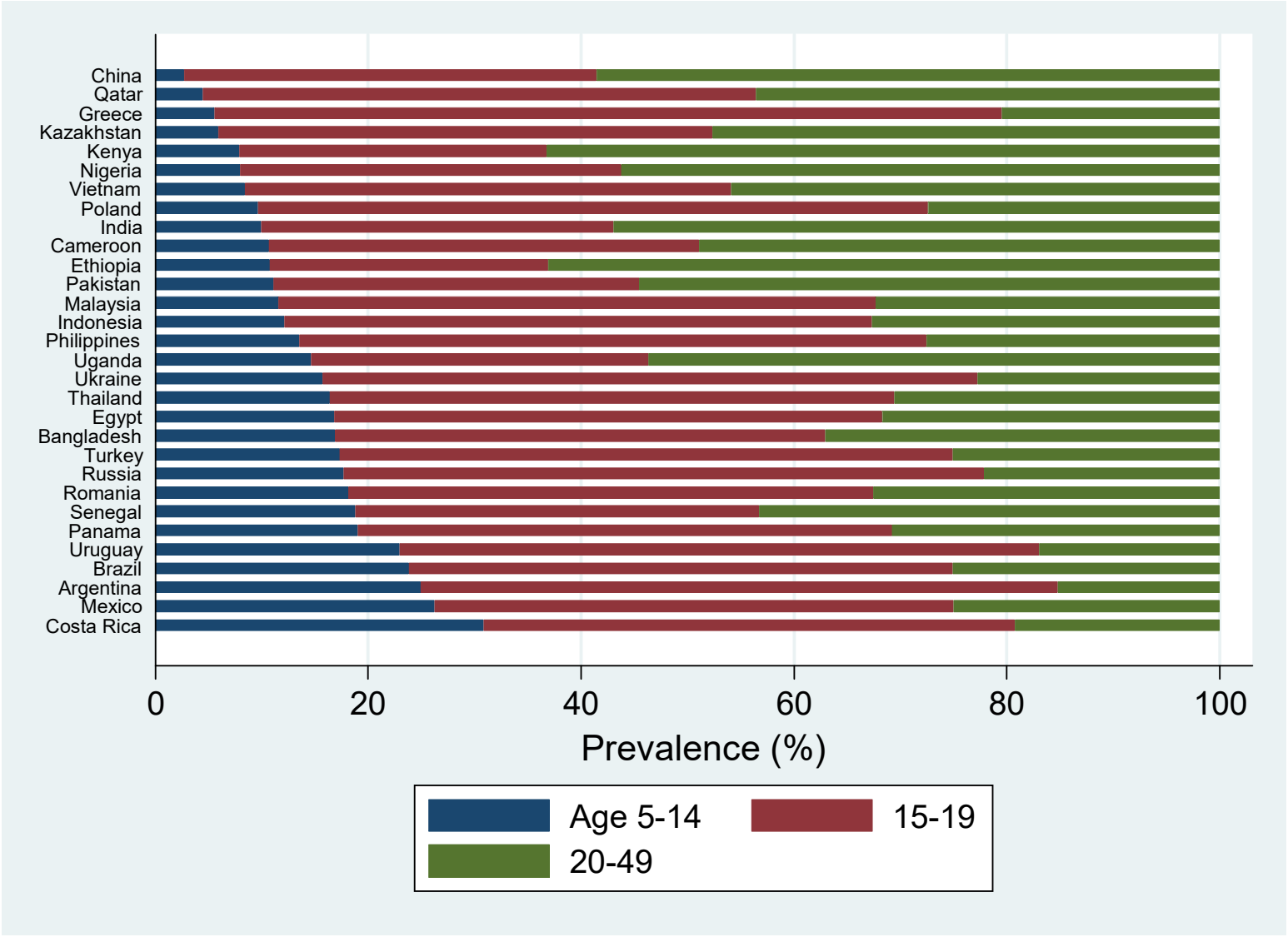


Figure 6.8: Age at smoking initiation among female cigarette smokers in GATS surveys, by country

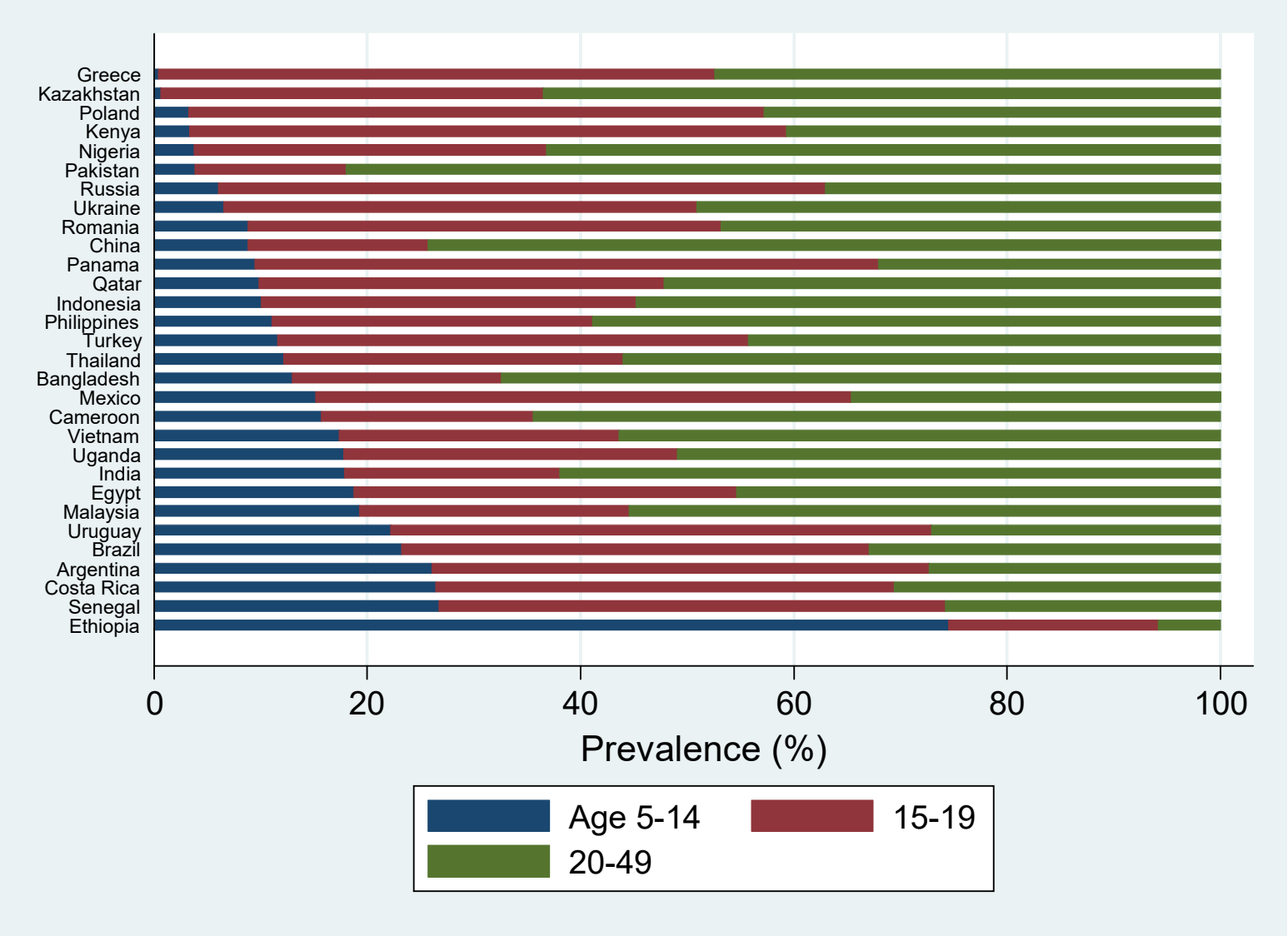


Figure 6.9: Adult smokers who began smoking daily at ages 10-14 in GATS surveys, by country

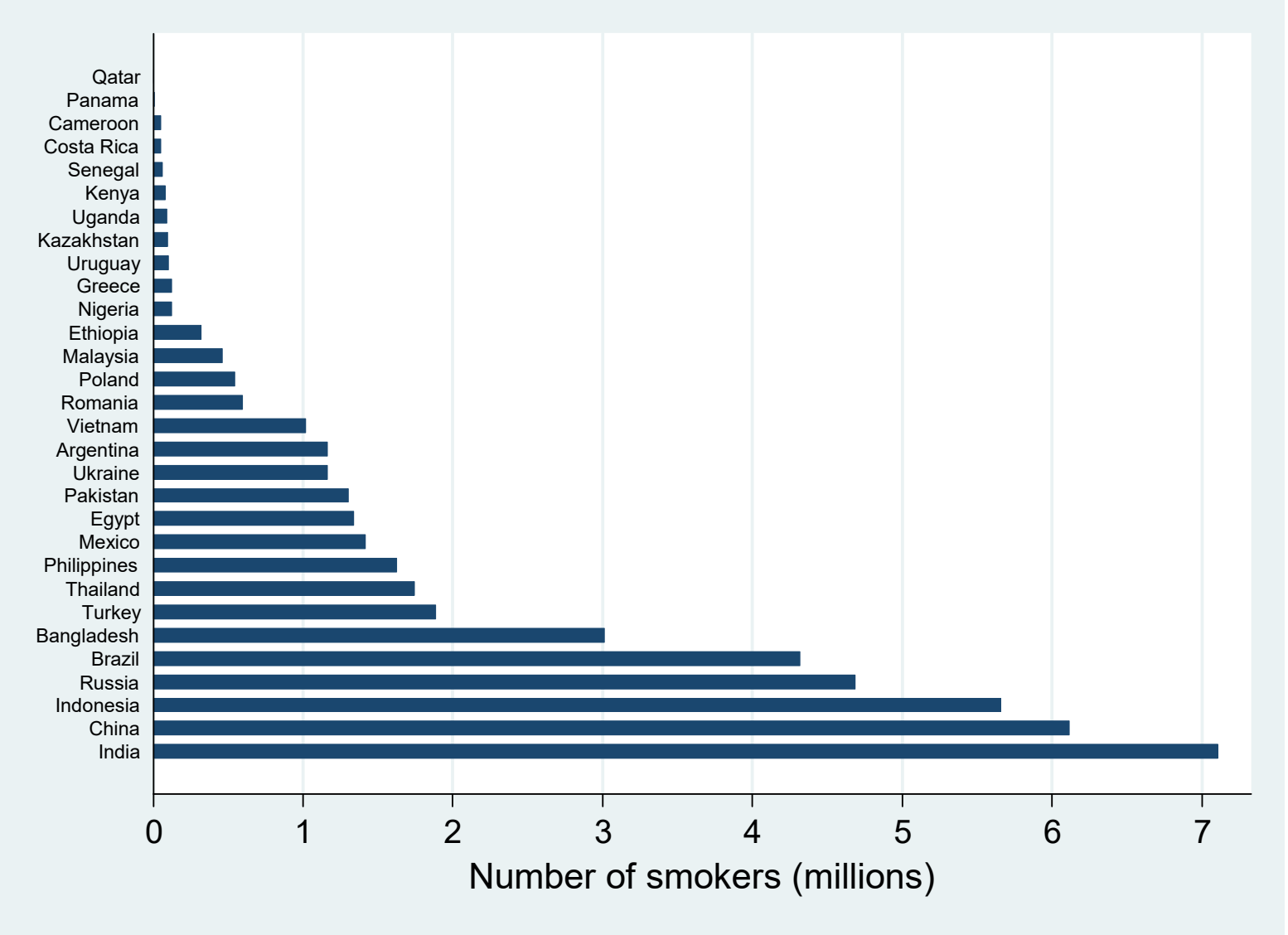


Figure 6.10: Adult smokers who began smoking daily at ages 5-9 in GATS surveys, by country

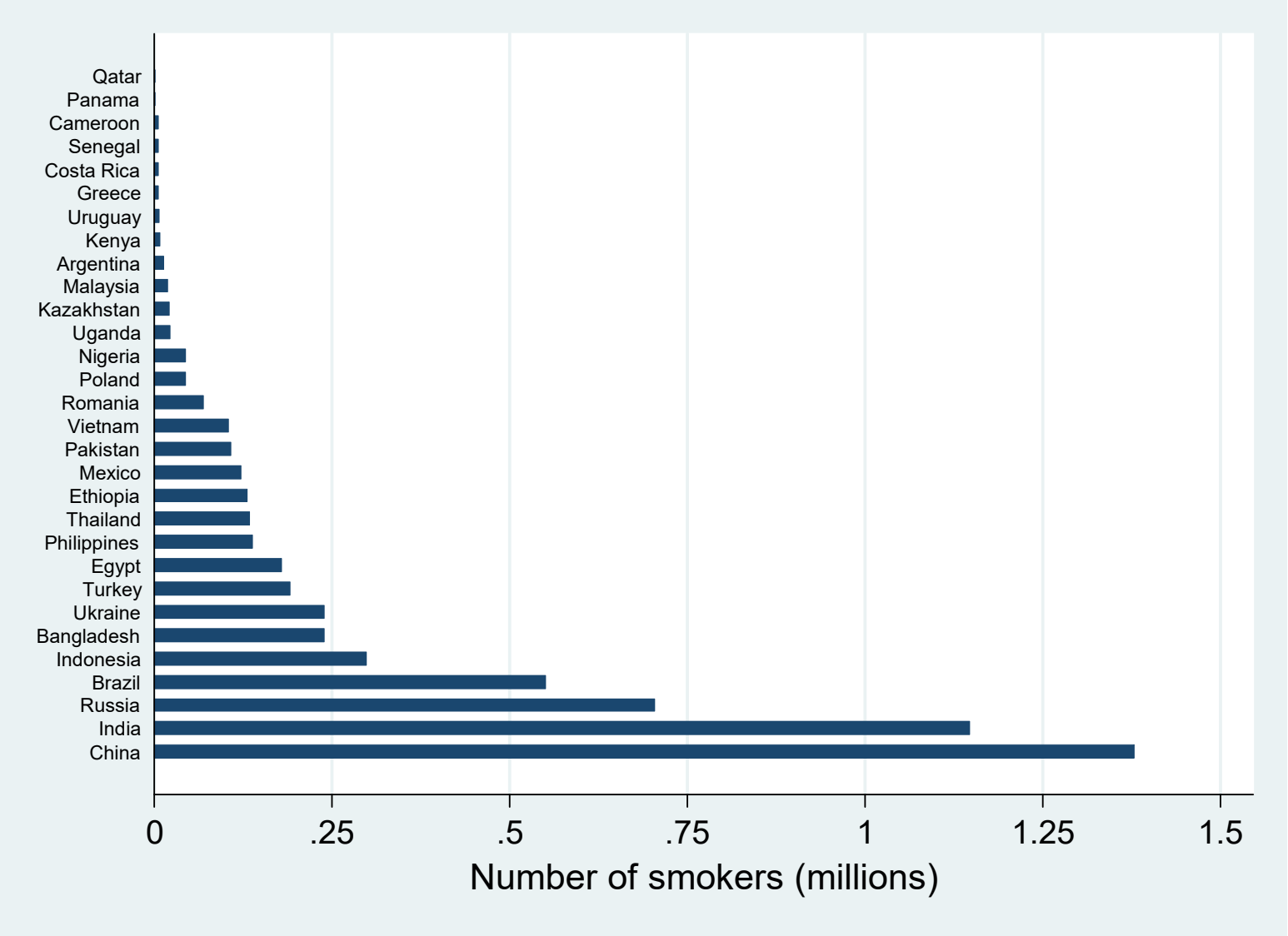


Figure 6.11: Quit ratio among ever-smoking men in GATS surveys, by country

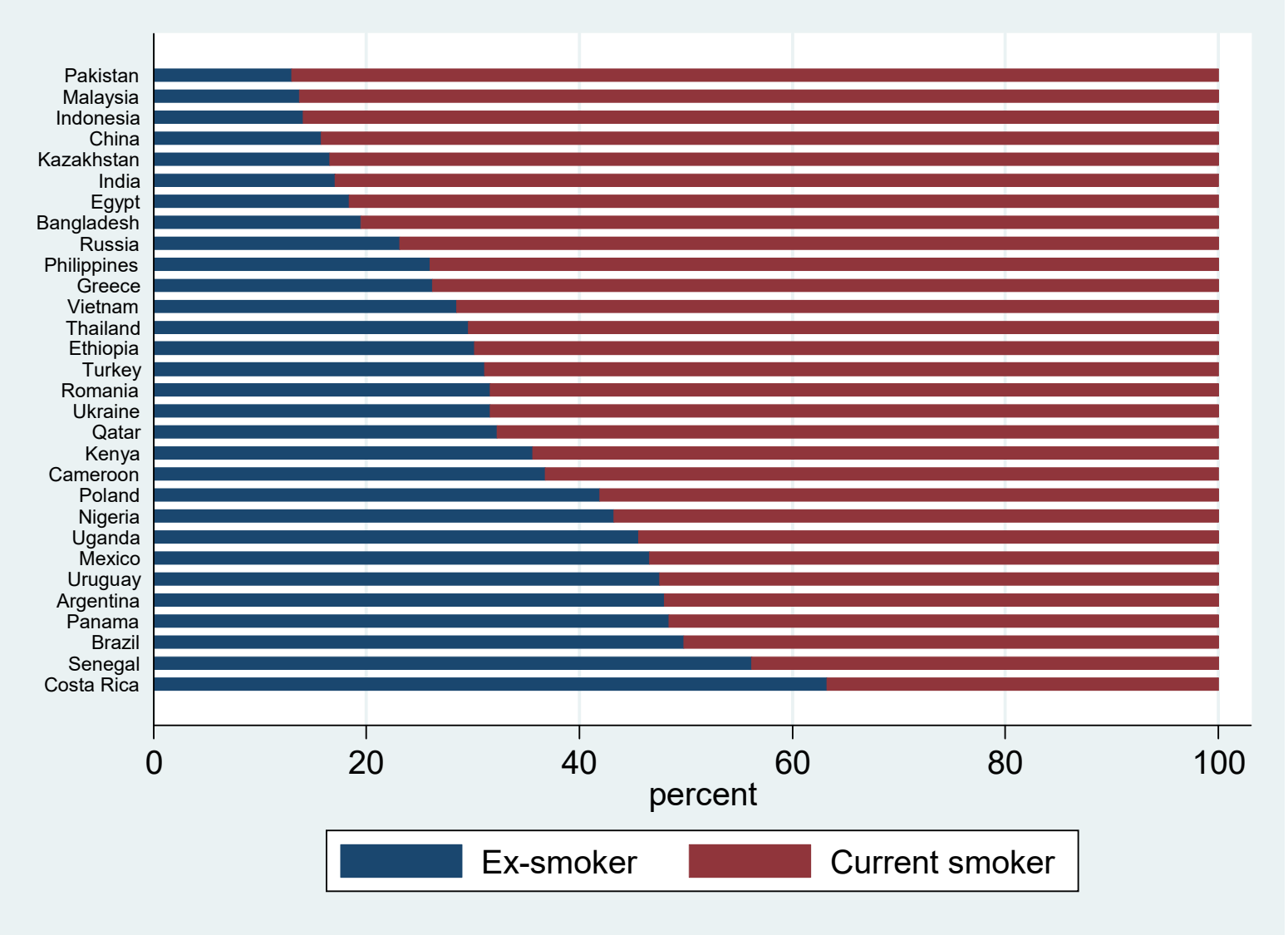


Figure 6.12: Quit ratio among ever-smoking women in GATS surveys, by country

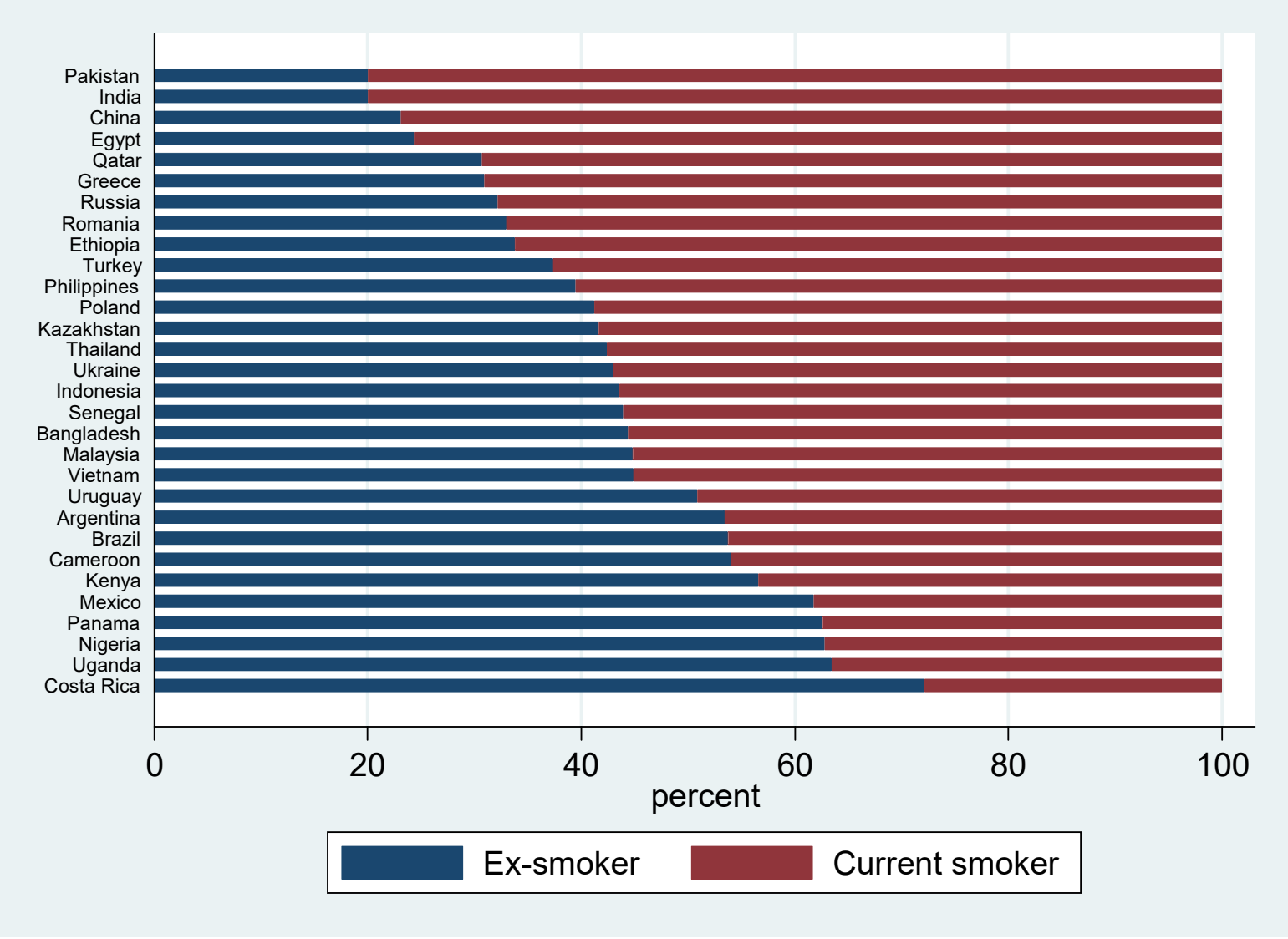


Figure 6.13: Time to first smoking after waking among men in GATS surveys, by country

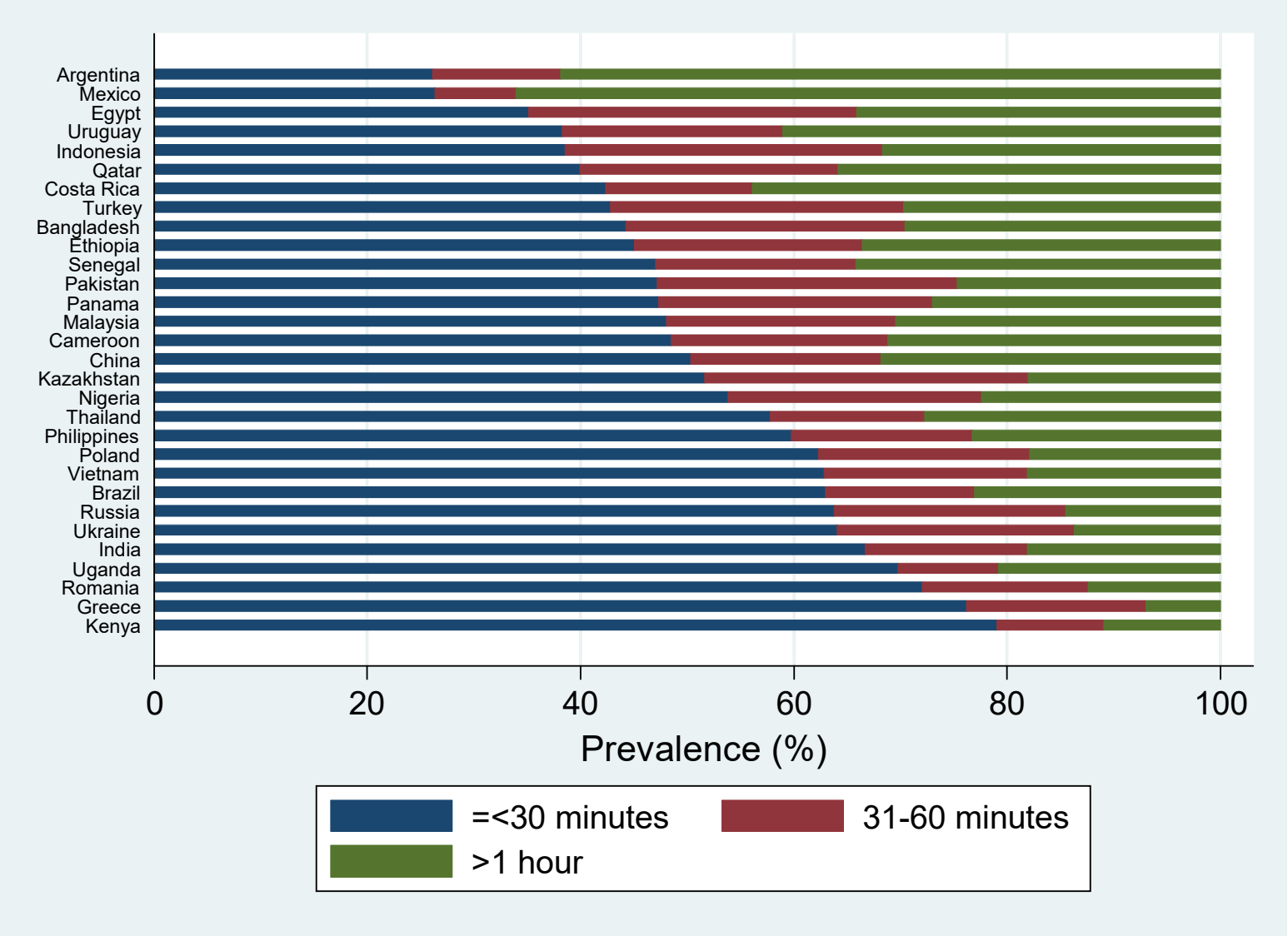
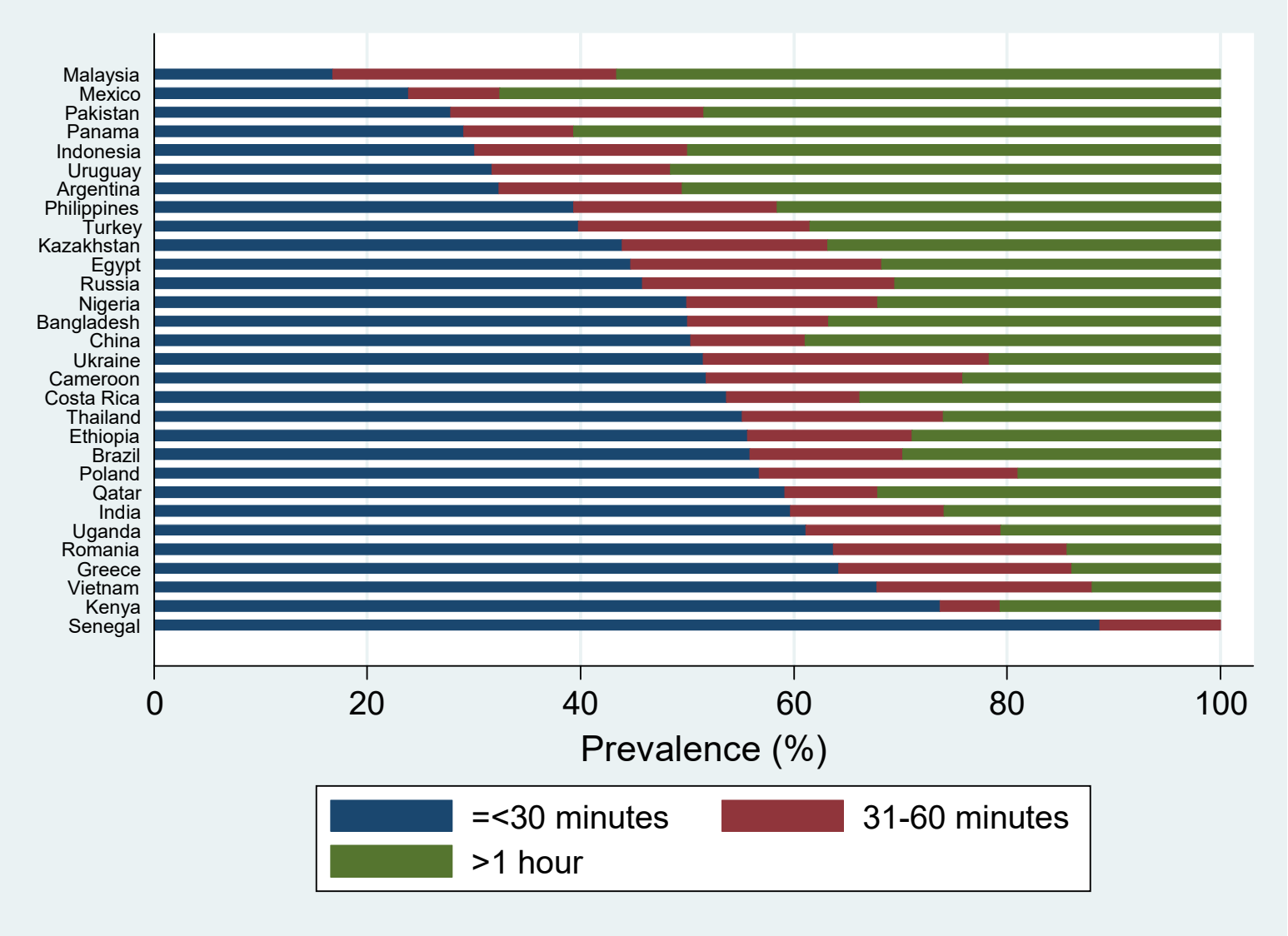


Figure 6.14: Time to first smoking after waking among women in GATS surveys, by country



References

1. GBD Collaborators. Smoking prevalence and attributable disease burden in 195 countries and territories, 1990-2015: a systematic analysis from the Global Burden of Disease Study 2015. *Lancet* 2017; **389**(10082): 1885-906.
2. Jha P, Peto R. Global effects of smoking, of quitting, and of taxing tobacco. *N Engl J Med* 2014; **370**(1): 60-8.
3. Doll R, Peto R, Boreham J, Sutherland I. Mortality in relation to smoking: 50 years' observations on male British doctors. *BMJ* 2004; **328**(7455): 1519.
4. Jha P, Ramasundarahettige C, Landsman V, et al. 21st-century hazards of smoking and benefits of cessation in the United States. *N Engl J Med* 2013; **368**(4): 341-50.
5. Banks E, Joshy G, Weber MF, et al. Tobacco smoking and all-cause mortality in a large Australian cohort study: findings from a mature epidemic with current low smoking prevalence. *BMC Med* 2015; **13**: 38.
6. Jha P. Avoidable global cancer deaths and total deaths from smoking. *Nat Rev Cancer* 2009; **9**(9): 655-64.
7. Thun M, Peto R, Boreham J, Lopez AD. Stages of the cigarette epidemic on entering its second century. *Tob Control* 2012; **21**(2): 96-101.
8. Thun MJ, Carter BD, Feskanich D, et al. 50-year trends in smoking-related mortality in the United States. *N Engl J Med* 2013; **368**(4): 351-64.
9. Chen Z, Peto R, Zhou M, et al. Contrasting male and female trends in tobacco-attributed mortality in China: evidence from successive nationwide prospective cohort studies. *Lancet* 2015; **386**(10002): 1447-56.
10. Doll R, Hill AB. Smoking and carcinoma of the lung; preliminary report. *Br Med J* 1950; **2**(4682): 739-48.
11. Doll R. Uncovering the effects of smoking: historical perspective. *Stat Methods Med Res* 1998; **7**(2): 87-117.
12. Wynder EL, Graham EA. Tobacco smoking as a possible etiologic factor in bronchiogenic carcinoma; a study of 684 proved cases. *J Am Med Assoc* 1950; **143**(4): 329-36.
13. Schrek R, Baker LA, et al. Tobacco smoking as an etiologic factor in disease. *Cancer Res* 1950; **10**(1): 49-58.
14. Mills CA, Porter MM. Tobacco smoking habits and cancer of the mouth and respiratory system. *Cancer Res* 1950; **10**(9): 539-42.
15. Levin ML, Goldstein H, Gerhardt PR. Cancer and tobacco smoking; a preliminary report. *J Am Med Assoc* 1950; **143**(4): 336-8.
16. Keating C. Smoking Kills: The Revolutionary Life of Richard Doll. Oxford: Signal Books Limited; 2009.
17. Doll R, Hill AB. The mortality of doctors in relation to their smoking habits; a preliminary report. *Br Med J* 1954; **1**(4877): 1451-5.
18. Hammond EC, Horn D. The relationship between human smoking habits and death rates: a follow-up study of 187,766 men. *J Am Med Assoc* 1954; **155**(15): 1316-28.
19. Kreyberg L. Lung cancer and tobacco smoking in Norway. *Br J Cancer* 1955; **9**(4): 495-510.
20. Pernu J. An epidemiological study of cancer of the digestive organs and respiratory system. A study based on 7078 cases. *Ann Med Intern Fenn Suppl* 1960; **49(Suppl 33)**: 1-117.
21. Lancaster HO. Cancer statistics in Australia: II. Respiratory stystem. *Med J Aust* 1962; **49(1)**: 1006-11.
22. Koulumies M. Smoking and pulmonary carcinoma. *Acta radiol* 1953; **39**(3): 255-60.
23. Schwartz D, Denoix PF. [French investigations on the etiology of bronchopulmonary cancer; role of tobacco]. *Sem Hop* 1957; **33**(62/7): 3630-43.

24. Lickint F. [Etiology and prophylaxis of cancer of the lungs]. *Strahlentherapie* 1953; **91**(4): 551-4.
25. Doll R, Hill AB. Mortality in Relation to Smoking: Ten Years' Observations of British Doctors. *Br Med J* 1964; **1**(5396): 1460-7 CONCL.
26. Doll R, Peto R. Mortality in relation to smoking: 20 years' observations on male British doctors. *Br Med J* 1976; **2**(6051): 1525-36.
27. Doll R, Peto R, Wheatley K, Gray R, Sutherland I. Mortality in relation to smoking: 40 years' observations on male British doctors. *BMJ* 1994; **309**(6959): 901-11.
28. Hammond EC. Smoking in Relation to Mortality and Morbidity. Findings in First Thirty-Four Months of Follow-up in a Prospective Study Started in 1959. *J Natl Cancer Inst* 1964; **32**: 1161-88.
29. Shopland DR, Eyre HJ, Pechacek TF. Smoking-attributable cancer mortality in 1991: is lung cancer now the leading cause of death among smokers in the United States? *J Natl Cancer Inst* 1991; **83**(16): 1142-8.
30. Peto R, Lopez AD, Boreham J, Thun M, Heath C, Jr. Mortality from tobacco in developed countries: indirect estimation from national vital statistics. *Lancet* 1992; **339**(8804): 1268-78.
31. Doll R, Gray R, Hafner B, Peto R. Mortality in relation to smoking: 22 years' observations on female British doctors. *Br Med J* 1980; **280**(6219): 967-71.
32. Garfinkel L, Stellman SD. Smoking and lung cancer in women: findings in a prospective study. *Cancer Res* 1988; **48**(23): 6951-5.
33. Fisher RA. Lung cancer and cigarettes. *Nature* 1958; **182**(4628): 108.
34. Thun MJ, Day-Lally CA, Calle EE, Flanders WD, Heath CW, Jr. Excess mortality among cigarette smokers: changes in a 20-year interval. *Am J Public Health* 1995; **85**(9): 1223-30.
35. Pirie K, Peto R, Reeves GK, Green J, Beral V, Million Women Study C. The 21st century hazards of smoking and benefits of stopping: a prospective study of one million women in the UK. *Lancet* 2013; **381**(9861): 133-41.
36. Parish S, Collins R, Peto R, et al. Cigarette smoking, tar yields, and non-fatal myocardial infarction: 14,000 cases and 32,000 controls in the United Kingdom. The International Studies of Infarct Survival (ISIS) Collaborators. *BMJ* 1995; **311**(7003): 471-7.
37. Tang JL, Morris JK, Wald NJ, Hole D, Shipley M, Tunstall-Pedoe H. Mortality in relation to tar yield of cigarettes: a prospective study of four cohorts. *BMJ* 1995; **311**(7019): 1530-3.
38. Stellman SD, Garfinkel L. Lung cancer risk is proportional to cigarette tar yield: evidence from a prospective study. *Prev Med* 1989; **18**(4): 518-25.
39. Harris JE, Thun MJ, Mondul AM, Calle EE. Cigarette tar yields in relation to mortality from lung cancer in the cancer prevention study II prospective cohort, 1982-8. *BMJ* 2004; **328**(7431): 72.
40. Doll R, Peto R. Cigarette smoking and bronchial carcinoma: dose and time relationships among regular smokers and lifelong non-smokers. *J Epidemiol Community Health* 1978; **32**(4): 303-13.
41. Peto J. That the effects of smoking should be measured in pack-years: misconceptions 4. *Br J Cancer* 2012; **107**(3): 406-7.
42. Doll R, Peto R. The causes of cancer: quantitative estimates of avoidable risks of cancer in the United States today. *J Natl Cancer Inst* 1981; **66**(6): 1191-308.
43. Sakata R, McGale P, Grant EJ, Ozasa K, Peto R, Darby SC. Impact of smoking on mortality and life expectancy in Japanese smokers: a prospective cohort study. *BMJ* 2012; **345**: e7093.
44. Yang JJ, Yu D, Wen W, et al. Tobacco Smoking and Mortality in Asia: A Pooled Meta-analysis. *JAMA Netw Open* 2019; **2**(3): e191474.
45. Jha P, Chaloupka FJ. The economics of global tobacco control. *BMJ* 2000; **321**(7257): 358-61.
46. Norheim OF, Jha P, Admasu K, et al. Avoiding 40% of the premature deaths in each country, 2010-30: review of national mortality trends to help quantify the UN sustainable development goal for health. *Lancet* 2015; **385**(9964): 239-52.

47. Harris F, MacKintosh AM, Anderson S, et al. Effects of the 2003 advertising/promotion ban in the United Kingdom on awareness of tobacco marketing: findings from the International Tobacco Control (ITC) Four Country Survey. *Tob Control* 2006; **15 Suppl 3**: iii26-33.
48. Saffer H, Chaloupka F. The effect of tobacco advertising bans on tobacco consumption. *J Health Econ* 2000; **19**(6): 1117-37.
49. Hammond D. Health warning messages on tobacco products: a review. *Tob Control* 2011; **20**(5): 327-37.
50. Peto R, Darby S, Deo H, Silcocks P, Whitley E, Doll R. Smoking, smoking cessation, and lung cancer in the UK since 1950: combination of national statistics with two case-control studies. *BMJ* 2000; **321**(7257): 323-9.
51. IHME. Tobacco Viz. 2018. https://vizhub.healthdata.org/tobacco/.
52. Roemer R, Taylor A, Lariviere J. Origins of the WHO Framework Convention on Tobacco Control. *Am J Public Health* 2005; **95**(6): 936-8.
53. Shibuya K, Ciecierski C, Guindon E, et al. WHO Framework Convention on Tobacco Control: development of an evidence based global public health treaty. *BMJ* 2003; **327**(7407): 154-7.
54. World Health Organisation. WHO Report on the Global Tobacco Epidemic, 2008: The MPOWER package. Geneva, 2008.
55. Jha P, Chaloupka FJ, Corrao M, Jacob B. Reducing the burden of smoking world-wide: effectiveness of interventions and their coverage. *Drug Alcohol Rev* 2006; **25**(6): 597-609.
56. WHO. WHO Report on the Global Tobacco Epidemic, 2015: World Health Organization, 2015.
57. Gravely S, Giovino GA, Craig L, et al. Implementation of key demand-reduction measures of the WHO Framework Convention on Tobacco Control and change in smoking prevalence in 126 countries: an association study. *Lancet Public Health* 2017; **2**(4): e166-e74.
58. Levy DT, Li Y, Yuan Z. Impact of nations meeting the MPOWER targets between 2014 and 2016: an update. *Tob Control* 2019.
59. Nishtar S, Niinisto S, Sirisena M, et al. Time to deliver: report of the WHO Independent High-Level Commission on NCDs. *Lancet* 2018; **392**(10143): 245-52.
60. Ghebreyesus TA. Progress in beating the tobacco epidemic. *Lancet* 2019.
61. GBD Risk Factors Collaborators. Global, regional, and national comparative risk assessment of 84 behavioural, environmental and occupational, and metabolic risks or clusters of risks, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet* 2017; **390**(10100): 1345-422.
62. Jha P, Jacob B, Gajalakshmi V, et al. A nationally representative case-control study of smoking and death in India. *N Engl J Med* 2008; **358**(11): 1137-47.
63. Alam DS, Jha P, Ramasundarahettige C, et al. Smoking-attributable mortality in Bangladesh: proportional mortality study. *Bull World Health Organ* 2013; **91**(10): 757-64.
64. Sitas F, Egger S, Bradshaw D, et al. Differences among the coloured, white, black, and other South African populations in smoking-attributed mortality at ages 35-74 years: a case-control study of 481,640 deaths. *Lancet* 2013; **382**(9893): 685-93.
65. Asma S, Lozano R, Chatterji S, et al. Monitoring the health-related Sustainable Development Goals: lessons learned and recommendations for improved measurement. *Lancet* 2020; **395**(10219): 240-6.
66. Chen ZM, Peto R, Iona A, et al. Emerging tobacco-related cancer risks in China: A nationwide, prospective study of 0.5 million adults. *Cancer* 2015; **121 Suppl 17**: 3097-106.
67. Gajalakshmi V, Peto R, Kanimozhi VC, Whitlock G, Veeramani D. Cohort Profile: the Chennai prospective study of mortality among 500,000 adults in Tamil Nadu, South India. *Int J Epidemiol* 2007; **36**(6): 1190-5.
68. Zaridze D, Lewington S, Boroda A, et al. Alcohol and mortality in Russia: prospective observational study of 151,000 adults. *Lancet* 2014; **383**(9927): 1465-73.
69. Rossi M, Negri E, La Vecchia C, Campos H. Smoking habits and the risk of non-fatal acute myocardial infarction in Costa Rica. *Eur J Cardiovasc Prev Rehabil* 2011; **18**(3): 467-74.

70. Sotos-Prieto M, Baylin A, Campos H, Qi L, Mattei J. Lifestyle Cardiovascular Risk Score, Genetic Risk Score, and Myocardial Infarction in Hispanic/Latino Adults Living in Costa Rica. *J Am Heart Assoc* 2016; **5**(12).
71. Piegas LS, Avezum A, Pereira JC, et al. Risk factors for myocardial infarction in Brazil. *Am Heart J* 2003; **146**(2): 331-8.
72. Lanas F, Avezum A, Bautista LE, et al. Risk factors for acute myocardial infarction in Latin America: the INTERHEART Latin American study. *Circulation* 2007; **115**(9): 1067-74.
73. Ciruzzi M, Schargrodsky H, Pramparo P, et al. Attributable risks for acute myocardial infarction in four countries of Latin America. *Medicina (B Aires)* 2003; **63**(6): 697-703.
74. Moura MA, Bergmann A, Aguiar SS, Thuler LC. The magnitude of the association between smoking and the risk of developing cancer in Brazil: a multicenter study. *BMJ Open* 2014; **4**(2): e003736.
75. Kfouri SA, Eluf Neto J, Koifman S, et al. Fraction of head and neck cancer attributable to tobacco and alcohol in cities of three Brazilian regions. *Rev Bras Epidemiol* 2018; **21**: e180005.
76. Boing AF, Antunes JL, de Carvalho MB, et al. How much do smoking and alcohol consumption explain socioeconomic inequalities in head and neck cancer risk? *J Epidemiol Community Health* 2011; **65**(8): 709-14.
77. Ferreira Antunes JL, Toporcov TN, Biazevic MG, Boing AF, Scully C, Petti S. Joint and independent effects of alcohol drinking and tobacco smoking on oral cancer: a large case-control study. *PLoS One* 2013; **8**(7): e68132.
78. Szymanska K, Hung RJ, Wunsch-Filho V, et al. Alcohol and tobacco, and the risk of cancers of the upper aerodigestive tract in Latin America: a case-control study. *Cancer Causes Control* 2011; **22**(7): 1037-46.
79. Sanchez-Zamorano LM, Flores-Luna L, Angeles-Llerenas A, et al. Healthy lifestyle on the risk of breast cancer. *Cancer Epidemiol Biomarkers Prev* 2011; **20**(5): 912-22.
80. Llibre-Guerra JC, Valhuerdi Cepero A, Fernandez Concepcion O, Llibre-Guerra JJ, Gutierrez RF, Llibre-Rodriguez JJ. Stroke incidence and risk factors in Havana and Matanzas, Cuba. *Neurologia* 2015; **30**(8): 488-95.
81. de Oliveira C, Marmot MG, Demakakos P, Vaz de Melo Mambrini J, Peixoto SV, Lima-Costa MF. Mortality risk attributable to smoking, hypertension and diabetes among English and Brazilian older adults (The ELSA and Bambui cohort ageing studies). *Eur J Public Health* 2016; **26**(5): 831-5.
82. Otiniano ME, Ottenbacher KJ, Markides KS, Ray LA, Du XL. Self-reported heart attack in Mexican-American elders: examination of incidence, prevalence, and 7-year mortality. *J Am Geriatr Soc* 2003; **51**(7): 923-9.
83. Wei M, Mitchell BD, Haffner SM, Stern MP. Effects of cigarette smoking, diabetes, high cholesterol, and hypertension on all-cause mortality and cardiovascular disease mortality in Mexican Americans. The San Antonio Heart Study. *Am J Epidemiol* 1996; **144**(11): 1058-65.
84. Sorlie PD, Garcia-Palmieri MR, Costas R, Cruz-Vidal M, Havlik R. Cigarette smoking and coronary heart disease in Puerto Rico. *Prev Med* 1982; **11**(3): 304-16.
85. Kittner SJ, Garcia-Palmieri MR, Costas R, Jr., Cruz-Vidal M, Abbott RD, Havlik RJ. Alcohol and coronary heart disease in Puerto Rico. *Am J Epidemiol* 1983; **117**(5): 538-50.
86. Armas Rojas N, Lacey B, Lewington S, et al. Cohort Profile: the Cuba Prospective Study. *Int J Epidemiol* 2019. DOI: 10.1093/ije/dyy297.
87. Institute for Health Metrics and Evaluation. Global Burden of Disease (GBD). http://www.healthdata.org/gbd (accessed August 5 2019).
88. Cooper RS, Kennelly JF, Ordunez-Garcia P. Health in Cuba. *Int J Epidemiol* 2006; **35**(4): 817-24.
89. Global BMI Mortality Collaboration, Di Angelantonio E, Bhupathiraju Sh N, et al. Body-mass index and all-cause mortality: individual-participant-data meta-analysis of 239 prospective studies in four continents. *Lancet* 2016; **388**(10046): 776-86.

90. Asvold BO, Bjorngaard JH, Carslake D, et al. Causal associations of tobacco smoking with cardiovascular risk factors: a Mendelian randomization analysis of the HUNT Study in Norway. *Int J Epidemiol* 2014; **43**(5): 1458-70.
91. Jha P. Reliable direct measurement of causes of death in low- and middle-income countries. *BMC Med* 2014; **12**: 19.
92. Plummer M. Improved estimates of floating absolute risk. *Stat Med* 2004; **23**(1): 93-104.
93. Gnatiuc L, Alegre-Diaz J, Wade R, et al. General and Abdominal Adiposity and Mortality in Mexico City: Prospective Study of 150 000 Adults. *Ann Intern Med* 2019.
94. Bonet M. New survey results enhance Cuba's NCD surveillance: Mariano Bonet MD, Director, National Hygiene, Epidemiology and Microbiology Institute. *MEDICC Rev* 2011; **13**(4): 11-3.
95. Franco M, Bilal U, Ordunez P, et al. Population-wide weight loss and regain in relation to diabetes burden and cardiovascular mortality in Cuba 1980-2010: repeated cross sectional surveys and ecological comparison of secular trends. *BMJ* 2013; **346**: f1515.
96. Bonet M, Varona P. III Encuesta nacional de factores de riesgo y actividades preventivas de enfermedades no transmisibles. Cuba 2010-2011. La Habana, Cuba: Editorial Ciencias Medica; 2015.
97. Lugo NS. Consumo, precio y segmentación del mercado de cigarrillos. Cuba, 2017. *Revista Cubana de Salud Pública* 2018; **44**(4).
98. Choi SH, Stommel M. Impact of Age at Smoking Initiation on Smoking-Related Morbidity and All-Cause Mortality. *Am J Prev Med* 2017; **53**(1): 33-41.
99. Huxley RR, Yatsuya H, Lutsey PL, Woodward M, Alonso A, Folsom AR. Impact of age at smoking initiation, dosage, and time since quitting on cardiovascular disease in african americans and whites: the atherosclerosis risk in communities study. *Am J Epidemiol* 2012; **175**(8): 816-26.
100. Christensen CH, Rostron B, Cosgrove C, et al. Association of Cigarette, Cigar, and Pipe Use With Mortality Risk in the US Population. *JAMA Intern Med* 2018; **178**(4): 469-76.
101. Chang CM, Corey CG, Rostron BL, Apelberg BJ. Systematic review of cigar smoking and all cause and smoking related mortality. *BMC Public Health* 2015; **15**: 390.
102. Tapia-Conyer R, Kuri-Morales P, Alegre-Diaz J, et al. Cohort profile: the Mexico City Prospective Study. *Int J Epidemiol* 2006; **35**(2): 243-9.
103. Alegre-Diaz J, Herrington W, Lopez-Cervantes M, et al. Diabetes and Cause-Specific Mortality in Mexico City. *N Engl J Med* 2016; **375**(20): 1961-71.
104. Herrington WG, Alegre-Diaz J, Wade R, et al. Effect of diabetes duration and glycaemic control on 14-year cause-specific mortality in Mexican adults: a blood-based prospective cohort study. *Lancet Diabetes Endocrinol* 2018; **6**(6): 455-63.
105. Encuesta Nacional de Salud 2000. Cuernavaca, Mexico: Instituto Nacional de Salud Publica, 2003.
106. Giovino GA, Mirza SA, Samet JM, et al. Tobacco use in 3 billion individuals from 16 countries: an analysis of nationally representative cross-sectional household surveys. *Lancet* 2012; **380**(9842): 668-79.
107. Lajous M, Ortiz-Panozo E, Monge A, et al. Cohort Profile: The Mexican Teachers' Cohort (MTC). *Int J Epidemiol* 2015.
108. Inoue-Choi M, Liao LM, Reyes-Guzman C, Hartge P, Caporaso N, Freedman ND. Association of Long-term, Low-Intensity Smoking With All-Cause and Cause-Specific Mortality in the National Institutes of Health-AARP Diet and Health Study. *JAMA Intern Med* 2017; **177**(1): 87-95.
109. Ng M, Freeman MK, Fleming TD, et al. Smoking prevalence and cigarette consumption in 187 countries, 1980-2012. *JAMA* 2014; **311**(2): 183-92.
110. Reynales-Shigematsu LM, Fleischer NL, Thrasher JF, et al. Effects of tobacco control policies on smoking prevalence and tobacco-attributable deaths in Mexico: the SimSmoke model. *Rev Panam Salud Publica* 2015; **38**(4): 316-25.
111. Reynales-Shigematsu LM, Guerrero-Lopez CM, Hernandez Avila M, Irving H, Jha P. Divergence and convergence in cause-specific premature adult mortality in Mexico and US Mexican Hispanics from 1995 to 2015: analyses of 4.9 million individual deaths. *Int J Epidemiol* 2017.

112. Carter BD, Freedman ND, Jacobs EJ. Smoking and mortality--beyond established causes. *N Engl J Med* 2015; **372**(22): 2170.
113. Centers for Disease Control and Prevention. National Health Interview Survey. 2019. https://www.cdc.gov/nchs/nhis/index.htm (accessed February 27 2019).
114. Survey Description, National Health Interview Survey (NHIS) Public Use Data Release. Hyattsville, Maryland: National Center for Health Statistics, 2015.
115. The Linkage of National Center for Health Statistics Survey Data to the National Death Index - 2015 Linked Mortality File (LMF): Methodology Overview and Analytic Considerations. Hyattsville, Maryland: National Center for Health Statistics, 2019.
116. Lochner K, Hummer RA, Bartee S, Wheatcroft G, Cox C. The public-use National Health Interview Survey linked mortality files: methods of reidentification risk avoidance and comparative analysis. *Am J Epidemiol* 2008; **168**(3): 336-44.
117. Connor Gorber S, Schofield-Hurwitz S, Hardt J, Levasseur G, Tremblay M. The accuracy of self-reported smoking: a systematic review of the relationship between self-reported and cotinine-assessed smoking status. *Nicotine Tob Res* 2009; **11**(1): 12-24.
118. Baltar VT, Xun WW, Chuang SC, et al. Smoking, secondhand smoke, and cotinine levels in a subset of EPIC cohort. *Cancer Epidemiol Biomarkers Prev* 2011; **20**(5): 869-75.
119. Vartiainen E, Seppala T, Lillsunde P, Puska P. Validation of self reported smoking by serum cotinine measurement in a community-based study. *J Epidemiol Community Health* 2002; **56**(3): 167-70.
120. Rodriquez EJ, Oh SS, Perez-Stable EJ, Schroeder SA. Changes in Smoking Intensity Over Time by Birth Cohort and by Latino National Background, 1997-2014. *Nicotine Tob Res* 2016; **18**(12): 2225-33.
121. Kaplan RC, Bangdiwala SI, Barnhart JM, et al. Smoking among U.S. Hispanic/Latino adults: the Hispanic community health study/study of Latinos. *Am J Prev Med* 2014; **46**(5): 496-506.
122. Jain RB. Trends in serum cotinine concentrations among daily cigarette smokers: data from NHANES 1999-2010. *Sci Total Environ* 2014; **472**: 72-7.
123. Fenelon A. Revisiting the Hispanic mortality advantage in the United States: the role of smoking. *Soc Sci Med* 2013; **82**: 1-9.
124. Warren CW, Lea V, Lee J, Jones NR, Asma S, McKenna M. Change in tobacco use among 13-15 year olds between 1999 and 2008: findings from the Global Youth Tobacco Survey. *Glob Health Promot* 2009; **16**(2 Suppl): 38-90.
125. Xi B, Liang Y, Liu Y, et al. Tobacco use and second-hand smoke exposure in young adolescents aged 12-15 years: data from 68 low-income and middle-income countries. *Lancet Glob Health* 2016; **4**(11): e795-e805.
126. Kamsa-Ard S, Promthet S, Lewington S, et al. Association between smoking and mortality: Khon Kaen cohort study, Thailand. *Asian Pac J Cancer Prev* 2013; **14**(4): 2643-7.
127. Etemadi A, Khademi H, Kamangar F, et al. Hazards of cigarettes, smokeless tobacco and waterpipe in a Middle Eastern Population: a Cohort Study of 50 000 individuals from Iran. *Tob Control* 2017; **26**(6): 674-82.
128. Sitas F, Urban M, Bradshaw D, Kielkowski D, Bah S, Peto R. Tobacco attributable deaths in South Africa. *Tob Control* 2004; **13**(4): 396-9.
129. Mbulo L, Palipudi KM, Andes L, et al. Secondhand smoke exposure at home among one billion children in 21 countries: findings from the Global Adult Tobacco Survey (GATS). *Tob Control* 2016; **25**(e2): e95-e100.
130. Palipudi KM, Gupta PC, Sinha DN, et al. Social determinants of health and tobacco use in thirteen low and middle income countries: evidence from Global Adult Tobacco Survey. *PLoS One* 2012; **7**(3): e33466.
131. Heatherton TF, Kozlowski LT, Frecker RC, Rickert W, Robinson J. Measuring the heaviness of smoking: using self-reported time to the first cigarette of the day and number of cigarettes smoked per day. *Br J Addict* 1989; **84**(7): 791-9.

132. Transdisciplinary Tobacco Use Research Center Tobacco D, Baker TB, Piper ME, et al. Time to first cigarette in the morning as an index of ability to quit smoking: implications for nicotine dependence. *Nicotine Tob Res* 2007; **9 Suppl 4**: S555-70.
133. Jamal A, Phillips E, Gentzke AS, et al. Current Cigarette Smoking Among Adults - United States, 2016. *MMWR Morb Mortal Wkly Rep* 2018; **67**(2): 53-9.
134. Adult smoking habits in the UK: 2018. 2 July 2019. https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/healthandlifeexpectancies/bulletins/adultsmokinghabitsingreatbritain/2018 (accessed 1 October 2019).
135. National Drug Strategy Household Survey (NDSHS) 2016 key findings data tables. http://www.aihw.gov.au/alcohol-and-other-drugs/data-sources/ndshs-2016/data/ (accessed 1 October 2019 2019).
136. Fry A, Littlejohns TJ, Sudlow C, et al. Comparison of Sociodemographic and Health-Related Characteristics of UK Biobank Participants With Those of the General Population. *Am J Epidemiol* 2017; **186**(9): 1026-34.
137. Lindsted KD, Fraser GE, Steinkohl M, Beeson WL. Healthy volunteer effect in a cohort study: temporal resolution in the Adventist Health Study. *J Clin Epidemiol* 1996; **49**(7): 783-90.
138. Benowitz NL. Biomarkers of environmental tobacco smoke exposure. *Environ Health Perspect* 1999; **107 Suppl 2**: 349-55.
139. Benowitz NL, Dains KM, Dempsey D, Havel C, Wilson M, Jacob P, 3rd. Urine menthol as a biomarker of mentholated cigarette smoking. *Cancer Epidemiol Biomarkers Prev* 2010; **19**(12): 3013-9.
140. Jacob P, 3rd, Goniewicz ML, Havel CM, Schick SF, Benowitz NL. Nicotelline: a proposed biomarker and environmental tracer for particulate matter derived from tobacco smoke. *Chem Res Toxicol* 2013; **26**(11): 1615-31.
141. Benowitz NL, Bernert JT, Caraballo RS, Holiday DB, Wang J. Optimal serum cotinine levels for distinguishing cigarette smokers and nonsmokers within different racial/ethnic groups in the United States between 1999 and 2004. *Am J Epidemiol* 2009; **169**(2): 236-48.
142. Campuzano JC, Hernandez-Avila M, Jaakkola MS, et al. Determinants of salivary cotinine levels among current smokers in Mexico. *Nicotine Tob Res* 2004; **6**(6): 997-1008.
143. Pichini S, Basagana XB, Pacifici R, et al. Cord serum cotinine as a biomarker of fetal exposure to cigarette smoke at the end of pregnancy. *Environ Health Perspect* 2000; **108**(11): 1079-83.
144. Klesges RC, Debon M, Ray JW. Are self-reports of smoking rate biased? Evidence from the Second National Health and Nutrition Examination Survey. *J Clin Epidemiol* 1995; **48**(10): 1225-33.
145. Verde Z, Reinoso-Barbero L, Chicharro L, et al. Effects of cigarette smoking and nicotine metabolite ratio on leukocyte telomere length. *Environ Res* 2015; **140**: 488-94.
146. Xu T, Holzapfel C, Dong X, et al. Effects of smoking and smoking cessation on human serum metabolite profile: results from the KORA cohort study. *BMC Med* 2013; **11**: 60.
147. Clarke R, Bennett DA. Commentary: Mendelian randomization study of tobacco smoking and cardiovascular risk factors: hazards of tobacco smoking greatly outweigh any benefits. *Int J Epidemiol* 2014; **43**(5): 1471-2.
148. Linneberg A, Jacobsen RK, Skaaby T, et al. Effect of Smoking on Blood Pressure and Resting Heart Rate: A Mendelian Randomization Meta-Analysis in the CARTA Consortium. *Circ Cardiovasc Genet* 2015; **8**(6): 832-41.
149. Gentzke AS, Creamer M, Cullen KA, et al. Vital Signs: Tobacco Product Use Among Middle and High School Students - United States, 2011-2018. *MMWR Morb Mortal Wkly Rep* 2019; **68**(6): 157-64.
150. Soneji S, Barrington-Trimis JL, Wills TA, et al. Association Between Initial Use of e-Cigarettes and Subsequent Cigarette Smoking Among Adolescents and Young Adults: A Systematic Review and Meta-analysis. *JAMA Pediatr* 2017; **171**(8): 788-97.

151. Miech R, Patrick ME, O'Malley PM, Johnston LD. E-cigarette use as a predictor of cigarette smoking: results from a 1-year follow-up of a national sample of 12th grade students. *Tob Control* 2017; **26**(e2): e106-e11.
152. Centers for Disease Control and Prevention. Outbreak of Lung Injury Associated with E-Cigarette Use, or Vaping. September 19, 2019 (accessed September 25 2019).
153. Perrine CG, Pickens CM, Boehmer TK, et al. Characteristics of a Multistate Outbreak of Lung Injury Associated with E-cigarette Use, or Vaping - United States, 2019. *MMWR Morb Mortal Wkly Rep* 2019; **68**(39): 860-4.
154. Kennedy RD, Awopegba A, De Leon E, Cohen JE. Global approaches to regulating electronic cigarettes. *Tob Control* 2017; **26**(4): 440-5.
155. Thrasher JF, Abad-Vivero EN, Barrientos-Gutierrez I, et al. Prevalence and Correlates of E-Cigarette Perceptions and Trial Among Early Adolescents in Mexico. *J Adolesc Health* 2016; **58**(3): 358-65.
156. Palipudi KM, Mbulo L, Morton J, et al. Awareness and Current Use of Electronic Cigarettes in Indonesia, Malaysia, Qatar, and Greece: Findings From 2011-2013 Global Adult Tobacco Surveys. *Nicotine Tob Res* 2016; **18**(4): 501-7.
157. Vardavas CI, Filippidis FT, Agaku IT. Determinants and prevalence of e-cigarette use throughout the European Union: a secondary analysis of 26 566 youth and adults from 27 Countries. *Tob Control* 2015; **24**(5): 442-8.
158. Liu X, Bragg F, Yang L, et al. Smoking and smoking cessation in relation to risk of diabetes in Chinese men and women: a 9-year prospective study of 0.5 million people. *Lancet Public Health* 2018; **3**(4): e167-e76.
159. Will JC, Galuska DA, Ford ES, Mokdad A, Calle EE. Cigarette smoking and diabetes mellitus: evidence of a positive association from a large prospective cohort study. *Int J Epidemiol* 2001; **30**(3): 540-6.
160. Tomlinson M, Solomon W, Singh Y, et al. The use of mobile phones as a data collection tool: a report from a household survey in South Africa. *BMC Med Inform Decis Mak* 2009; **9**: 51.
161. Schuz J, Elliott P, Auvinen A, et al. An international prospective cohort study of mobile phone users and health (Cosmos): design considerations and enrolment. *Cancer Epidemiol* 2011; **35**(1): 37-43.
162. Collins R. What makes UK Biobank special? *Lancet* 2012; **379**(9822): 1173-4.
163. Barr ML, van Ritten JJ, Steel DG, Thackway SV. Inclusion of mobile phone numbers into an ongoing population health survey in New South Wales, Australia: design, methods, call outcomes, costs and sample representativeness. *BMC Med Res Methodol* 2012; **12**: 177.
164. Zaridze D, Brennan P, Boreham J, et al. Alcohol and cause-specific mortality in Russia: a retrospective case-control study of 48,557 adult deaths. *Lancet* 2009; **373**(9682): 2201-14.
165. Millwood IY, Walters RG, Mei XW, et al. Conventional and genetic evidence on alcohol and vascular disease aetiology: a prospective study of 500 000 men and women in China. *Lancet* 2019; **393**(10183): 1831-42.
166. Gajalakshmi V, Lacey B, Kanimozhi V, Sherliker P, Peto R, Lewington S. Body-mass index, blood pressure, and cause-specific mortality in India: a prospective cohort study of 500 810 adults. *Lancet Glob Health* 2018; **6**(7): e787-e94.
167. Doll R. Cancer in five continents. *Proc R Soc Med* 1972; **65**(1): 49-55.
168. Palipudi KM, Morton J, Hsia J, et al. Methodology of the Global Adult Tobacco Survey - 2008-2010. *Glob Health Promot* 2016; **23**(2 Suppl): 3-23.

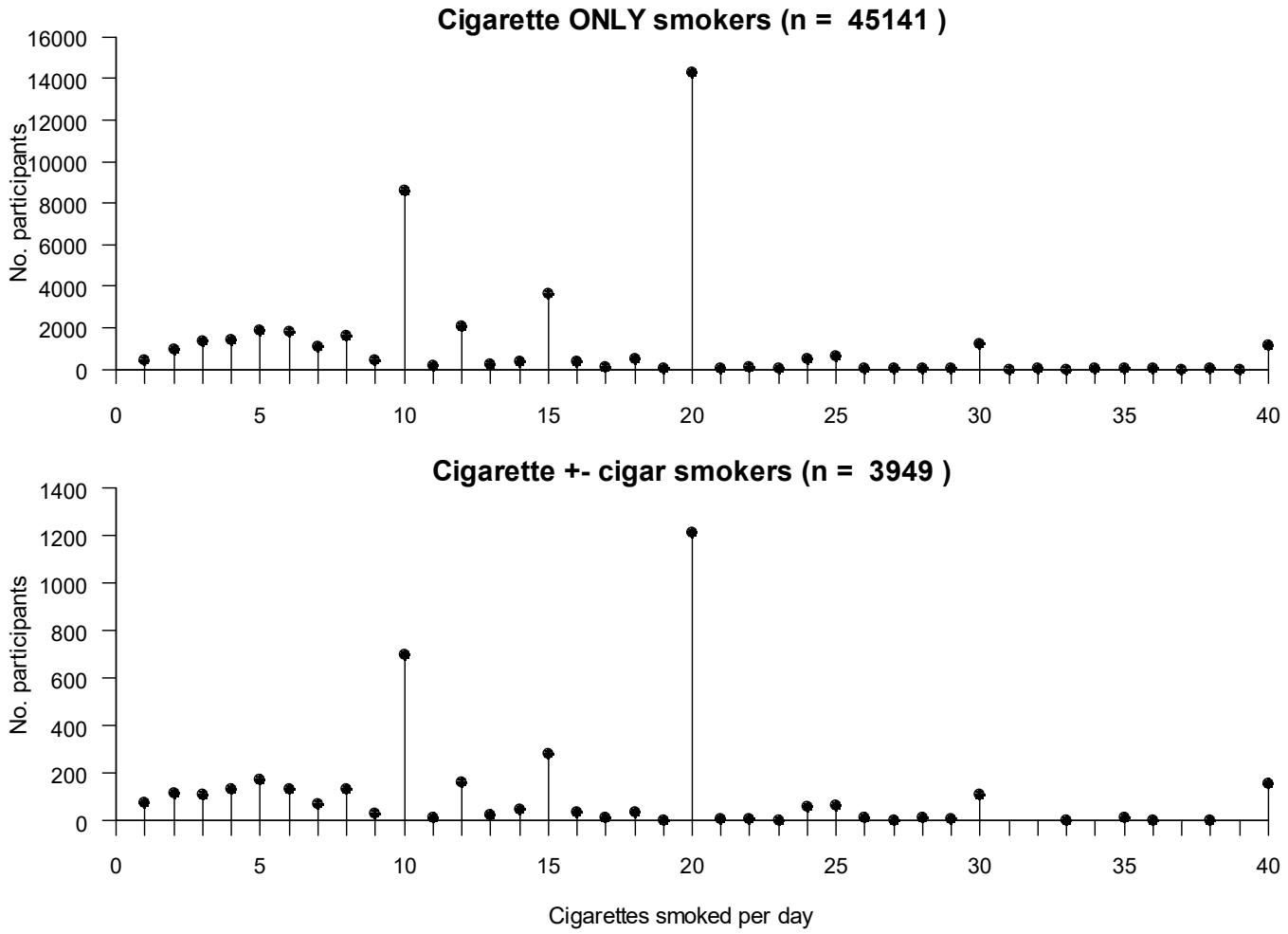
Appendix A: Additional results for the Cuba Prospective Study

Table A1: Baseline characteristics of the participants aged 70-89 at recruitment, by sex and smoking status

	Never smoker	Ex smoker	Smoker of cigars only	Cigarette smoker, by age started				
				Any	5-9	10-14	15-19	20+
Participants	15 112	2105	1224	4550	346	1695	1551	958
Age, years	77 (5)	77 (5)	77 (5)	76 (5)	76 (5)	77 (5)	76 (5)	76 (5)
Male (%)	32%	66%	89%	67%	77%	73%	67%	50%
Completed high school	10%	11%	6%	12%	6%	10%	15%	14%
Consume alcohol weekly	3%	7%	12%	15%	15%	17%	16%	12%
BMI kg/m²	24 (4)	24 (4)	24 (4)	23 (4)	23 (5)	23 (4)	23 (4)	24 (5)
SBP (mmHg)	130 (18)	130 (18)	130 (17)	128 (17)	128 (18)	129 (17)	128 (16)	129 (17)
Smoking at entry								
Age started smoking (years)	-	18 (9)	16 (8)	16 (7)	8 (1)	12 (1)	16 (1)	26 (10)
Cigarettes/day*	-	-	-	13 (9)	15 (10)	15 (9)	13 (8)	11 (8)
Cigar smokers*	-	-	100%	16%	22%	19%	15%	12%
Smoking at resurvey†								
Current smokers at resurvey	2%	20%	83%	80%	100%	67%	76%	91%
Cigarettes/day among those still smoking	-	-	-	12 (8)	17 (7)	15 (11)	10 (7)	9 (4)
Cigarettes/day among all resurveyed	0 (2)	1 (4)	2 (8)	9 (8)	17 (7)	9 (11)	8 (8)	7 (5)

Percent or mean (SD) shown.
Excludes those age <70 at recruitment and those with missing information on smoking or covariates.
† 486 participants were resurveyed about 6 years after baseline. At baseline, 400 participants reported never smoking, 19 ex-smoking, 23 cigar smoking only, and 44 cigarette smoking (mean 13/day).
*Smokers of cigars only consumed 2.5 (SD 1) cigars/day, and the few cigarette smokers who also smoked cigars consumed 1.3 (SD 0.5) cigars/day.

Figure A1: Cigarettes smoked per day at recruitment, by product smoked



Note: Y axes are different scales

Figure A2: Cigars smoked per day at recruitment, by product smoked

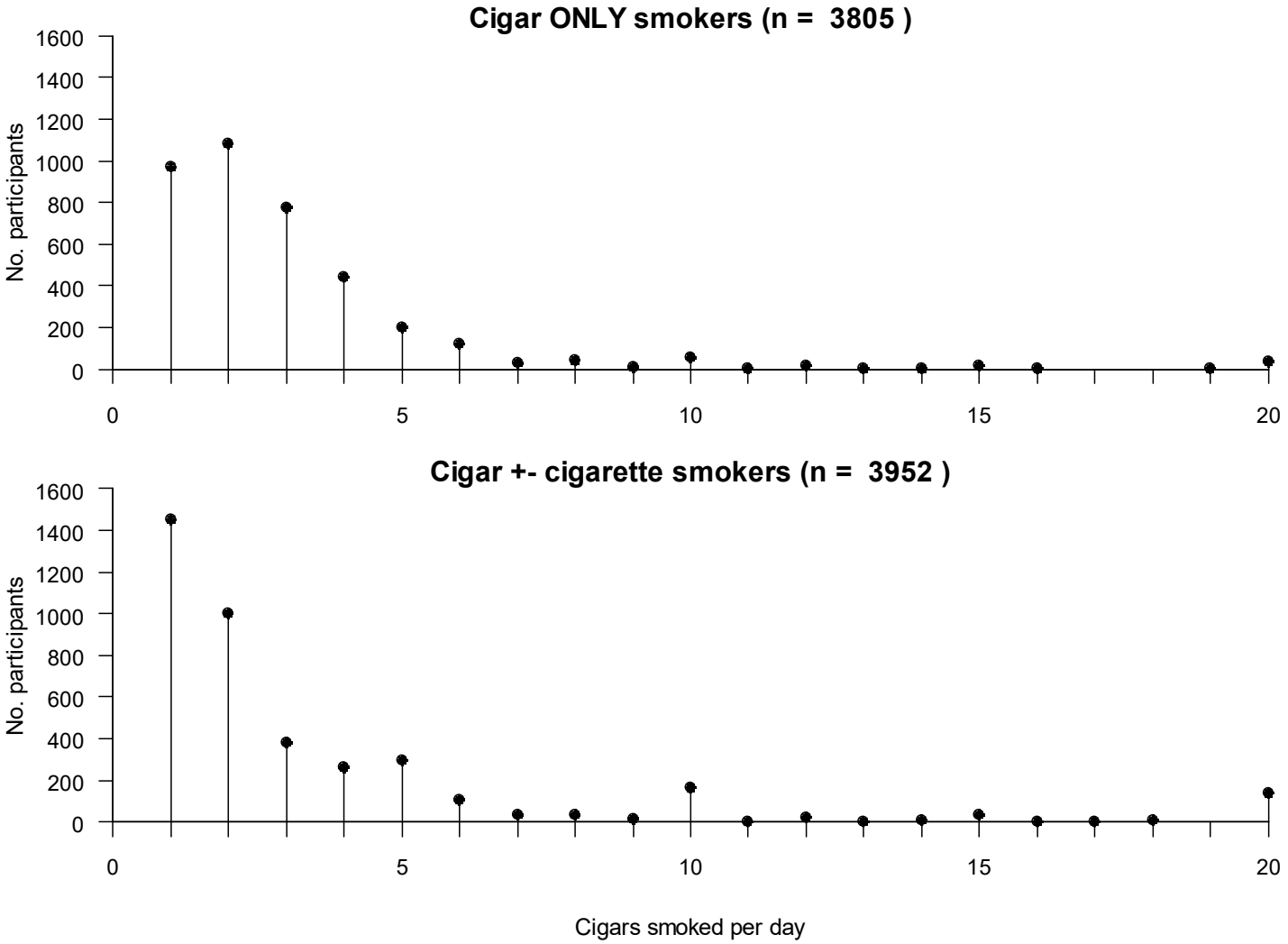


Figure A3: All-cause mortality rate ratio at ages 70-89, cigarette smoker versus never smoker, by age started smoking and usual amount smoked

Participants with chronic disease at recruitment were included in these analyses.

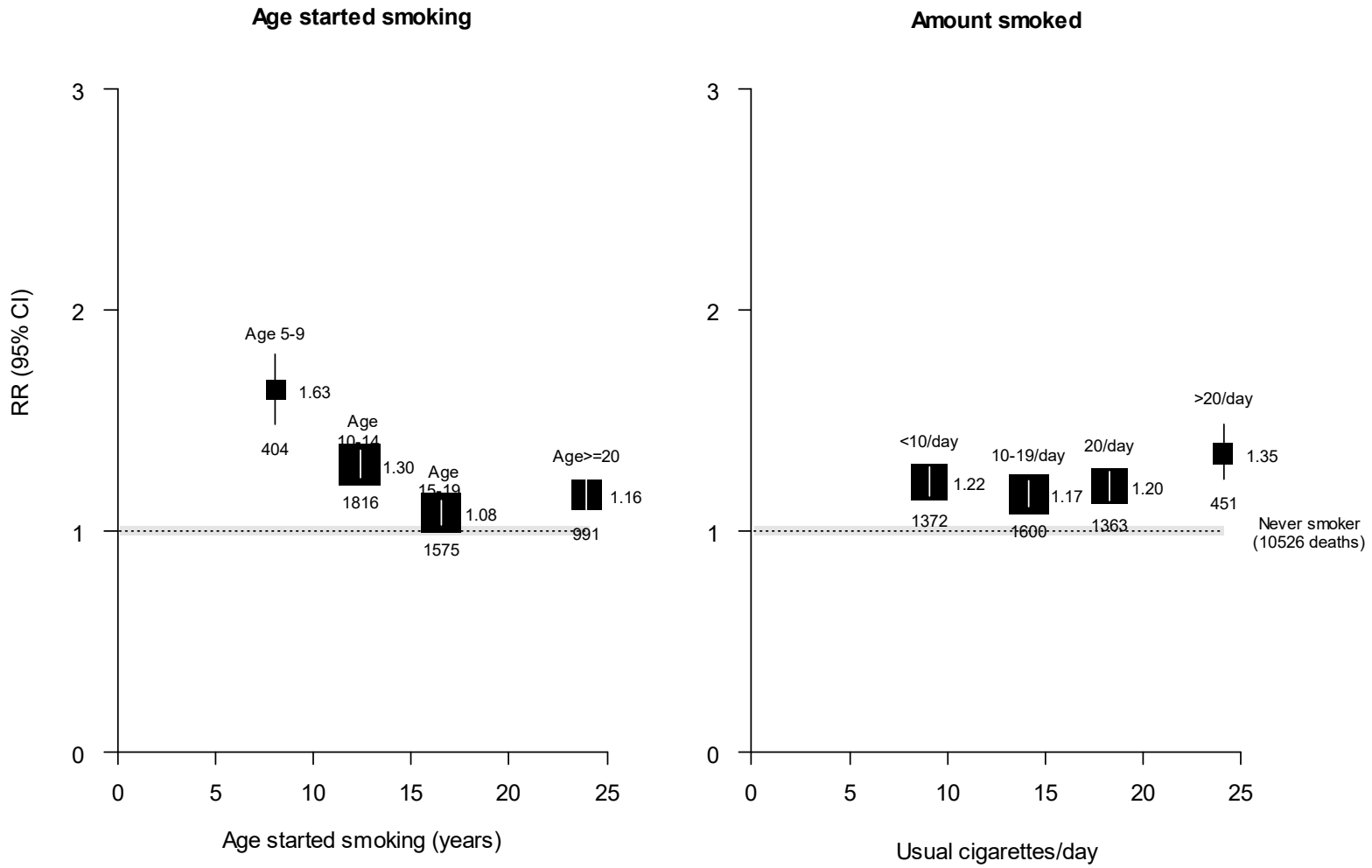


Figure A4: All-cause mortality rate ratio at ages 70-89, cigarette smoker versus never smoker, by age started smoking and usual amount smoked

Participants with chronic disease at recruitment were excluded from these analyses.

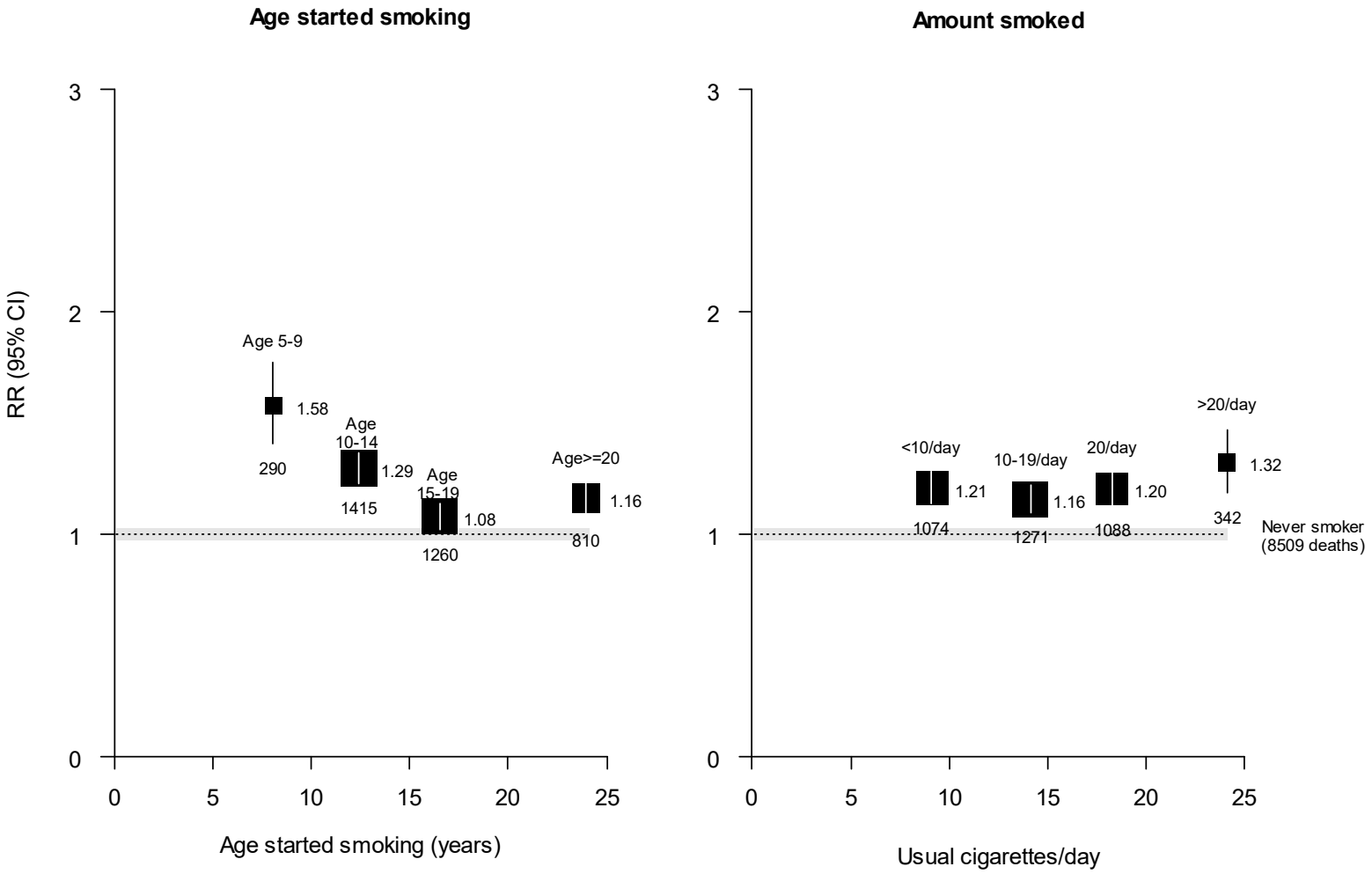
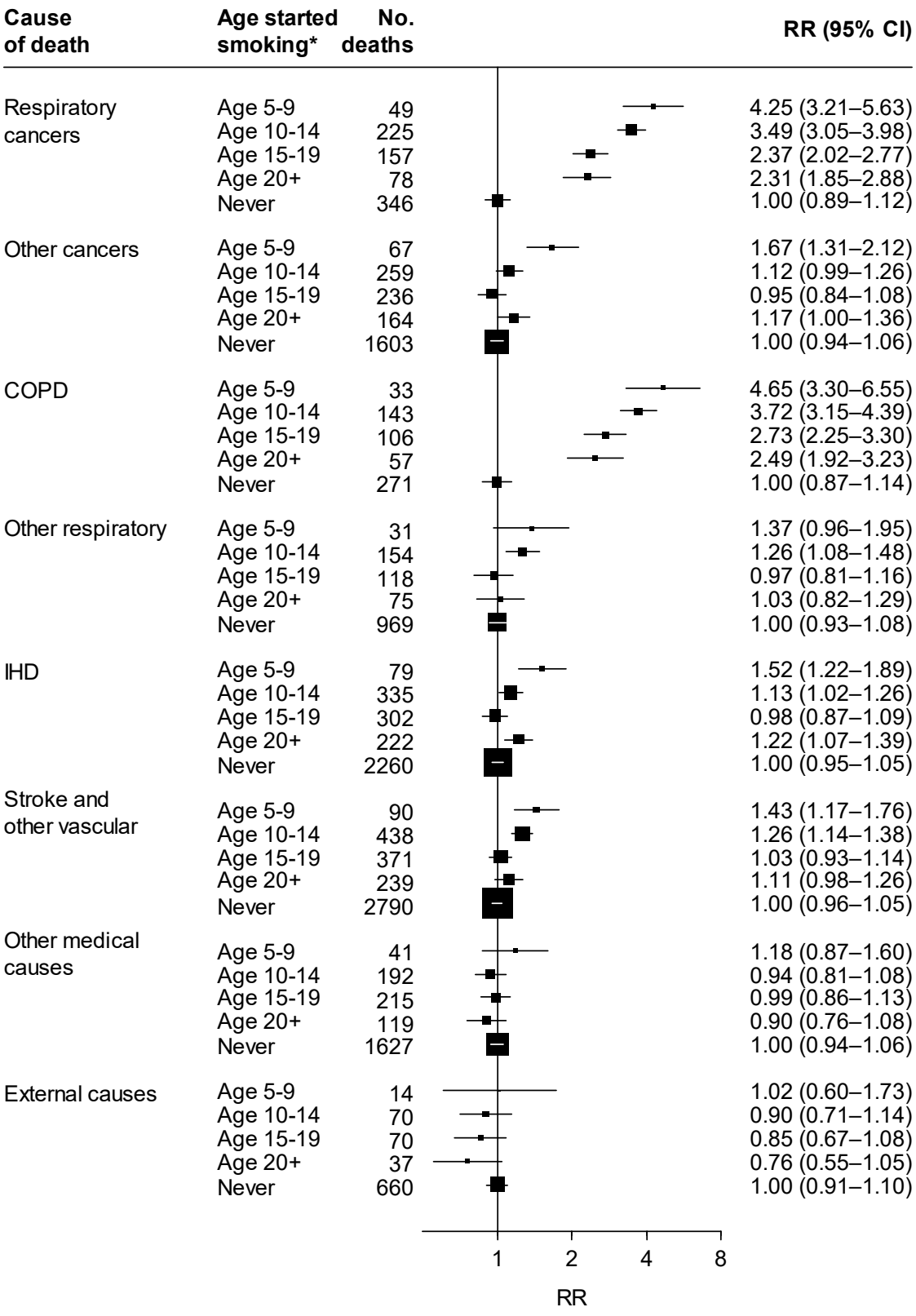
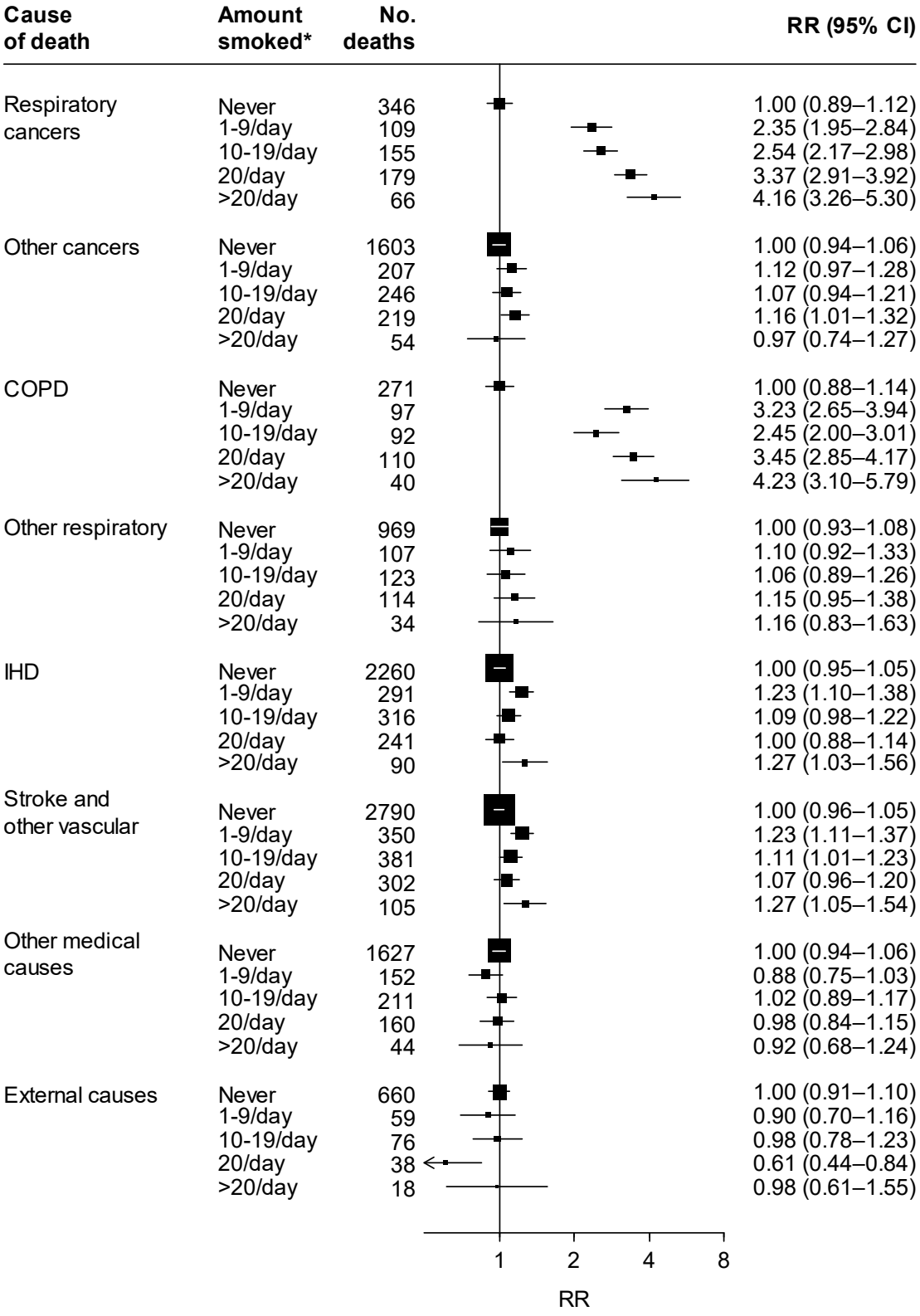


Figure A5: Cause-specific mortality rate ratio at ages 70-89, cigarette versus never smoker, by age started smoking



RRs associated with never smoking, compared to starting to smoke on average at ages 8, 12, 17, and 24 years, respectively.

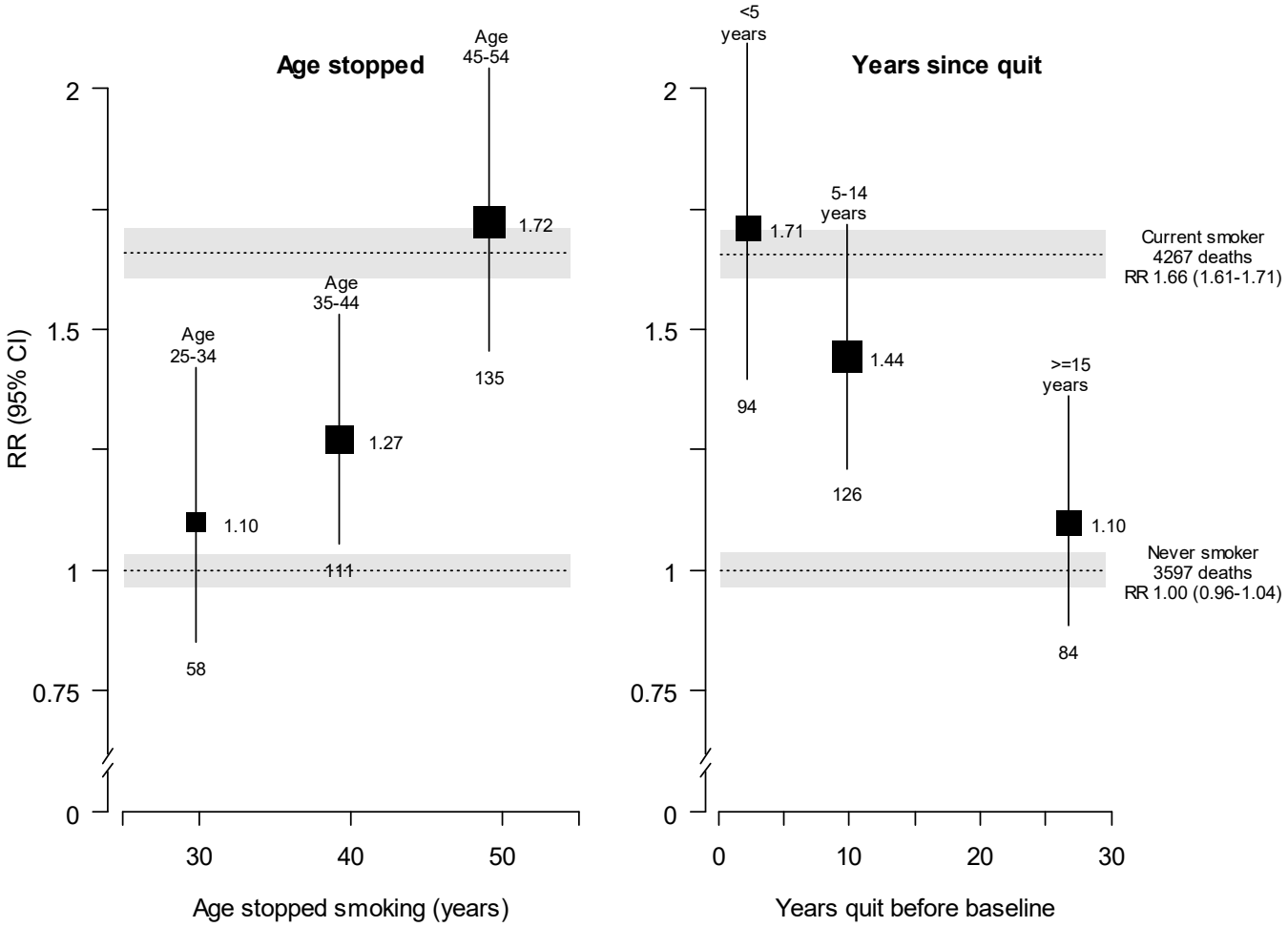
Figure A6: Cause-specific mortality rate ratio at ages 70-89, cigarette versus never smoker, by amount smoked



RRs associated with never smoking and smoking 9, 14, 18, and 24 cigarettes per day on average at resurvey, respectively.

Figure A7: All-cause mortality rate ratio at ages 30-69, cigarette smoker versus never smoker, by age at cessation and years since quitting

Participants with chronic disease at recruitment were included in these analyses.



Appendix B: Additional results for the Mexico City Prospective Study

Table B1: Baseline characteristics of the participants at ages 35-89, by sex and smoking status at recruitment, among those with chronic disease at recruitment

	Men						Women					
	Never	**Ex >3 years**	**Ex 1-3 years**	**< Daily**	**Daily < 10**	**Daily 10+**	**Never**	**Ex >3 years**	**Ex 1-3 years**	**< Daily**	**Daily < 10**	**Daily 10+**
Participants	1707	3724	632	936	1494	828	12 128	3257	518	998	1265	410
Smoking behaviour												
Mean cigarettes/day ‡	-	11.2 (14.4)	11.5 (12.9)	0.2 (0.5)	4.1 (2.1)	17.1 (8.7)	-	5.7 (10.0)	6.1 (8.2)	0.1 (0.2)	3.6 (2.0)	15.7 (7.6)
Age started smoking (years)	-	17.5 (6.3)	17.7 (6.2)	19.4 (7.8)	18.8 (8.2)	16.8 (6.3)	-	23.5 (9.4)	24.7 (11.6)	26.7 (12.4)	23.4 (11.1)	20.5 (8.0)
Ever tried to quit	-	-	-	43%	45%	38%	-	-	-	41%	52%	43%
Inhale smoke	-	-	-	61%	73%	82%	-	-	-	42%	59%	71%
Others smoke inside home	16%	20%	26%	29%	35%	41%	30%	34%	41%	48%	56%	62%
Age, years	61 (12)	64 (12)	58 (12)	55 (11)	57 (11)	56 (11)	61 (12)	60 (12)	55 (11)	52 (10)	53 (12)	53 (11)
Resident in Iztapalapa	64%	60%	56%	62%	58%	54%	65%	59%	55%	65%	57%	53%
Education attained												
University	17%	14%	16%	16%	13%	20%	4%	7%	9%	8%	9%	12%
High school	18%	17%	19%	23%	23%	22%	11%	17%	19%	21%	24%	27%
Elementary school	50%	54%	52%	52%	51%	46%	58%	56%	55%	58%	54%	51%
Other	15%	15%	13%	8%	12%	11%	27%	21%	16%	13%	13%	10%
Drinking behaviour												
Drink at least weekly	42%	40%	42%	53%	46%	43%	42%	46%	49%	60%	55%	52%
Drink less than weekly	36%	48%	46%	35%	42%	44%	18%	31%	31%	22%	20%	26%
Former drinker	6%	6%	6%	10%	9%	8%	2%	3%	5%	6%	6%	5%
Never drinker	16%	7%	6%	3%	3%	4%	38%	20%	14%	12%	18%	17%
Anthropometry and blood pressure												
BMI kg/m²	27.7 (4.3)	27.8 (4.3)	27.5 (5.0)	27.8 (4.3)	27.4 (4.3)	26.9 (4.7)	29.4 (5.3)	30.0 (5.7)	30.0 (5.2)	30.6 (6.0)	29.2 (5.7)	28.7 (5.9)
Waist-to-hip ratio	0.97 (0.07)	0.97 (0.08)	0.97 (0.07)	0.96 (0.06)	0.97 (0.07)	0.96 (0.06)	0.90 (0.07)	0.90 (0.07)	0.90 (0.07)	0.89 (0.07)	0.89 (0.08)	0.89 (0.07)
Systolic BP (mmHg)	133 (18)	134 (18)	133 (18)	131 (17)	132 (17)	131 (17)	135 (19)	134 (19)	130 (18)	130 (17)	130 (19)	129 (17)
Diastolic BP (mmHg)	85 (11)	85 (11)	85 (11)	85 (11)	85 (11)	84 (11)	85 (11)	85 (11)	83 (10)	84 (10)	83 (11)	82 (10)

Table B2: Baseline characteristics of the participants at ages 35-69, by sex and smoking status at recruitment, among those without chronic disease at recruitment)

	Men						Women					
	Never	**Ex >3 years**	**Ex 1-3 years**	**< Daily**	**Daily < 10**	**Daily 10+**	**Never**	**Ex >3 years**	**Ex 1-3 years**	**< Daily**	**Daily < 10**	**Daily 10+**
Participants	7815	9825	2176	5957	7416	4517	47 956	10 883	2481	6911	8226	2435
Smoking behaviour												
Mean cigarettes/day	-	8.2 (12.1)	8.2 (11.1)	0.2 (0.5)	4.1 (2.2)	16.3 (7.5)	-	3.7 (7.4)	4.2 (7.1)	0.1 (0.3)	3.7 (2.0)	14.4 (6.0)
Age started smoking (years)	-	17.4 (5.2)	18.0 (6.0)	18.8 (6.2)	18.1 (6.1)	16.8 (5.0)	-	21.3 (7.2)	22.8 (9.0)	23.6 (9.2)	21.3 (8.0)	19.5 (6.6)
Ever tried to quit	-	-	-	40%	43%	34%	-	-	-	42%	53%	43%
Inhale smoke	-	-	-	64%	77%	85%	-	-	-	52%	68%	77%
Others smoke inside home	16%	18%	22%	26%	35%	43%	32%	32%	39%	44%	56%	61%
Age, years	48 (10)	51 (10)	48 (10)	46 (9)	47 (9)	48 (9)	48 (10)	48 (9)	45 (8)	44 (8)	45 (8)	47 (8)
Resident in Iztapalapa	56%	55%	50%	59%	56%	52%	63%	58%	53%	60%	56%	48%
Education attained												
University	35%	27%	27%	27%	23%	25%	11%	18%	19%	16%	17%	23%
High school	26%	26%	29%	34%	32%	30%	24%	29%	34%	37%	35%	34%
Elementary school	33%	40%	37%	36%	39%	39%	52%	44%	41%	42%	42%	39%
Other	6%	7%	6%	4%	6%	5%	12%	9%	7%	5%	5%	5%
Drinking behaviour												
Drink at least weekly	49%	48%	48%	51%	48%	46%	52%	58%	59%	66%	61%	57%
Drink less than weekly	28%	39%	41%	35%	39%	41%	12%	21%	22%	16%	19%	22%
Former drinker	9%	9%	8%	11%	10%	9%	4%	5%	6%	8%	8%	8%
Never drinker	14%	5%	3%	2%	3%	4%	32%	16%	13%	10%	12%	13%
Anthropometry and blood pressure												
BMI kg/m²	28.2 (4.3)	28.5 (4.2)	28.6 (4.5)	28.6 (4.3)	27.8 (4.4)	27.0 (4.4)	29.8 (5.2)	29.9 (5.4)	29.6 (5.1)	29.8 (5.2)	29.0 (5.2)	28.0 (5.3)
Waist-to-hip ratio	0.95 (0.07)	0.95 (0.07)	0.95 (0.06)	0.95 (0.06)	0.94 (0.06)	0.95 (0.07)	0.87 (0.07)	0.87 (0.07)	0.87 (0.07)	0.86 (0.06)	0.86 (0.07)	0.86 (0.07)
Systolic BP (mmHg)	127 (15)	128 (15)	127 (14)	126 (14)	126 (14)	126 (15)	125 (16)	124 (16)	123 (15)	121 (14)	121 (15)	122 (15)
Diastolic BP (mmHg)	84 (10)	85 (10)	84 (10)	84 (9)	83 (10)	84 (10)	82 (10)	82 (10)	81 (10)	80 (10)	80 (10)	80 (10)

Analyses excluded data from participants with previously diagnosed chronic disease (chronic kidney disease, ischemic heart disease, stroke, cirrhosis, cancer, emphysema, or diabetes).

Table B3: Smoking characteristics at resurvey according to baseline smoking status, ages 35-69

	Never smoker	Ex smoker	Smoke < daily	Smoke < 10 cigs/day	Smoke 10+ cigs/day daily
Participants	4381	1651	870	1016	381
Smoking status at resurvey					
Never smoker	4001 (91%)	442 (27%)	165 (19%)	53 (5%)	7 (2%)
Ex smoker (not for illness)	284 (6%)	866 (52%)	389 (45%)	276 (27%)	50 (13%)
Ex smoker (for illness)	22 (1%)	202 (12%)	66 (8%)	138 (14%)	47 (12%)
Smoke < daily	34 (1%)	58 (4%)	156 (18%)	125 (12%)	14 (4%)
Smoke < 10 cigs/day	10 (<0.5%)	44 (3%)	66 (8%)	351 (35%)	107 (28%)
Smoke 10+ cigs/day	1 (<0.5%)	6 (<0.5%)	3 (<0.5%)	55 (5%)	154 (40%)
Incomplete info $	29 (1%)	33 (2%)	25 (3%)	18 (2%)	2 (1%)
Average daily cigs					
At baseline	-	5.4 (9.6)	0.1 (0.3)	3.8 (2.1)	15.5 (7.2)
At resurvey, per smoker †	1.8 (4.4)	2.7 (3.9)	1.3 (2.0)	4.0 (3.8)	10.5 (6.9)
At resurvey, all participants ‡	0.0 (0.5)	0.2 (1.2)	0.4 (1.2)	2.1 (3.4)	7.6 (7.5)
Inhale smoke					
At baseline	-	-	505 (58%)	741 (73%)	310 (81%)
At resurvey	27 (1%)	81 (5%)	158 (18%)	388 (38%)	203 (53%)
Ever tried to quit §					
At baseline	-	-	324 (37%)	469 (46%)	148 (39%)
At resurvey	30 (1%)	83 (5%)	168 (19%)	395 (39%)	207 (54%)

N (%) or mean (SD) shown.
$ Where information given is insufficient to categorise exposure, e.g. when participant reports ex-smoking but not years since quitting, or current smoking but not frequency of smoking
† Among those self-reporting current smoking at resurvey
‡ Among all participants, i.e. those reporting never or ex smoking assigned zero cigarettes/day
§ Only current smokers asked about quit behaviour at baseline.

Table B4: Smoking characteristics at resurvey according to baseline smoking status, among those with chronic disease at recruitment, ages 35-89

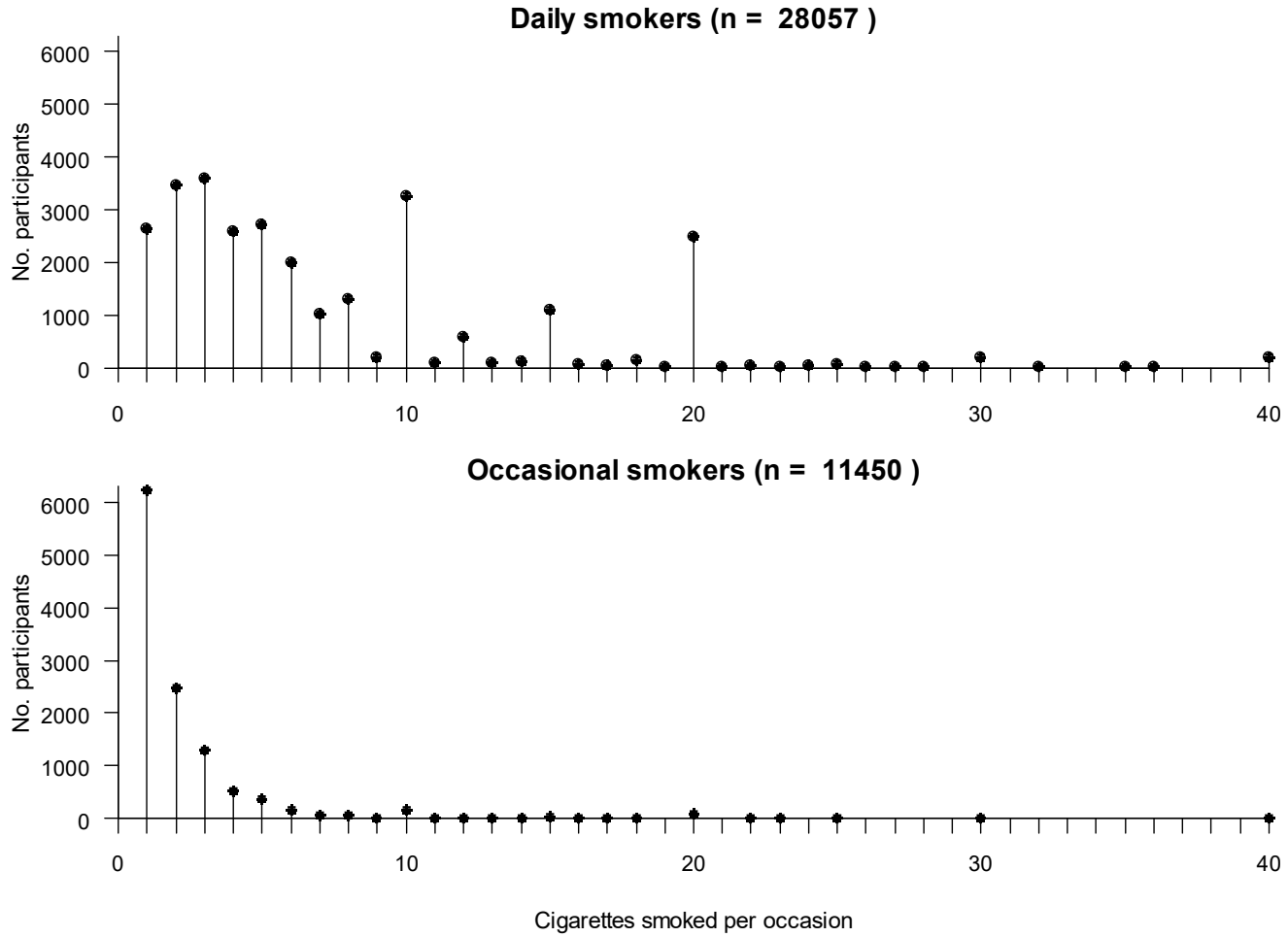
	Never smoker	Ex smoker	Smoke < daily	Smoke < 10 cigs/day	Smoke 10+ cigs/day daily
Participants	688	340	108	138	51
Smoking status at resurvey					
Never smoker	646 (94%)	81 (24%)	28 (26%)	5 (4%)	3 (6%)
Ex smoker (not for illness)	30 (4%)	177 (52%)	49 (45%)	40 (29%)	13 (25%)
Ex smoker (for illness)	3 (<0.5%)	62 (18%)	13 (12%)	37 (27%)	5 (10%)
Smoke < daily	0 (0%)	10 (3%)	7 (6%)	13 (9%)	1 (2%)
Smoke < 10 cigs/day	3 (<0.5%)	7 (2%)	6 (6%)	37 (27%)	11 (22%)
Smoke 10+ cigs/day	0 (0%)	0 (0%)	0 (0%)	2 (1%)	17 (33%)
Incomplete info $	6 (1%)	3 (1%)	5 (5%)	4 (3%)	1 (2%)
Average daily cigs					
At baseline	-	7.8 (11.6)	0.1 (0.3)	3.6 (2.0)	16.8 (8.2)
At resurvey, per smoker †	2.3 (0.6)	1.7 (2.0)	1.7 (1.7)	3.0 (2.6)	11.0 (7.5)
At resurvey, all participants ‡	<0.1 (0.2)	0.1 (0.6)	0.2 (0.8)	1.1 (2.2)	6.4 (7.9)
Inhale smoke					
At baseline	-	-	53 (49%)	88 (64%)	42 (82%)
At resurvey	3 (<0.5%)	10 (3%)	7 (6%)	37 (27%)	19 (37%)
Ever tried to quit §					
At baseline	-	-	32 (30%)	52 (38%)	24 (47%)
At resurvey	2 (<0.5%)	10 (3%)	11 (10%)	41 (30%)	22 (43%)

N (%) or mean (SD) shown.
$ Where information given is insufficient to categorise exposure, e.g. when participant reports ex-smoking but not years since quitting, or current smoking but not frequency of smoking
† Among those self-reporting current smoking at resurvey
‡ Among all participants, i.e. those reporting never or ex smoking assigned zero cigarettes/day
§ Only current smokers asked about quit behaviour at baseline.

Table B5: Population attributable fraction (PAF) at ages 35-69, by disease and sex

	Male			Female		
Cause of death	No. of deaths smokers/non-smokers	RR (95% CI)	PAF (%)	No. of deaths smokers/non-smokers	RR (95% CI)	PAF (%)
Respiratory cancers	62/7	2.29 (1.04-5.05)	51	30/37	1.31 (0.79-2.16)	11
COPD	57/9	1.47 (0.72-2.99)	27	35/56	1.48 (0.95-2.30)	12
Gastrointestinal	102/27	0.83 (0.54-1.27)	-17	82/119	1.35 (1.01-1.81)	11
IHD	598/116	1.28 (1.05-1.57)	19	283/473	1.19 (1.02-1.39)	6
Stroke	160/26	1.56 (1.02-2.38)	31	110/165	1.22 (0.95-1.56)	7
Other vascular	66/10	1.60 (0.82-3.13)	33	31/59	1.08 (0.68-1.69)	2
Other cancers	293/82	0.91 (0.71-1.16)	-8	323/611	0.85 (0.74-0.98)	-6
Renal	459/117	0.96 (0.78-1.18)	-3	299/517	1.14 (0.98-1.32)	4
Hepatobiliary	363/81	0.95 (0.74-1.21)	-5	139/178	1.36 (1.08-1.71)	12
Other respiratory	187/41	1.10 (0.78-1.56)	8	106/206	0.93 (0.73-1.19)	-2
Acute diabetic	128/24	1.20 (0.77-1.88)	14	66/106	1.18 (0.86-1.63)	6
Other medical	165/39	1.03 (0.72-1.47)	2	117/188	1.19 (0.93-1.51)	6
Any medical	2640/580	1.09 (1.00-1.20)	7	1621/2716	1.10 (1.03-1.18)	3

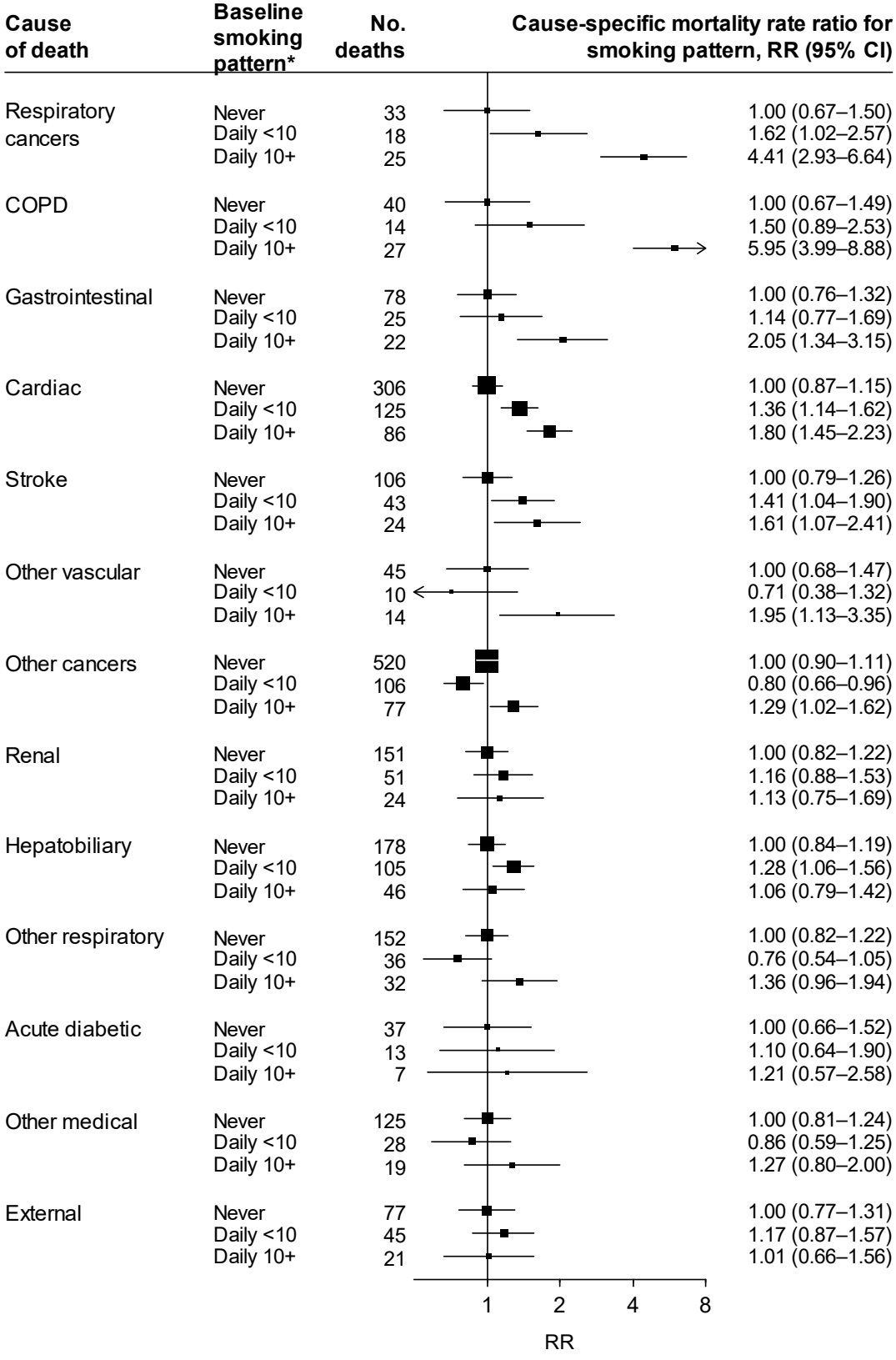
Figure B1: Cigarettes smoked per occasion by smoking frequency among current smokers



Note: Excludes those who reported smoking >40 cigarettes per occasion (65 daily smokers and 2 occasional smokers)

Figure B2: Risk of cause-specific mortality among daily vs never smokers by amount smoked, ages 35-69

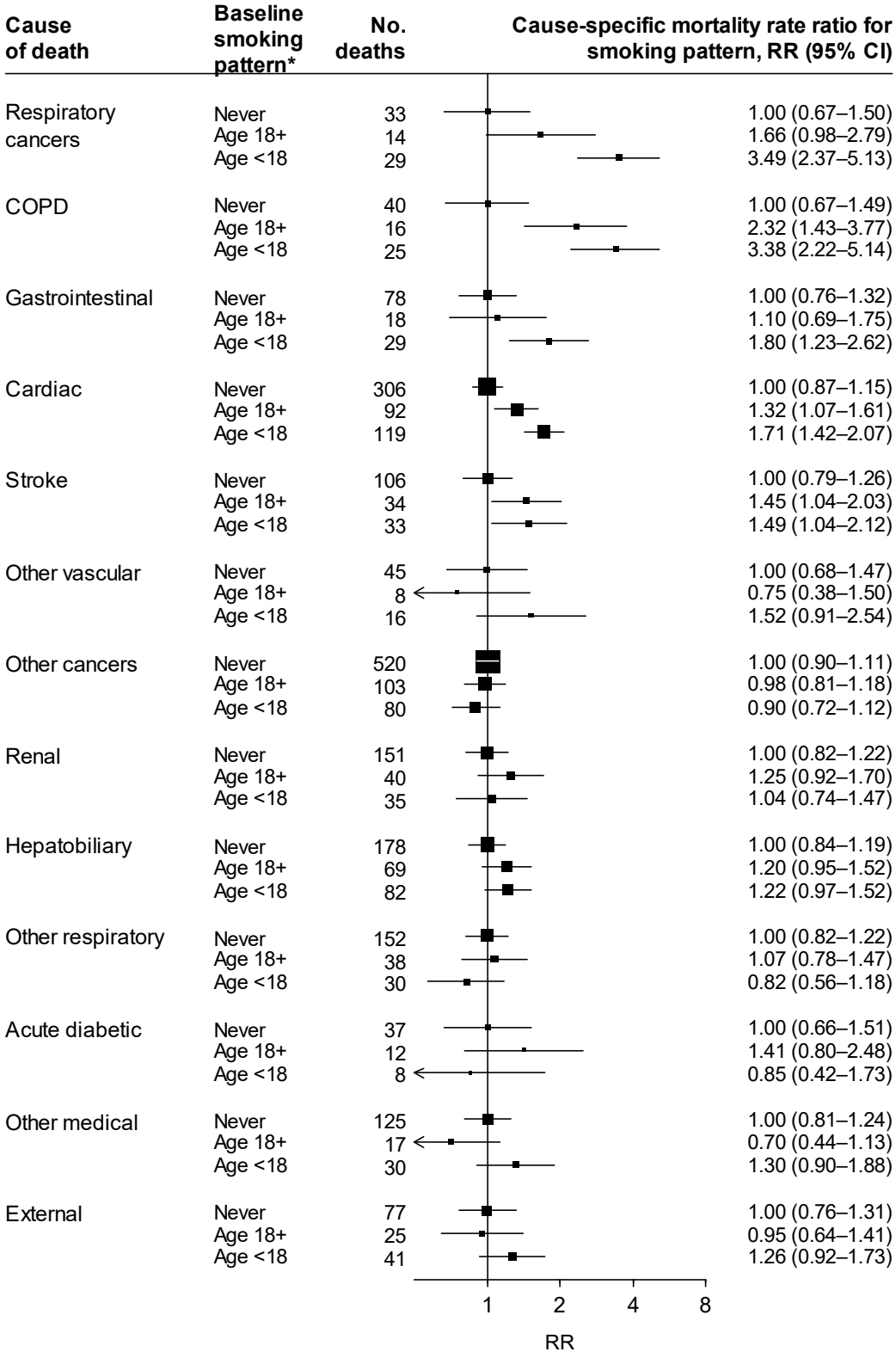
Participants with chronic disease at recruitment were excluded from these analyses.



*RRs associated with smoking 0, 4, and 16 cigarettes/day on average at baseline, respectively.

Figure B3: Risk of cause-specific mortality among daily vs never smokers by age started, ages 35-69

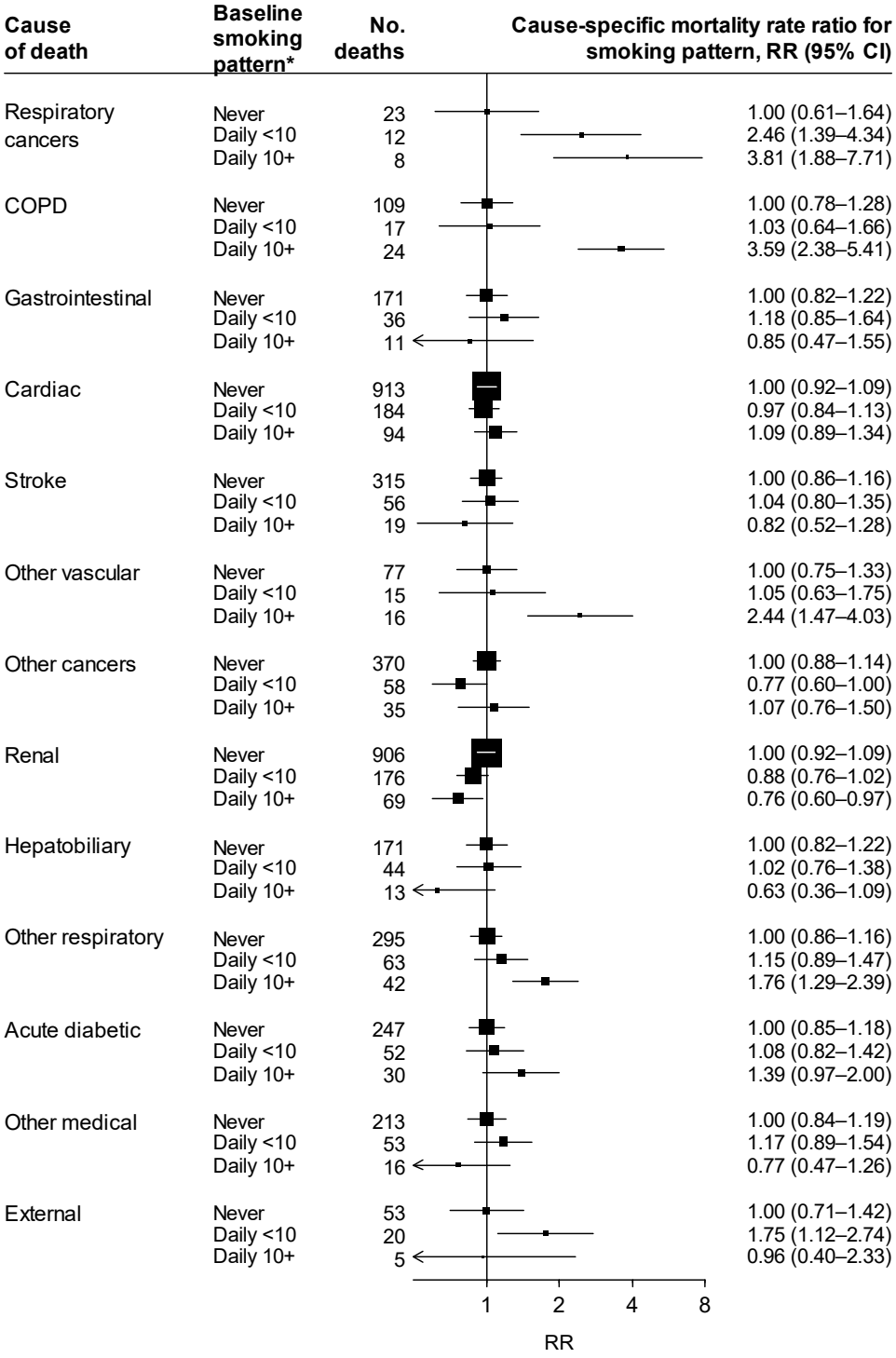
Participants with chronic disease at recruitment were excluded from these analyses.



*RRs associated with starting to smoke at ages 24 or 14 years on average, respectively, compared to never smoking.

Figure B4: Risk of cause-specific mortality among daily vs never smokers by amount smoked, ages 35-89

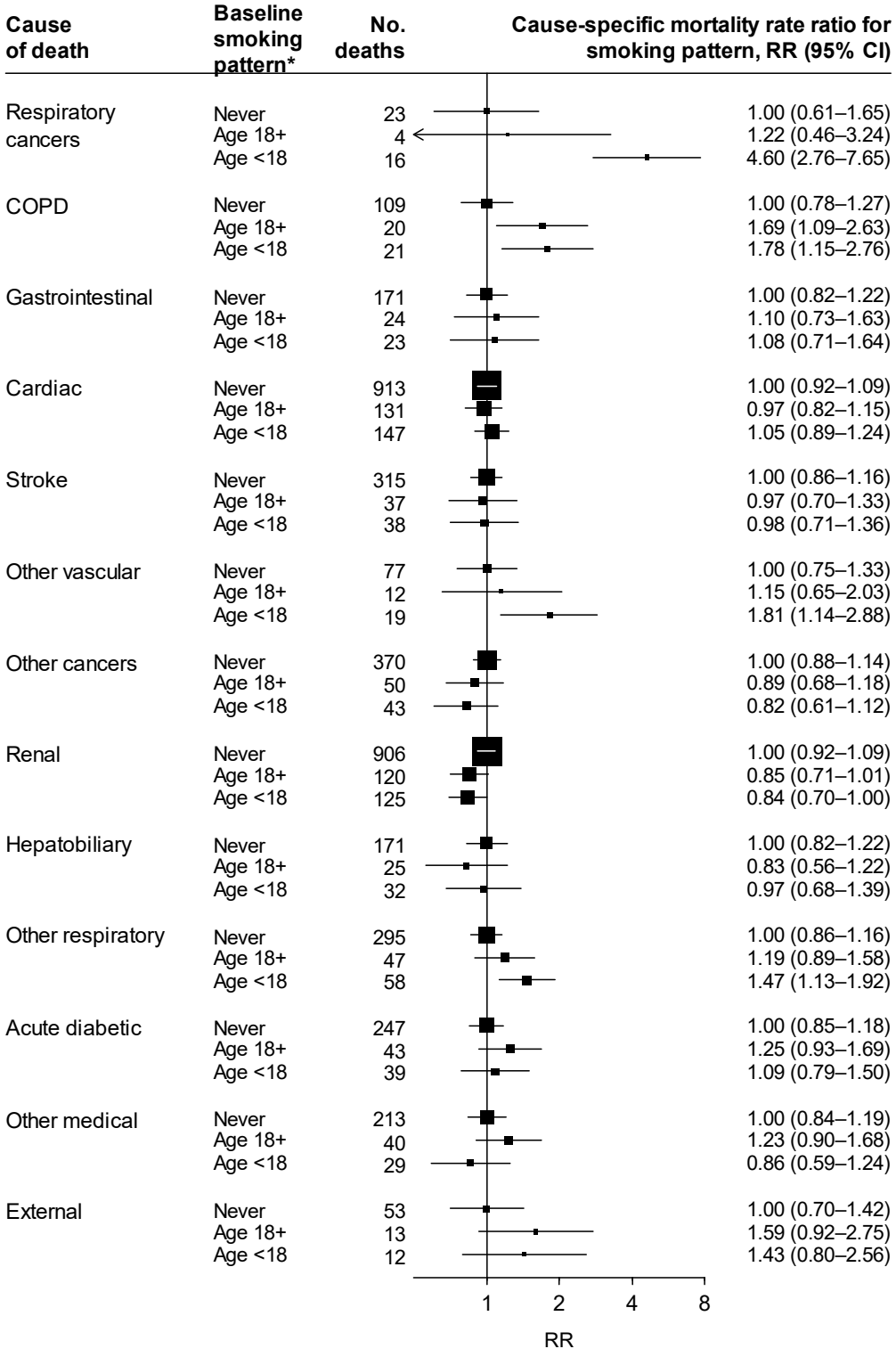
Only participants with chronic disease at recruitment were included in these analyses.



*RRs associated with smoking 0, 4, and 16 cigarettes/day on average at baseline, respectively.

Figure B5: Risk of cause-specific mortality among daily vs never smokers by age started, ages 35-89

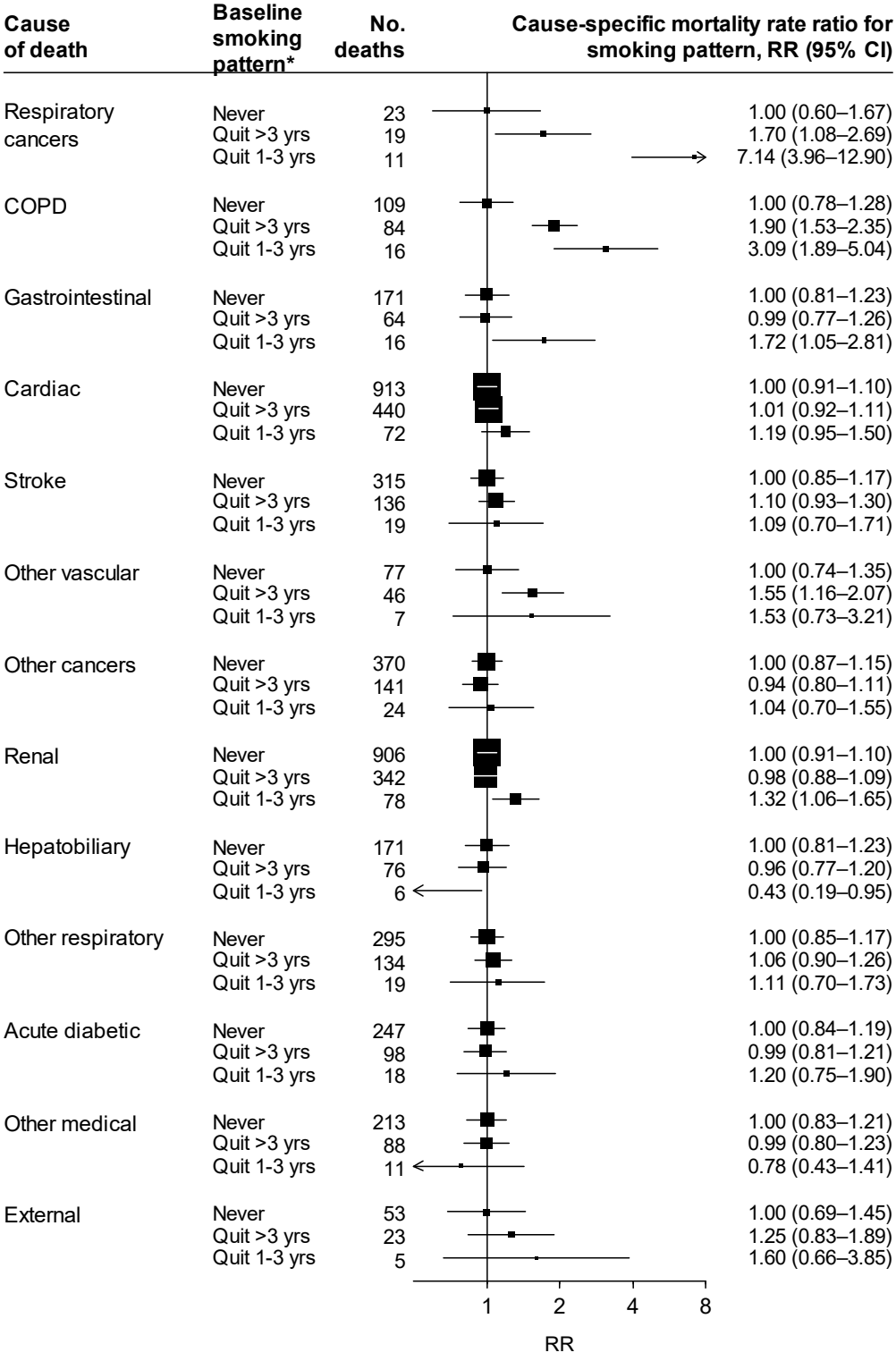
Only participants with chronic disease at recruitment were included in these analyses.



*RRs associated with starting to smoke at ages 24 or 14 years on average, respectively, compared to never smoking.

Figure B6: Risk of cause-specific mortality, never vs ex-daily smokers by quitting duration, ages 35-89

Only participants with chronic disease at recruitment were included in these analyses.



*Precise number of years since quitting was not reported.

Appendix C: Additional results for the US National Health Interview Survey prospective study

Table C1: National estimates of never, ex, non-daily, and daily smokers by race in the United States in 2018 (in millions)

	Mexican American	Other Hisp.	Non-Hisp. White	Non-Hisp. Black	Other race	Total
Never	19.0	11.8	91.5	20.8	16.5	160.0
Ex	3.0	2.8	42.0	3.9	3.3	55.0
Non-daily	1.0	0.5	4.8	1.5	0.8	8.7
Daily	1.4	1.1	18.7	2.7	1.7	25.5
Total	24.3	16.3	156.9	29.0	22.3	248.7

Estimates limited to adults (ie, those aged 18 or older at the time of survey).

Estimates may not sum due to rounding.

Excludes the estimated 0.7 million participants with insufficient information to be categorised.

Table C2: Unweighted sample size of never, ex, non-daily, and daily smokers by race in the United States in 2018 (NHIS)

	Mexican American	Other Hisp.	Non-Hisp. White	Non-Hisp. Black	Other race	Total
Never	1,391	952	9,673	1,880	1,396	15,292
Ex	257	236	4,939	480	371	6,283
Non-daily	81	47	562	154	80	924
Daily	105	94	2,136	312	187	2,834
Total	1,834	1,329	17,310	2,826	2,034	25,333

Estimates limited to adults (ie, those aged 18 or older at the time of survey).

Excludes the 84 participants with insufficient information to be categorised.

Figure C1: Mortality rate ratio at ages 30-69 for major smoking-related causes, daily vs never smokers at recruitment, by age started smoking

Analyses conducted among all participants (pooled across race), further adjusting for race. Analyses include those with illness at recruitment.

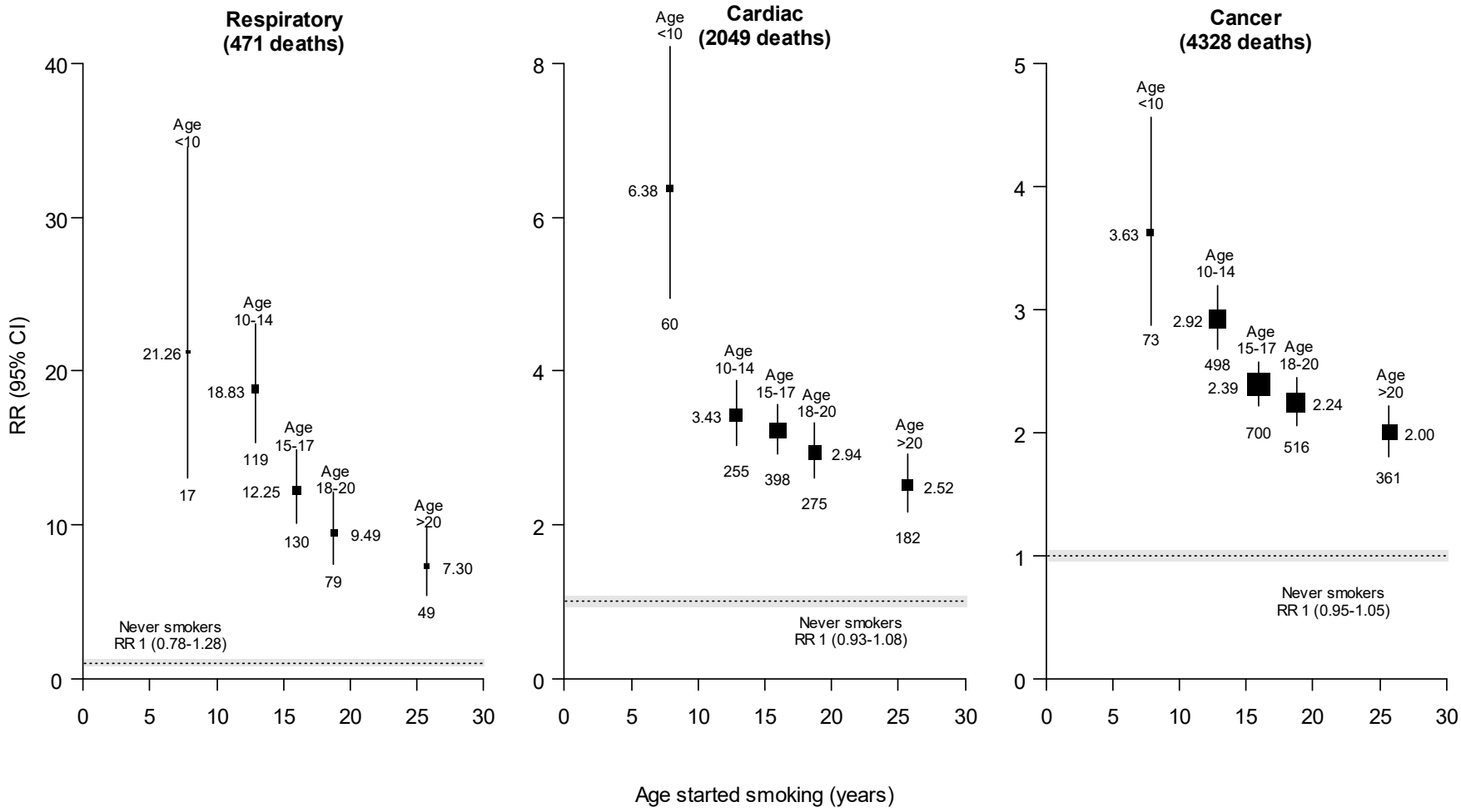


Figure C2: All-cause mortality rate ratio at ages 70-89, daily vs never smokers, by age started smoking

Analyses conducted among all participants (pooled across race), further adjusting for race. Analyses include those with illness at recruitment.

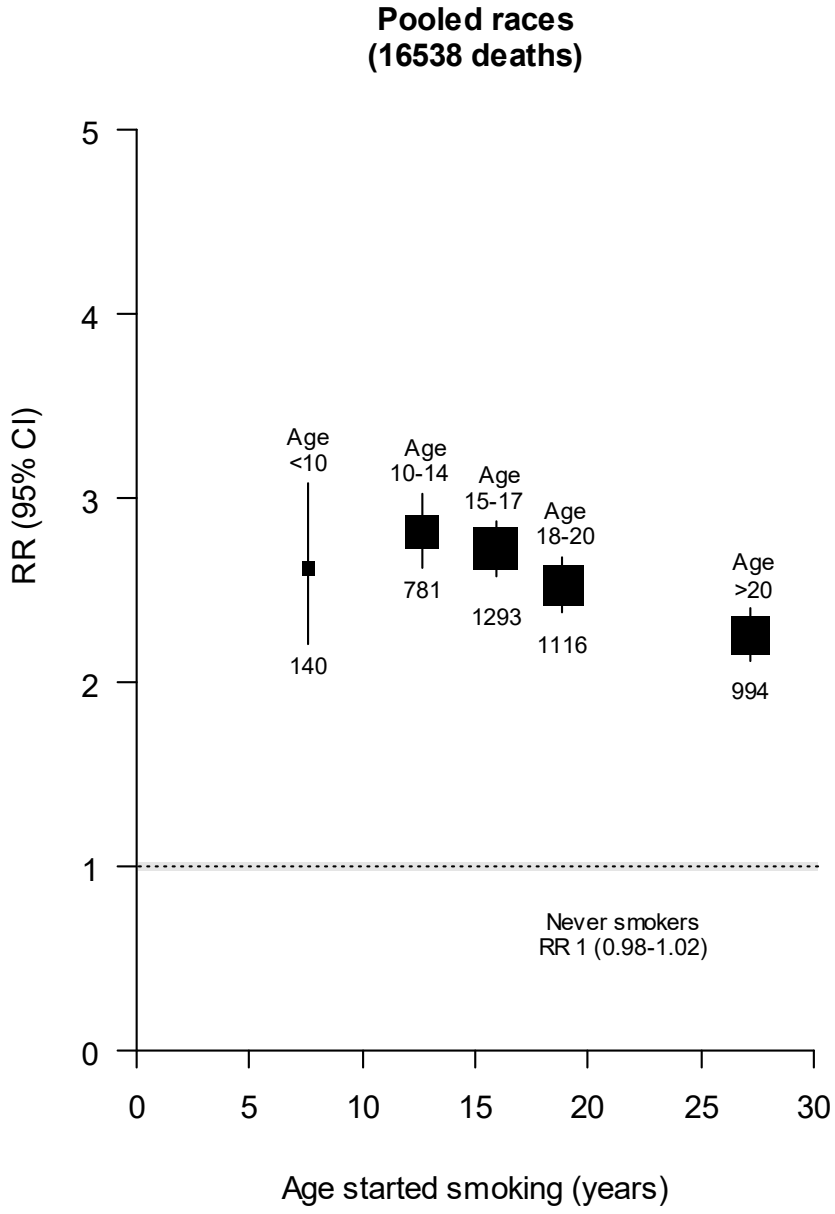


Figure C3: All-cause mortality rate ratio at ages 70-89, daily vs never smokers, by amount smoked

Analyses conducted among all participants (pooled across race), further adjusting for race. Analyses include those with illness at recruitment.

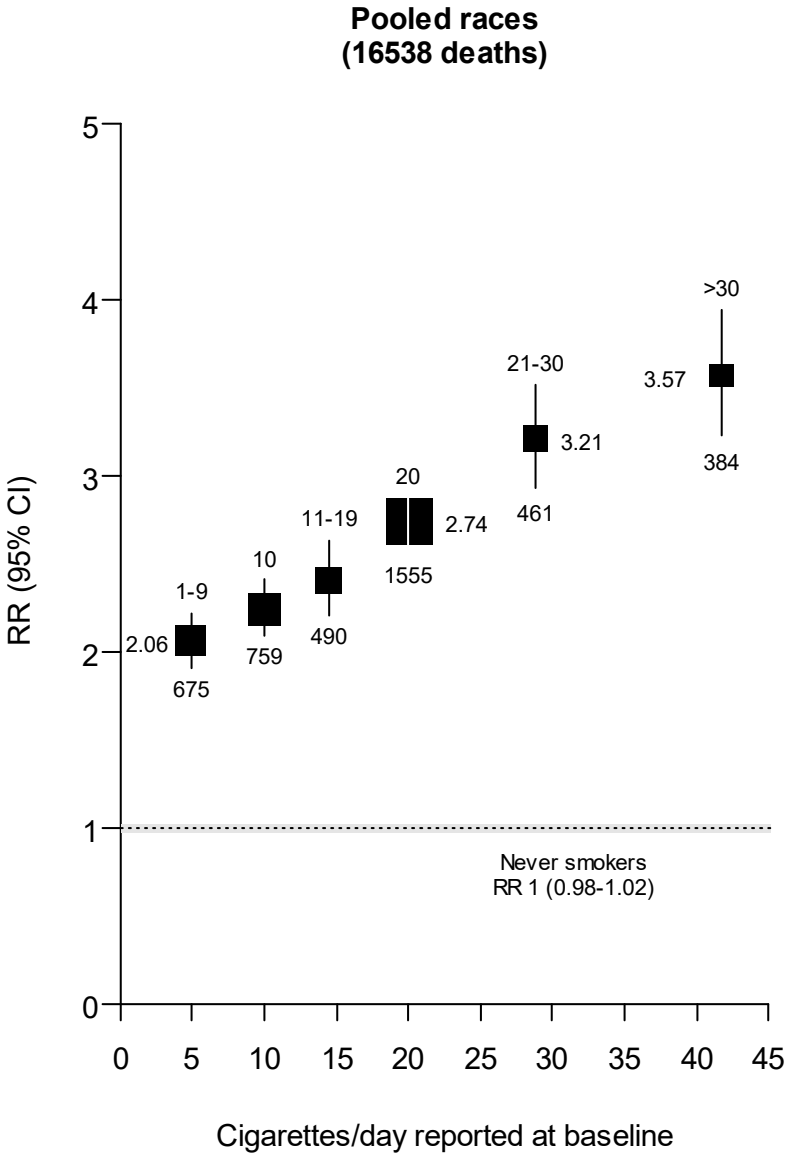


Figure C4: Mortality rate ratio at ages 30-69 for major smoking-related causes, daily vs never smokers at recruitment, by amount smoked

Analyses conducted among all participants (pooled across race), further adjusting for race. Analyses include those with illness at recruitment.

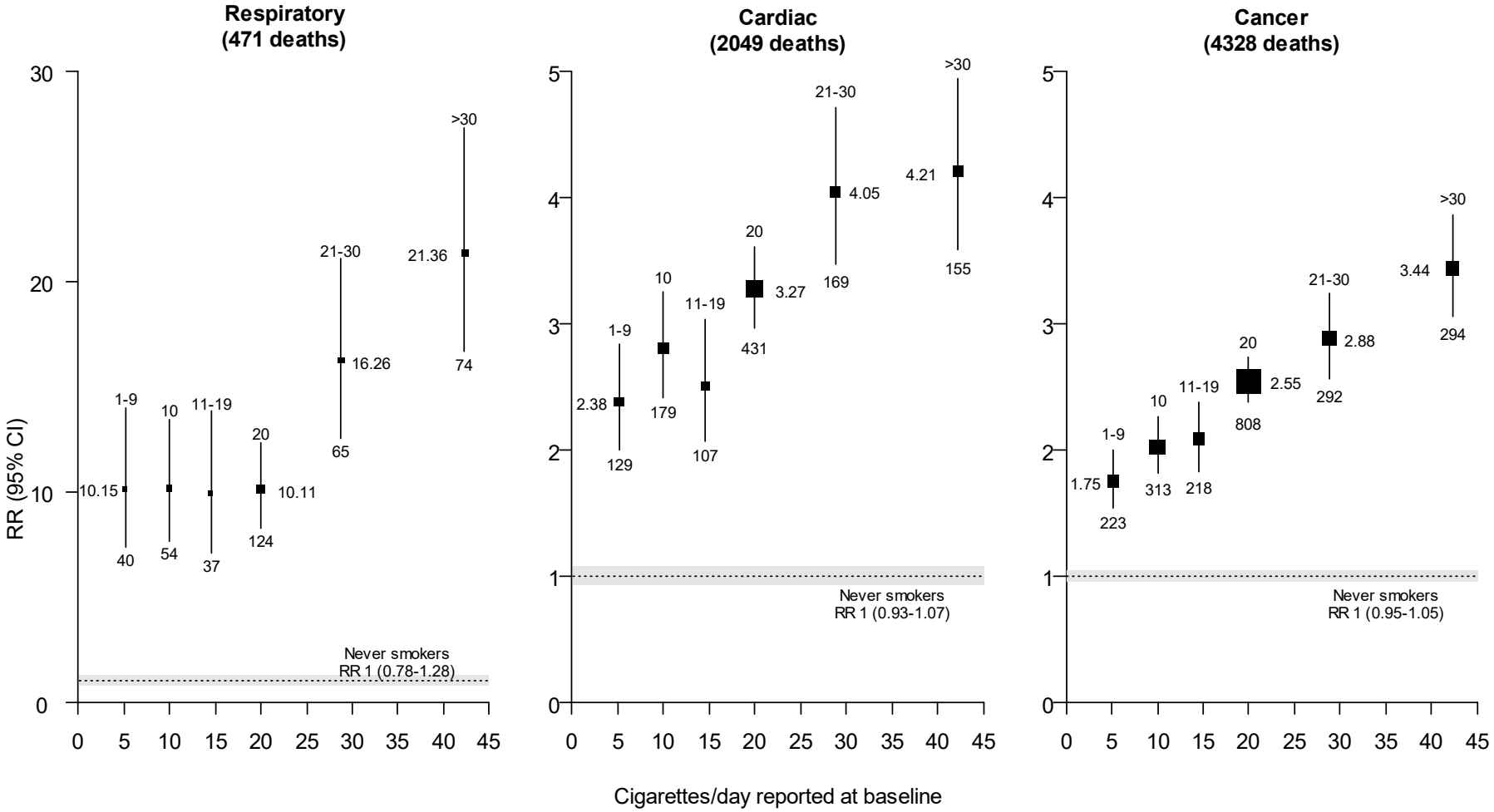


Figure C5: Mortality rate ratio at ages 30-89 for major smoking-related causes, by age quit smoking among those quitting at ages 15-54, compared to never and current smokers at recruitment

Participants with chronic disease at recruitment were excluded from these analyses.

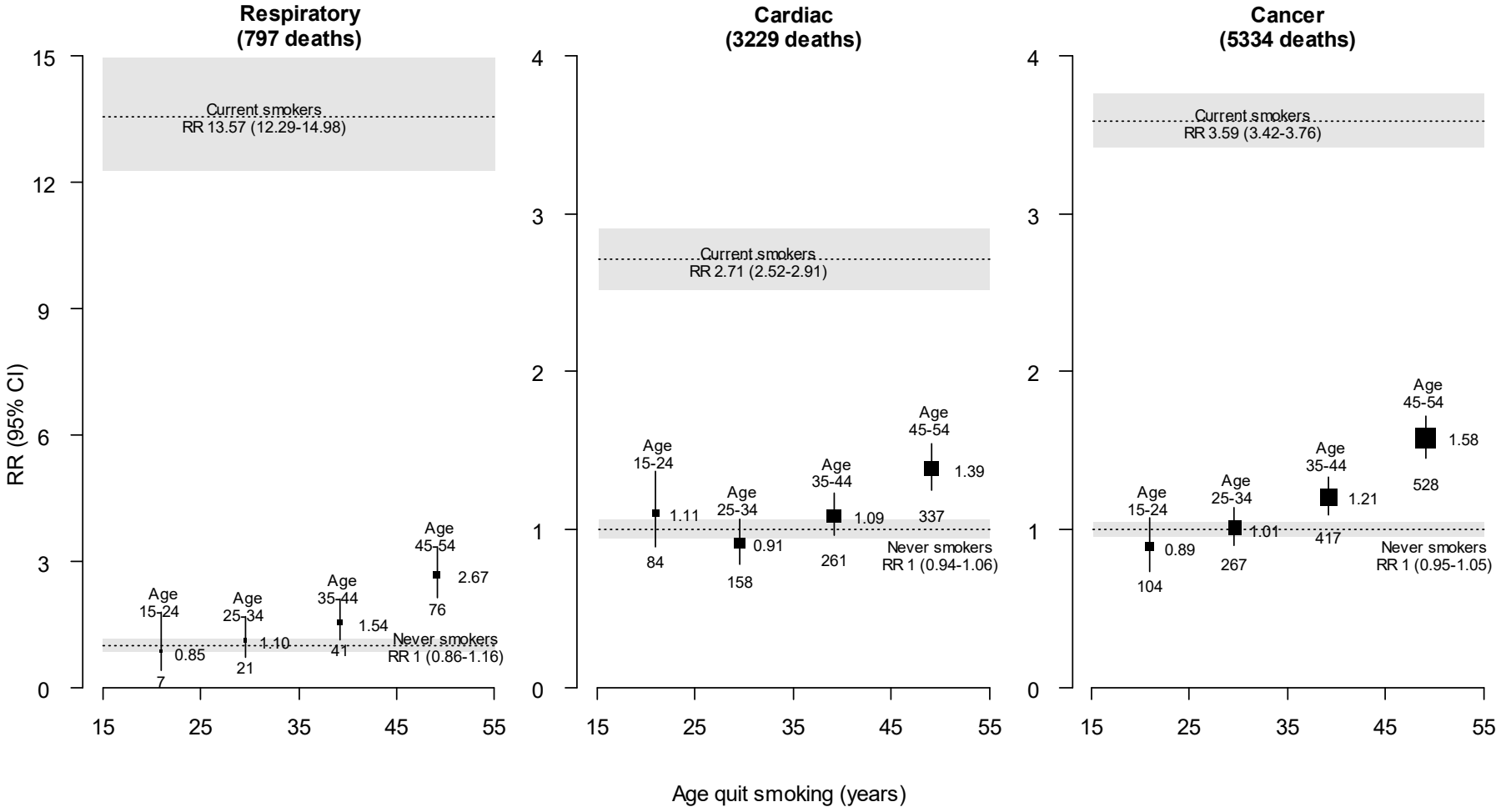


Figure C6: Mortality rate ratio at ages 30-89 for major smoking-related causes, by years since quitting smoking, compared to never and current smokers at recruitment

Participants with chronic disease at recruitment were excluded from these analyses.

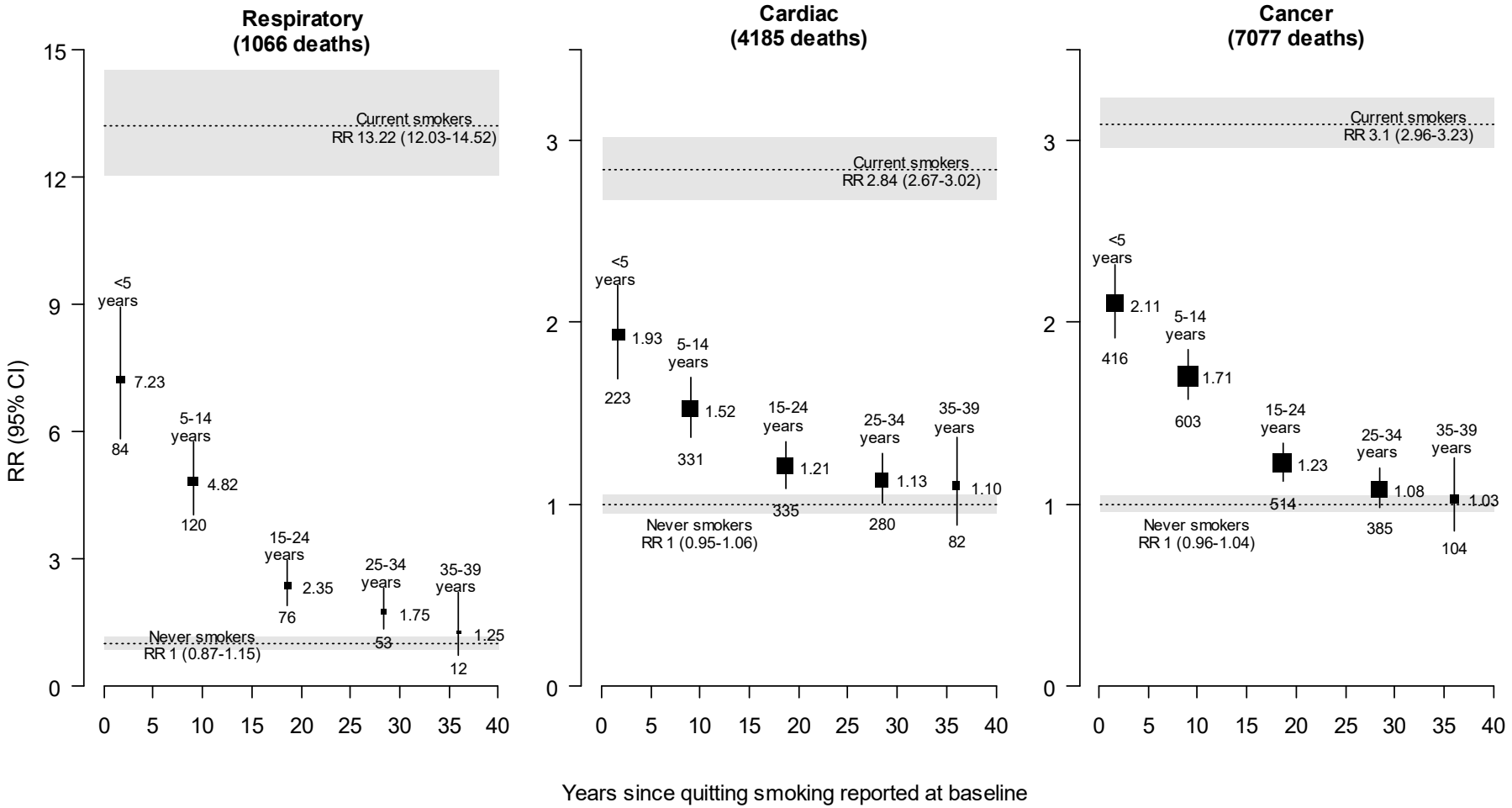


Figure C7: All-cause mortality rate ratio by age quit smoking among those quitting at ages 15-54, compared to never and current smokers at baseline, age at risk 55-89

Analyses conducted among all participants (pooled across race), further adjusting for race. Analyses include those with illness at recruitment.

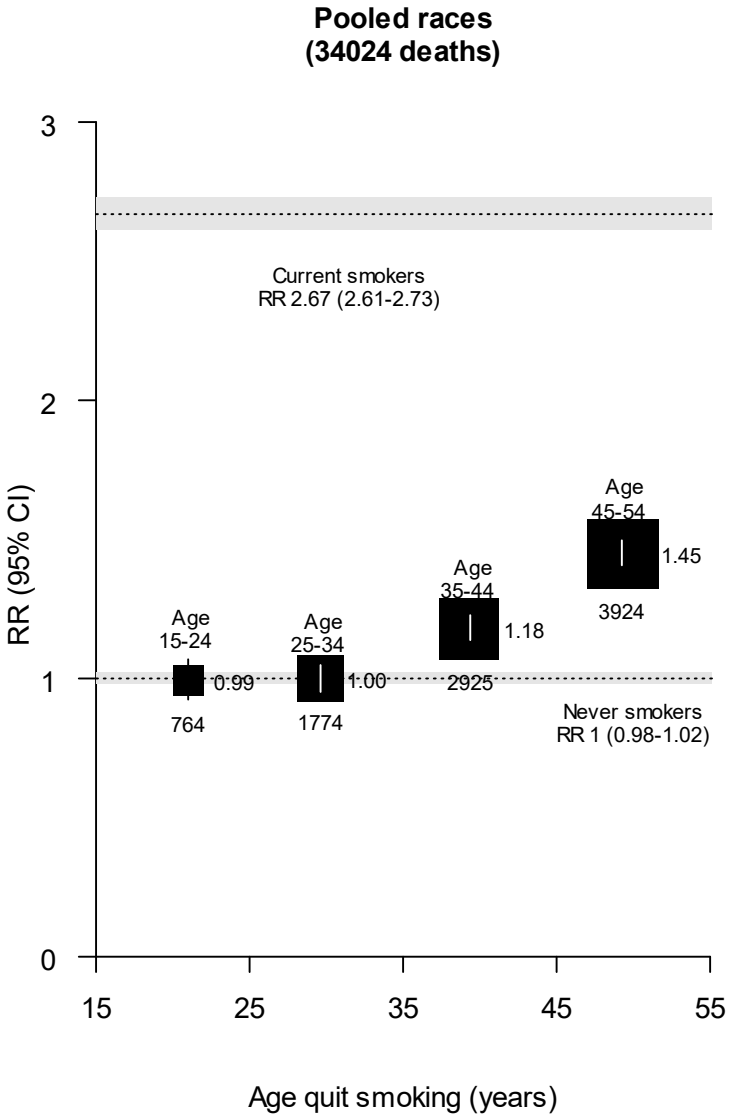
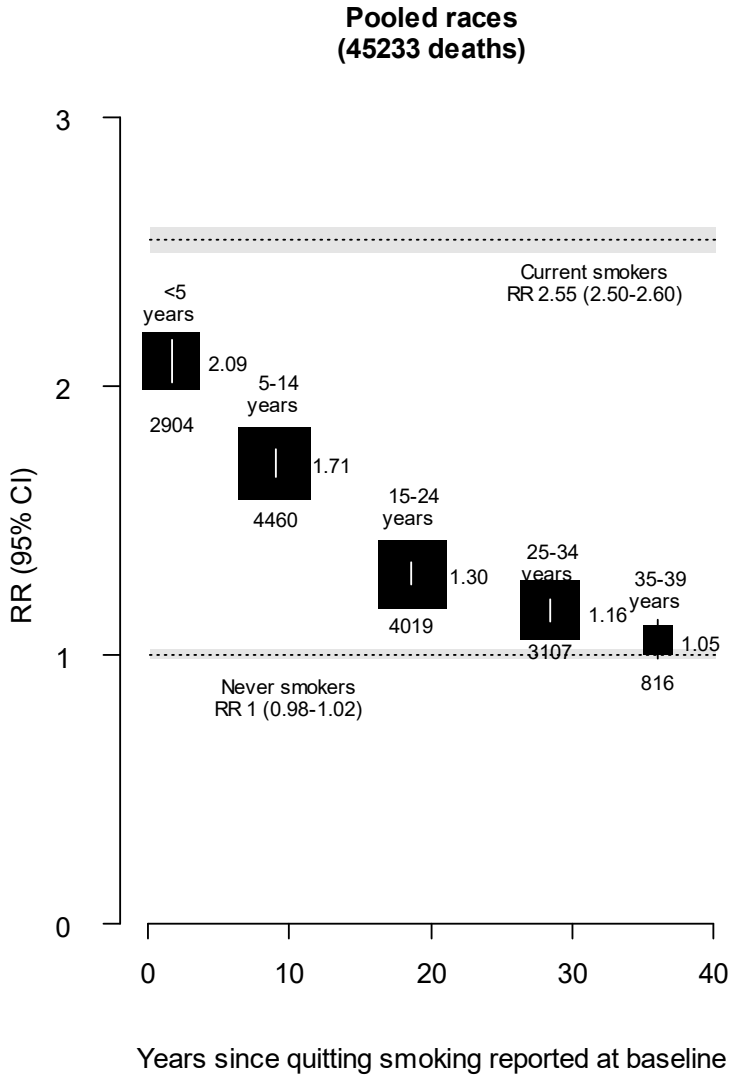


Figure C8: All-cause mortality rate ratio by years since quitting smoking, compared to never and current smokers at baseline, age at risk 30-89

Analyses conducted among all participants (pooled across race), further adjusting for race. Analyses include those with illness at recruitment.



Appendix D: Additional information and methods for the Global Adult Tobacco Surveys

Background of the Global Adult Tobacco Surveys (GATS)

The purpose, design, and conduct of the GATS surveys have been previously described,[168] and are briefly summarised here. In 1998, the World Health Organisation (WHO) and US Centers for Disease Control and Prevention (CDC) launched a partnership to monitor tobacco use, with an initial focus on school-based surveys. In 2008, the GATS survey was launched as a means of conducting surveillance of tobacco use and policies to address it in countries around the world, with a focus on adult use. The survey is designed to monitor each of the six major tobacco control measures introduced by the WHO's MPOWER package: Monitor tobacco use and prevention policies; Protect people from tobacco smoke; Offer help to quit tobacco use; Warn about the dangers of tobacco; Enforce bans on tobacco advertising, promotion, and sponsorship; and Raise taxes on tobacco.

The GATS survey is a nationally-representative household survey of adults aged 15 years or older. The survey is questionnaire-based, and limited to self-reported information. Designed to be highly standardised across countries, but also to customisable to meet country-specific needs and interests, the survey is comprised of a core questionnaire and additional questions that may be added to address local priorities. The GATS survey methodology was also standardised to maximise comparability across countries, including the sample design, analysis plan, field interviewer protocol, and data release policies. At the time of writing, the survey was the most comprehensive, standardised, cross-country survey of tobacco use among adults in predominantly low- and middle-income countries. A report detailing findings from GATS surveys conducted in 22 countries (which contain more than half of the global adult population) can be found in the GATS Atlas online (http://gatsatlas.org/).

Methods

Files containing microdata (ie, individual participant data) and corresponding codebooks for each country were downloaded from the CDC's Global Tobacco Surveillance System Data website (https://www.cdc.gov/tobacco/global/gtss/gtssdata/index.html). Where the survey had been conducted more than once in a country, the most recent publicly-available survey round was used for analyses (as of 1 September 2019).

Although the GATS surveys were designed to track progress on each of the major WHO evidence-based tobacco control measures introduced to advance the Framework Convention on Tobacco Control (MPOWER), analyses reported in **Chapter 6** focus on monitoring use (as the topic most directly relevant to the findings of this thesis). Wherever possible, analyses were conducted such that they aligned with other analyses conducted throughout this thesis.

Statistical analysis

Respondents' smoking frequency was determined by answers given to the following question: "Do you currently smoke tobacco on a daily basis, less than daily, or not at all?" Those who responded that they smoked daily or less than daily were categorised accordingly. Those who said they currently did not smoke were then asked, "In the past, have you smoked tobacco on a daily basis, less than daily, or not at all?" Those who answered prior daily or non-daily smoking were considered ex-smokers. Analyses of age started smoking, amount smoked, and time to first cigarette after waking were restricted to daily smokers.

Variables reflecting the amount smoked and age started smoking were derived such that they most closely aligned with the categorisations used in the primary findings of interest (eg, 'light' smoking in Mexico, childhood smoking in Cuba). Therefore, the number of

cigarettes smoked per day was grouped into three categories: 1-9 (ie, <10) cigs/day, 10-19 cigs/day, and 20-60 cigs/day. As the predominant focus of this thesis was cigarette smoking, analyses of amount smoked was restricted to cigarettes, either manufactured or hand-rolled (though smoking habits regarding several other tobacco products were also recorded separately, some of which were country-specific). Age at smoking initiation was categorised into four age groups: 5-9, 10-14, 15-19, and 20-49 years. However, some analyses report smoking initiation at ages 5-14 to reflect overall childhood smoking. As tobacco product smoked may change over time, analyses of age at which respondents began smoking tobacco daily did not take product currently (or predominantly) smoked into consideration.

Analyses of the proportion of ever smokers who had quit (or 'quit ratio') included all current and ex-smokers, irrespective of smoking frequency (ie, daily or non-daily) or product smoked. These analyses did not take into account age at smoking cessation or duration of cessation. Finally, analyses of time to first smoking after waking were recorded categorically during data collection as ≤5 minutes, 6-30 minutes, 31-60 minutes, or >60 minutes after waking. However, analyses joined these first two groups (≤30 minutes) to improve precision of estimates.

Overall analyses were conducted among all respondents in order to produce nationally-representative results of 'adult' smokers (here, aged≥15 years). Because smoking prevalence and patterns are strongly determined by sex in most populations, however, the main results are presented by sex. Overall and sex-specific estimates used sampling weights to account for the complex sample design of the surveys (taking into account probability of selection and participation), resulting in nationally-representative population estimates. All analyses were conducted using Stata v.13.1.

Appendix E: Questionnaires for the Cuba Prospective Study and the Mexico City Prospective Study

Cuba Prospective Study Questionnaire (for Baseline and Resurvey)

CUBAN STUDY OF TOBACCO AND HEALTH: PROSPECTIVE

Province Nº ☐ 1-9 Municipality Nº ☐☐ 0-99 Health Area Nº ☐ 0-9 Identity Card Nº ☐☐☐☐☐☐-☐☐☐☐☐ Date of birth (y,m,d)

Name of interviewee:

Address:

Street Nº Municipality Province Telephone

Name of friend:

Address:

Street Nº Municipality Province Telephone

CAPITAL LETTERS, PLEASE!

Sex (1=masc, 2=fem) ☐ 1-2

Skin colour(1=white, 2=black, 3=mixed, 4=other) ☐ 1-4

Marital status (1=single/widowed, 2=married) ☐ 1-2

Education completed (0=none, 1=primary, 2=secondary, 3=technician, 4=skilled worker, 5=high school, 6=university) ☐ 0-6

Occupational category (0=none, 1=manager, 2=professional, 3=technician, 4=service worker, 5=agricultural, 6=industrial, 7=military; if housewife, retired or unemployed, we would like to know if you were employed in the past) ☐ 0-7

(☐ 0-7)

What I'd like to do is measure your blood pressure while you are sitting down, ask you a few questions about tobacco and alcohol, and then do your blood pressure again. Blood pressure (single measurement, seated): ☐☐☐ / ☐☐☐ SBP / DBP

Now, TOBACCO: Have you EVER smoked regularly (i.e. on most days for at least a year)? ☐ **Y=Yes, N=No**
If NO, then go directly to the ALCOHOL questions

AT ABOUT WHAT AGE DID YOU FIRST START SMOKING TOBACCO REGULARLY? ☐☐ years of age

Since then, have you ever STOPPED smoking for a year or more? ☐ Y=Yes, N=No

About how many cigarettes a day did you smoke last month? (0=none) ☐☐ daily

About how many cigars a day did you smoke last month? (0=none) ☐☐ daily

If NO SMOKING last month, at about what age did you last stop smoking? (blank = not applicable) ☐☐ years of age

Now I'd like to ask you some questions about ALCOHOL:

About how many days a week would you drink any rum or other spirits? (0=none or usually less than once a week) ☐ 0-7

On a day when you drink rum (or other spirits), about how many glasses ("lineas") would you drink? ☐☐ lineas

About how many days a week would you drink beer? (0=none or usually less than once a week) ☐ 0-7

On a day when you do drink beer, about how many bottles would you drink? ☐☐ bottles

Finally, MEDICAL MATTERS: I'd like to ask about various different illnesses. What I want to know for each illness is whether you've ever had medical attention for it. If you've never heard of the illness, then just say NO.

Asthma	☐ Y/N	Stroke (definite/probable)	☐ Y/N	Cirrhosis of the liver	☐ Y/N
Emphysema/C.O.P.D.	☐ Y/N	Myocardial infarction "	☐ Y/N	Chronic kidney disease	☐ Y/N
Diabetes	☐ Y/N	Angina pectoris "	☐ Y/N	Skin cancer	☐ Y/N
Peptic ulcer	☐ Y/N	Hypertension	☐ Y/N	Other cancer*	☐ Y/N

*If Other cancer, where in your body did the cancer start? ☐ 0-9
(1=mouth, 2=larynx, 3=lung, 4=oesphagus, 5=stomach, 6=pancreas, 7=cervix, 8=bladder, 9=breast, 0=other, more than 1 site, or not known)

Have you been taking any regular medication for more than 1 month? If Yes: What such medicines are you now taking? (Mark with an **X**)

Digoxin	☐	Nitrates	☐	Oral hypoglycaemic	☐	O.Contraceptive	☐	Ca. antagonist	☐
Beta blocker	☐	ACE inhibitor	☐	Insulin	☐	Antidepressant	☐	Bronchodilator	☐
Diuretic	☐	Aspirin	☐	Steroids	☐	Tranquillisers	☐	Antacid/H2	☐

THANK YOU VERY MUCH FOR YOUR ANSWERS. Now we just need one more blood pressure. Repeat blood pressure (single measurement, seated) ☐☐☐ / ☐☐☐ PAS / PAD

What was the last blood pressure in the patient's medical record **before** today?

Date of last BP (y,m,d) ☐☐☐☐☐☐ & ☐☐☐ / ☐☐☐ PAS / PAD

What is your approximate **HEIGHT** (either ask or measure)? ☐☐☐ cm

What is your approximate **WEIGHT** (either ask or measure)? ☐☐☐ kg

Please check that **ALL** the questions have been answered, so that there is no need to write back to you for extra details. **MANY THANKS!**

Date of interview ☐☐☐☐☐☐ (y,m,d)

Quality of co-operation (1=poor, 2=fair, 3=good) ☐ 1-3

Name of interviewer: please PRINT

Any comments: .....

ESTUDIO CUBANO SOBRE TABACO Y SALUD

Provincia No. ☐ 1-9 Municipio No. ☐☐ 0-99 Area de Salud No. ☐ 0-9 **№ 17524** **A**

Carnet de Identidad No. ☐☐☐☐☐☐–☐☐☐☐☐ Fecha de nacimiento (a.m.d) + 5 números
(11 números: excluir si no tiene C.I. de 6 + 5 números)

Nombre de la persona y dirección: _____
Calle No. Teléfono

Nombre del familiar y dirección: _____
Calle No. Municipio Provincia Teléfono

LETRA DE MOLDE, POR FAVOR

Sexo (1 = masc, 2 = fem.) ☐ 1-2
Color de piel (1 = blanco, 2 = negro, 3 = mestizo, 4 = otra) ☐ 1-4
Estado social (1 = soltero/viudo, 2 = casado) ☐ 1-2
Educación completada (0 = ninguno*, 1 = primaria, 2 = secundaria, 3 = técnico medio, 4 = obrero calificado, 5 = pre-universitario, 6 = universitario) ☐ 0-6
Categoría ocupacional (0 = ninguno*, 1 = dirigente, 2 = profesional, 3 = técnico, 4 = de servicio, 5 = agrícola, 6 = industria, 7 = militar) ☐ 0-7
* Si es cero (Ej. ama de casa, jubilado o desempleado), diga si ha trabajado alguna vez. ☐ 0-7

Quisiera medir su presión arterial mientras esta sentado, hacerle unas pocas preguntas sobre tabaco y alcohol, y medir su PA de nuevo.
Presión arterial (una sola toma, sentado): ☐☐☐ / ☐☐☐ PAS / PAD

Ahora, TABACO: ¿Ha fumado alguna vez regularmente (Ej. en la mayoría de los días)? ☐ S=Sí, N=No
Si NO, vaya directamente a las preguntas sobre ALCOHOL.

¿A QUE EDAD APROXIMADAMENTE COMENZO USTED A FUMAR REGULARMENTE? ☐☐ años de edad
¿Desde entonces ha dejado de fumar alguna vez por un año o más? ☐ S=Sí, N=No
¿Aproximadamente cuántos cigarrillos por día se fumó usted el último mes? (0 = ninguno) ☐☐ diariamente
¿Aproximadamente cuántos tabacos por día se fumó usted el último mes? (0 = ninguno) ☐☐ diariamente
¿Si no fumó el último mes, a que edad dejó usted de fumar? (en blanco si no es aplicable) ☐☐ años de edad

Ahora me gustaría hacerle algunas preguntas sobre ALCOHOL:

¿Aproximadamente cuántos días a la semana toma usted ron u otra bebida fuerte? (0 = cero días o usualmente menos de una vez por semana) ☐ 0-7
¿El día que usted toma ron (u otra bebida fuerte) cuántas líneas se toma? ☐☐ líneas
¿Aproximadamente cuántos días a la semana toma usted cerveza? (0 = cero días o usualmente menos de una vez por semana) ☐ 0-7
¿El día que usted toma cerveza, cuántas botellas se toma (aprox.) ☐☐ botellas

Finalmente, algunos aspectos médicos: Quisiera preguntarle sobre algunas enfermedades. Lo que quiero conocer es si usted alguna vez ha sido tratado por alguna de estas enfermedades. Si no lo ha sido, conteste No (N).

Asma	☐ S/N	A.V.E. (definitivo/probable)	☐ S/N	Cirrosis	☐ S/N
Enfisema (E.P.O.C.)	☐ S/N	Infarto miocardio "	☐ S/N	Enf. renal crónica	☐ S/N
Diabetes	☐ S/N	Angina pectoris "	☐ S/N	Cáncer de piel	☐ S/N
Ulcera péptica	☐ S/N	Hipertensión arterial	☐ S/N	Otro cáncer*	☐ S/N

¿*Si Otro Cáncer = Sí, dónde le comenzó el cáncer? ☐ 0-9
(1 = boca o faringe, 2 = laringe, 3 = pulmón, 4 = esófago, 5 = estómago, 6 = páncreas, 7 = cuello útero, 8 = vejiga, 9 = mama, 0 = otro, más de un sitio, o no conocido)

¿Esta usted tomando alguna medicina regularmente desde hace más de un mes? Si es Sí: ¿Qué medicinas esta tomando? (Marque con X)

Digoxina	☐	Nitratos	☐	Hipoglic. oral	☐	Anticoncept.	☐	Ca. Antag.	☐
Beta B.	☐	Inh. ACE	☐	Insulina	☐	Antidepres.	☐	Broncodilat.	☐
Diurético	☐	ASA	☐	Esteroides	☐	Ansiolítico	☐	Antiácido/H2	☐

MUCHAS GRACIAS POR SUS RESPUESTAS: Ahora necesito tomarle de nuevo su presión arterial. Repita la toma de PA (una sola toma sentado): ☐☐☐ / ☐☐☐ PAS/PAD

¿Cuál es la última cifra PA que tiene en su historia clínica antes de hoy?
Fecha de la última PA (a,m,d) ☐☐☐☐☐☐ & ☐☐☐ / ☐☐☐ PAS/PAD

Por favor chequee TODA la encuesta
¡MUCHAS GRACIAS!

¿Cuál es su *TALLA* aproximada? (medirlo o preguntarlo) ☐☐☐ cm
¿Cuál es su *PESO* aproximado? (pesarlo o preguntarlo) ☐☐☐ kg
Fecha de la entrevista ☐☐☐☐☐☐ (a,m,d)
Calidad de la cooperación (1 = pobre, 2 = regular, 3 = buena) ☐ 1-3

Nombre y apellidos del entrevistador: _____

¿Algún comentario? .....

PARA LA SIGUIENTE PERSONA, UTILICE EL REVERSO DE ESTA PAGINA →

Mexico City Prospective Study Baseline Questionnaire

Autumn 1998

EPIDEMIOLOGIA

FAMILY QUESTIONNAIRE

PROYECTO COYOACAN

THE NEXT QUESTIONS MUST BE REGISTERED ON THE KEYBOARD OF THE LOP BEFORE WE KNOCK IN THE HOUSE (01 TO 05)

01 AGEB 02 LIST AREA 03 BLOCK
04 HOUSE 05 FAMILY CODE

Good morning (evening), my name is and I work for the Ministry of Health in Project Coyoacan: in accordance with the information that we have send you, we are carrying out a survey about you and your family's health; I would like to ask you some questions about the people that normally live in this house.

5A Do people 35 years or older live in this house?

5001 YES

5002 NO

End of the survey

5003 REJECT TO THE INTERVIEWER
PASS TO RECOVERY QUESTIONNAIRE

5009 absentees

In case that you found people from 25 years or older, register that in your agenda

06 How many people live in this house permanently?
(Register on the LOP's keyboard, 9 if Not know or Not answer)

07 How long ago have you lived in here?

72 From 1 to 4 years

73 From 5 to 9 years

71 Less than 1 year

79 Not know, Not answer

74 More than 9 years

EPIDEMIOLOGIA

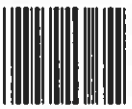
FAMILY QUESTIONNAIRE

PROYECTO COYOACAN

08 During the last ten years, how many times this family has changed address?



81
Never



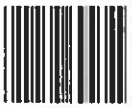
83
From 3 to 5 times



89
Not know, Not answer



82
From 1 to 2 times

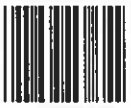


84
6 times or more

09 Do you rent or own this house?



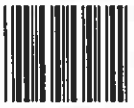
91
Owned



93
Other



92
Rented



99
Not know, Not answer

EPIDEMIOLOGIA

FAMILY QUESTIONNAIRE

PROYECTO COYOACAN

10 During the last ten years has anyone 35 years or older has died in this house?

101
YES

109
Not know, Not answer

102
NO

11 Results of the family questionnaire

111
Complete interview

112
Incomplete interview

113
Denied to give information

114
Inadequate informant

EPIDEMIOLOGIA

FAMILY QUESTIONNAIRE

PROYECTO COYOACAN

Would you please answer the following questions that I am going to ask you about every member of this family, starting with the head (he or she) of the family:

12 Register in LOP the Consecutive number +

13 Codes kinship cathegories +


131
Head of the family

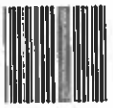

132
Husband/wife


133
Son/daughter


134
Grandson/grandaughter


135
Father/mother

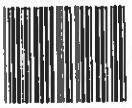

136
Brother/sister


137
Grandfather/grandmother

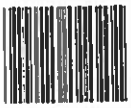

138
Uncle/aunt


139
Cousin


1310
Nephew/niece


1311
Father or mother in law


1312
Brother or sister in law


1399
Not know, Not answer

14 Digit the age and initials of the interviewee in LOP

(By example: Juan Pedro Martínez y Gonzales = JMG)

(9 if NK or NA)

EPIDEMIOLOGIA

INDIVIDUAL QUESTIONNAIRE

TO INITIATE IMEDIATLY DIGIT F3

PROYECTO COYOACAN

IDENTIFICATION CODE: REGISTER ON THE KEYBOARD TO LOAD THE INFORMATION: this information includes the AGEB + the list area + the block + the number of the house + the family code + the progressive number + The kinship category + the age + the cryovial number in case of blood sampling.

BEFORE YOU BEGIN 15 AGEB 16 LIST AREA 17 BLOCK 18 HOUSE 19 FAMILY CODE

20 Digit in LOP the consecutive number +

21 Codes for kinship cathegories +


211
Head of the familiy


212
Husband/wife


213
Son/daughter


214
Grandson/granddaughter


215
Father/mother


216
Brother/sister


217
Grandfather/grandmother


218
Uncle/aunt


219
Cousin


2110
Nephew/niece


2111
Father or mother in law


2112
Brother or sister in law


2199
Not know, Not answer

22 How old are you and what is your name?

(Digit the initials of the name in LOP ie. Juan José Camilo Mtz. Glez. = JMG)

(9 if NK or NA)

EPIDEMIOLOGIA

INDIVIDUAL QUESTIONNAIRE

PROYECTO COYOACAN

Now I would like to interview: _____.

Name of the interviewee

23 Sex

231
M

232
F

24 What is your marital status?

241
Single

242
Divorced

243
Living together

244
Separated

245
Married

246
Widowed

249
Not know, Not answer

EPIDEMIOLOGIA

INDIVIDUAL QUESTIONNAIRE

PROYECTO COYOACAN

25 What was your last school level?


251
Illiterate


256
Complete elementary


2511
Technical studies
with complete secondary level


252
Knows how to read

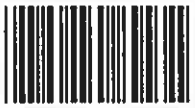

257
Technical studies
With complete elementary


2512
Complete high school


253
Knows how to read and write


258
Incomplete Secondary


2513
Technical studies
With complete high school


254
Incomplete elementary


259
Complete Secondary


2514
Incomplete university


255
Complete university

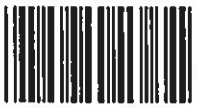

2510
Post graduate degree


2599
Not know, Not answer

26 How long have you been living in Coyoacan?


261
Less than 1 year


262
From 1 to 4 years


263
From 5 to 9 years


264
More than 9 years


269
Not know, Not answer

EPIDEMIOLOGIA

INDIVIDUAL QUESTIONNAIRE

PROYECTO COYOACAN

27 What is your present occupation?


271
Professional


272
Laborer or Handcrafter


273
Labor in the construction's industry


274
Technic


275
Helper or similar


276
Chauffeur/driver


277
Teacher


278
Public sector worker


279
Student


2710
Art worker


2711
Private sector worker


2712
Housework


2713
Bussinesman, manager


2714
Labor (blue collar)


2715
Retired


2716
Merchant/shopkeeper


2717
Domestic job


2718
Contractor


2719
Peddler


2720
Policeman or security guard


2721
Unemployed


2799
Not know, Not answer

EPIDEMIOLOGIA

INDIVIDUAL QUESTIONNAIRE

PROYECTO COYOACAN

28 Approximately, how much money do you earn every month?
Digit in LOP the quantity 9 if NK/NA)

29 Do you exercise or sport?



YES



NO
Pass to question 32



Not know, Not answer

30 How many times a week do you exercise?



Less than once a week



1 - 2 times a week



3 or more times a week



Not know, Not answer

31 When you exercise, How long, in minutes, do you work out?

311 Less than 30 min.

312 From 30 to 60 min.

313 More than 60 min.

319 Not know, Not answer

32 About your sleeping habits, How many hours do you sleep daily?

Digit in LOP the amount (9 if NK/NA)

In relation with your tobacco habits

33 Have you ever smoked tobacco?

331 YES

332 NO

Pass to question 44

339 Not know, Not answer

34 How old were you, when you started smoking regularly?

Digit in LOP the age (9 if NK/NA)

35 Do you smoke nowaday regularly (at present time)?



YES



NO

Pass to question 40



Not know, Not answer

36 With what frequency do you smoke?



Daily



Weekly



Monthly



Occasionally



Not know, Not answer

37 When you smoke, How many cigarettes do you smoke in a day?

Digit in LOP the amount (9 if NK/NA)

38 When you smoke tobacco, do you inhale?



YES



NO



Not know, Not answer

EPIDEMIOLOGIA

INDIVIDUAL QUESTIONNAIRE

PROYECTO COYOACAN

39 Have you ever quit smoking?


391

YES


392

NO

Pass to question 44


399

Not know, Not answer

40 How long ago did you quit the last time?


401

Less than 1 year


402

1 to 3 years


403

More than 3 years


409

Not know, Not answer

41 The last time that you quit smoking, How long did you give up?


411

Less than 1 year


412

1 to 2 years


413

More than 2 years


419

Not know, Not answer

42 When you smoked, With what frequency did you do it?



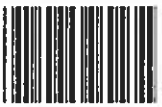
421
Daily



422
Weekly



423
Monthly



424
Occasionally



429
Not know, Not answer

43 How many cigarettes did you smoke?

Digit in LOP the amount (9 if NK/NA)

44 How many of the people that live in with you, smoke inside the house?

Digit in LOP the number of people (99 NK/NA)

EPIDEMIOLOGIA

INDIVIDUAL QUESTIONNAIRE

PROYECTO COYOACAN

About your alcohol consumption habits:

45 In the last 12 months, How often did you drink any beverage with alcohol? (Beer, Pulque, wine, brandy,whisky, rum, tequila, coolers, presidencolas, etc.)


451

Daily


452

2 or 3 times per month


453

1 to 5 times per year


454

3 or 4 times per week


455

more or less one time per month


456

Never in the last 12 months but yes before


457

1 or 2 times per week


458

6 to 11 times per year


459

Never has consumed alcohol

Pass to question 48


4599

Not know, Not answer

46 How many cups or glasses do you drink in one ocasion?

This means any combination of beer, wine, rum, etc.


461

1 to 2


462

3 to 4


463

5 to 7


464

8 to 11


465

12 or more


469

Not know, Not answer

47 What beverage did you drink most frequently?


471

Beer


472

Wine


473

Brandy,whiskey,tequila or rum


474

Pulque


475

Cooler


476

Pure alcohol


477

Other alcohol beverage


479

Not know, Not answer

EPIDEMIOLOGIA

INDIVIDUAL QUESTIONNAIRE

PROYECTO COYOACAN

48 Have you been diagnosed with any of the next diseases?
If answers affirmative, register in LOP the code of the disease, followed by the diagnostic decade code.

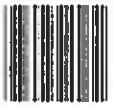

481
Emphysema


482
Heart attack


483
Angina pectoris


484
Diabetes


485
Asthma


486
Stroke


4813
Prostatic cancer


487
Chronic kidney insufficiency


488
Peptic ulcer


489
Cirrhosis of liver


4812
Other cancer


4810
Hypertension


4811
Lung cancer


4816
Oesophagus, stomach or intestine cancer


4814
Cervical cancer


4815
Breast cancer


4818
Peripheral arterial disease


4817
Oral, nasal or throat cancer


4899
Not know, Not answer


4819
None or end of the options

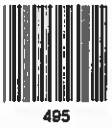
Pass to 50 if is a woman or to 58 if is a man

EPIDEMIOLOGIA

INDIVIDUAL QUESTIONNAIRE

PROYECTO COYOACAN

49 DECADE'S CODES OF THE DIAGNOSTIC



1950-1959



1960-1969



1970-1979



1980-1989



1990-1999



NOT REMEMBER

Pass to 50 if is a woman
or to 58 if is a man

EPIDEMIOLOGIA

INDIVIDUAL QUESTIONNAIRE

PROYECTO COYOACAN

QUESTIONS ONLY FOR WOMEN.

IF THE INTERVIEWEE IS A MAN PASS TO QUESTION 58.

Now I am going to ask you some questions about your life as a woman:

50 Do you still have your period (menstruation)?


501

YES

Pass to question 54


502

NO


509

Not know, Not answer

51 Have you had major surgery to remove your uterus?


511

YES


512

NO


519

Not know, Not answer

52 Have you had major surgery to remove your ovaries?


521

YES


522

NO


529

Not know, Not answer

EPIDEMIOLOGIA

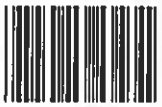
INDIVIDUAL QUESTIONNAIRE

PROYECTO COYOACAN

53 Have you recieved any hormonal therapy for the menopausic disturbs?



531
YES



532
NO

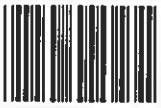


539
Not know, Not answer

54 From the next list of contraceptives, which one you have used more?



541
The Pill



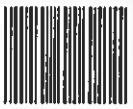
542
Syring hormonals



549
Not know, Not answer



543
Other



544
None

55 If you have had sex life, at what age did you have your first sexual relation?

Digit age in LOP (9 NK/NA)

56 How old were you when you get pregnant for first time?

Digit age in LOP (9 NK/NA)

57 How many pregnancies have you got in your life?

Digit quantity in LOP (9 NK/NA)

EPIDEMIOLOGIA

INDIVIDUAL QUESTIONNAIRE

PROYECTO COYOACAN

Now I am going to ask you some questions about your diet

58 How many days per week do you eat fruit and vegetables?


581
1 to 2


584
Never


582
3 to 4


589
Not know, Not answer


583
5 to 7

59 How many days per week do you eat fried food?
Like sopes, huearaches, fried beans, Quesadillas, etc.


591
1 to 2


594
Never


592
3 to 4

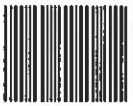

599
Not know, Not answer

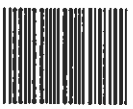

593
5 to 7

60 With what kind of oil do you cook your meals?

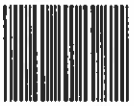

601
Olive oil

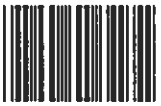

602
Other vegetal oil
(Cartamus, Soy)

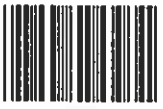

603
Butter


604
Margarine


605
Vegetal fat


606
Animal fat


607
None


609
Not know, Not answer

EPIDEMIOLOGIA

INDIVIDUAL QUESTIONNAIRE

PROYECTO COYOACAN

61 Have you take any medicine (including aspirins) during more than a month?


611

YES

Pass to question 63


612

NO


619

Not know, Not answer

62 When was the last time that you take any medicine for more time than a month?


621

2 to 3 months


623

More than 6 months


622

4 to 6 months


629

Not know, Not answer

63 Would you tell me the name of the medicine or prescription that you have **take for more than a month?**

Digit the code of the main prescriptions (maximum three)

(Not know, Not answer 999) (Other vgr. alternative medicine 001, None 000)

EPIDEMIOLOGIA

INDIVIDUAL QUESTIONNNAIRE

PROYECTO COYOACAN

To end the study I would like to make you some measures and take you a blood test:

64 Arterial Pressure (Three registers with 3 minutes between each one)

Digit the measurements consecutively

numbers below 100 have a 0 before ie. 90=090.

65 WEIGHT (kg)

66 HEIGHT (cm.)

67 Waist perimeter. (cm.)

68 Hip perimeter (cm.)

69 BLOOD SAMPLE'S RESULT

691

ADEQUATE BLOOD SAMPLE

692

INADEQUATE BLOOD SAMPLE

PASS TO 73

699

DENIED THE BLOOD SAMPLE

PASS TO 73

70 DATE AND HOUR OF THE BLOOD SAMPLE (write them consecutive)

71 DATE AND HOUR OF THE LAST MEAL (write them consecutive)

72 Cryovial code (only who accepted the blood sample)

REGISTER THIS LAST INFORMATION FROM THE ADHESIVE TAG OF THE VACUTAINER

EPIDEMIOLOGIA

INDIVIDUAL QUESTIONNAIRE

PROYECTO COYOACAN

73 Interview result



Denied to give information



Complete interview with anthropometry and blood sample



Complete interview without anthropometry and without blood sample



Complete interview with blood sample but without anthropometry



Complete interview with anthropometry but without blood sample

EPIDEMIOLOGIA

RECOVERY QUESTIONNAIRE

TO INITIATE IMMEDIATLY ACTIVATE F4

PROYECTO COYOACAN

IDENTIFICATION CODE: REGISTER ON THE KEYBOARD TO LOAD THE INFORMATION: this information includes the AGEB + the list area + the block + the number of the house

BEFORE YOU BEGIN 74 AGEB 75 LIST AREA 76 BLOCK 77 HOUSE

78 How many people 35 years or older live in this house?

Digit the quantity in LOP

If nobody is 35 years or older digit 77 GIVE THANKS AND STOP THE INTERVIEW

If still rejects the interview digit 88 GIVE THANKS AND STOP THE INTERVIEW

If not know or not answer digit 99 GIVE THANKS AND STOP THE INTERVIEW

79 How many men 35 years or older live in this house?
Digit quantity in LOP (99 NK/NA)

80 How many women 35 years or older live in this house?
Digit quantity in LOP (99 NK/NA)

81 How many people smoke in this house?
Digit quantity in LOP (99 NK/NA)

82 How many people drink alcoholic beverages in this house?
Digit quantity in LOP (99 NK/NA)

83 During the last ten years how many people 35 years or older has died in this house?
Digit quantity in LOP (99 NK/NA)

84 How many people in this house has actually any disease like hipetension, cancer or diabetes?
Digit quantity in LOP (99 NK/NA)

Mexico City Prospective Study Resurvey Questionnaire

Section 0: Information to be pre-loaded onto tablet prior to home visit (interview should only take place if all of the following are confirmed, ie, the correct person is definitely identified)

0.1 Study ID (registro): ☐☐☐☐☐☐☐☐ Baseline CADENA: ☐☐☐☐☐☐☐☐☐☐☐☐

0.2 Name：_____

0.3 Sex：Male ☐ Female ☐

0.4 Age at baseline ☐☐☐ Approximate age <u>today</u> ☐☐☐

0.5 Position in the family at baseline ☐☐☐

0.6 Address: District ☐☐ Basic geographical area ☐☐ Listed area ☐☐

Block ☐☐ Household ☐☐

Section 1: Availability / willingness to participate in resurvey

1.1 Upon arrival at the <u>expected</u> location of the house, the tablet should record the time, date, tablet number and GPS location.

1.2 Is (or was) this address the correct address for the person identified in section 0?

- ☐ Yes →***If ticked, Go to <u>1.3</u>***
- ☐ Nobody at home →***If ticked, Go to <u>section 9</u>***
- ☐ Person refuses to speak to interviewer →***If ticked, Go to <u>section 9</u>***
- ☐ Person speaks to interviewer but house is not the correct address →***If ticked, Go to <u>1.5</u>***
- ☐ House no longer exists →***If ticked, Go to <u>section 9</u>***

1.3 Is the person in section 0 at home?

- ☐ Yes →***If ticked, Go to <u>1.4</u>***
- ☐ No (they're dead) → ***If ticked, Go to <u>1.6</u>***
- ☐ No (but person is alive) → ***If ticked, Go to <u>Section 9 (but potentially arrange revisit)</u>***

1.4 Is the person willing to take part in the resurvey (now)?

- ☐ Yes →***If ticked, Go to <u>Section 2</u>***
- ☐ No →***If ticked, Go to <u>Section 9 (but potentially arrange revisit)</u>***

1.5 Has the person moved house or died?

- ☐ Moved house →***If ticked, record new address (if known) and the approximate date they moved: then go to <u>Section 9</u>***
- ☐ Died →***If ticked, Go to <u>1.6</u>***

1.6 Dead: Record summary information about the person and their death (and arrange a time for a revisit by mortality fieldworkers)

Approximate date of birth: Day / Month / Year | | | / | | | / | | | | |

or ☐ don't know/no answer

Approximate date of death: Day / Month / Year | | | / | | | / | | | | |

or ☐ don't know/no answer

(leave day/month blank if uncertain)

Place of death: …… or ☐ don't know/no answer

Cause of death: …… or ☐ don't know/no answer

(If known: describe your best estimate from what the informant can tell you, and say that a doctor may come later to ask for more details)

→ ***Then, Go to <u>Section 9</u>***

Section 2: Background information from person in section 0

2.1 Date of birth | | | / | | | / | | | | |

2.2 Marital status

Single	☐	Divorced	☐	Living together	☐
Separated	☐	Married	☐	Widowed	☐
Don't know / no answer	☐				

2.3 Reading and writing

Illiterate	☐	Able to read	☐	Able to read and write	☐

[If illiterate, then go straight to question 2.5]

2.4 Highest attained educational level

Incomplete elementary	☐	Complete elementary	☐	Technical with complete elementary	☐
Incomplete secondary	☐	Complete secondary	☐	Technical with secondary	☐
High school	☐	Technical with high school	☐	Incomplete university	☐
Complete university	☐	Post graduate degree	☐	Don't know / no answer	☐

2.5 Are you mostly retired?

☐ Yes
☐ No
☐ Don't know/no answer

2.6 What is your present main occupation (or your previous main occupation if you're now mostly retired)?

Professional	☐	Laborer or Handcrafter	☐
Labor in construction industry	☐	Technical	☐
Helper (or similar)	☐	Chauffeur / driver	☐
Teacher	☐	Public Sector Worker	☐
Student	☐	Art related worker	☐
Private sector worker	☐	Housework	☐
Businessman/manager	☐	Labor (blue collar)	☐
Merchant/shopkeeper	☐	Domestic job	☐
Peddler	☐	Policeman/security guard	☐
Landlord	☐	Unemployed	☐
Don't know / no answer	☐		

2.7 Approximately, how much money do <u>you</u> earn or receive per month (includes any pensions or benefits)?

Either | | | | | | | | pesos per month or ☐ don't know / no answer

2.8 Approximately, how much money does your <u>family</u> in this house earn per month?

Either | | | | | | | | pesos per month or ☐ don't know / no answer

2.9 Who is your main health service provider?

SSA	☐	IMSS	☐	ISSSTE	☐	SEDANA	☐
SECMAR	☐	PEMEX	☐	INI	☐	Private medic	☐
Medic dispensaries	☐	Other	☐	Seguro Popular	☐	None	☐
Don't know /no answer	☐						

Section 3: Lifestyle

PHYSICAL ACTIVITY

3.1 Do you do exercise or play sport?

- ☐ Yes
- ☐ No →***If ticked, Go to 3.4***
- ☐ Don't know/no response →***If ticked, Go to 3.4***

3.2 How many days a week do you do exercise or play sport?

☐☐ Days ☐ Don't know/no response

3.3 When you exercise or play sport, how long (in minutes) do you usually work out/play?

☐☐☐ Minutes ☐ Don't know/no response

3.4 How many hours a day would you normally spend on your feet?

☐☐ hours ☐ Don't know/no response

3.5 How many days a week do you walk for 10 minutes or more?

☐☐ days ☐ Don't know/no response

[If answer to question 3.5 is 0 days, then go to question 3.7]

3.6 On these day(s), how much time (in hours) do you spend walking?

☐☐ hours ☐ Don't know/no response

SLEEPING HABITS

3.7 How many hours do you normally sleep each night?

☐☐ Hours ☐ Don't know/no response

SMOKING

3.8 Have you ever smoked tobacco?

- ☐ Yes
- ☐ No →***If ticked, Go to 3.21***
- ☐ Don't know →***If ticked, Go to 3.21***

3.9 How old were you when you first started smoking?

☐☐ years ☐ Don't know/no response

3.10 Do you smoke now?

☐ Yes

☐ No →***If ticked, Go to <u>3.17</u>***

☐ Don't know / no answer →***If ticked, Go to <u>3.21</u>***

3.11 CURRENT SMOKERS: About how many days per week do you smoke?

Either ☐ days or ☐ Don't know / no answer

3.12 CURRENT SMOKERS: On the days that you smoke, how many cigarettes do you normally smoke in a day?

Either ☐☐☐ cigarettes or ☐ Don't know / no answer

3.13 CURRENT SMOKERS: When you smoke, do you inhale?

☐ No (mouth only)

☐ Yes (mouth and throat only)

☐ Yes (Into lung)

☐ Don't know / no answer

3.14 CURRENT SMOKERS: Have you ever tried to quit smoking?

☐ Yes

☐ No

☐ Don't know / no answer

3.15 CURRENT SMOKERS: Do you want to quit smoking sometime soon?

☐ Yes

☐ No

☐ Don't know / no answer

3.16 CURRENT SMOKERS: How many cigarettes have you smoked today?

Either ☐☐☐ cigarettes or ☐ Don't know / no answer

➔ **Then, go to <u>3.21</u>**

3.17 EX-SMOKERS: How many years ago did you last quit?

Either ☐☐ years ago or ☐ Don't know / no answer

3.18 EX-SMOKERS: What age were you when you last quit?

Either ☐☐ years old or ☐ Don't know / no answer

3.18b EX-SMOKERS: Did you quit because you were already ill?

☐ Yes

☐ No

☐ Don't know / no answer

3.19 EX-SMOKERS: How many days per week did you used to smoke?

Either ☐ days or ☐ Don't know / no answer

3.20 EX-SMOKERS: On the days that you used to smoke, how many cigarettes did you usually smoke in a day?

Either ☐☐☐ cigarettes or ☐ Don't know / no answer

3.21 How many of the other people who live with you smoke inside the house?

Either ☐☐ people or ☐ Don't know / no answer

DRINKING ALCOHOL

3.22 Have you drunk any alcohol in the past year?

☐ Yes

☐ No

☐ Don't know / no answer

3.23 How often do you drink alcohol?

Either ☐ days per week or ☐ Less frequently

or ☐ Used to drink alcohol but not any more
→***If ticked, Go to 3.29***

or ☐ Never drunk alcohol →***If ticked, Go to section 4***

or ☐ Don't know / no answer →***If ticked, Go to section 4***

3.24 CURRENT DRINKERS: When you drink, how many cups/glasses do you normally drink during a single session? (this means any combination of beer, wine, rum etc.)

Either ☐☐ drinks per session or ☐ Don't know / no answer

3.25 CURRENT DRINKERS: What do you drink most often?

- ☐ Beer
- ☐ Wine
- ☐ Spirits (eg, Brandy, whiskey, tequila or rum)
- ☐ Pulque
- ☐ Cooler
- ☐ Pure alcohol
- ☐ Other alcoholic beverage
- ☐ Don't know / no answer

3.26 CURRENT DRINKERS: Have you ever reduced the amount of alcohol you drink because you were ill?

- ☐ Yes
- ☐ No
- ☐ Don't know / no answer

3.27 CURRENT DRINKERS: Do you mainly drink with your meals, or at other times?

- ☐ Mainly with your meals
- ☐ Mainly not with your meals
- ☐ No regular pattern
- ☐ Don't know / no answer

3.28 CURRENT DRINKERS: During the past month, how often have you drunk alcohol in the morning?

- ☐ Never
- ☐ <1 day/week
- ☐ A few days a week
- ☐ Daily or almost daily
- ☐ Don't know / no answer

➔ **Then, go to section 4**

3.29 EX-DRINKERS: Did you stop drinking because you were ill?

- ☐ Yes
- ☐ No
- ☐ Don’t know / no answer

Section 4: Previous diseases

4.1 Have you ever been diagnosed with any of the following diseases (tick all that apply and tell us the year it was first diagnosed)?

	Yes	No	Don't know	Year FIRST diagnosed
Chronic obstructive pulmonary disease	☐	☐	☐	☐☐☐☐
Chronic bronchitis	☐	☐	☐	☐☐☐☐
Tuberculosis	☐	☐	☐	☐☐☐☐
Emphysema	☐	☐	☐	☐☐☐☐
Heart attack	☐	☐	☐	☐☐☐☐
Angina pectoris	☐	☐	☐	☐☐☐☐
Diabetes	☐	☐	☐	☐☐☐☐
Asthma	☐	☐	☐	☐☐☐☐
Stroke	☐	☐	☐	☐☐☐☐
Chronic kidney disease	☐	☐	☐	☐☐☐☐
Peptic ulcer	☐	☐	☐	☐☐☐☐
Cirrhosis of the liver	☐	☐	☐	☐☐☐☐
Hypertension	☐	☐	☐	☐☐☐☐
Parkinsons	☐	☐	☐	☐☐☐☐
Dementia	☐	☐	☐	☐☐☐☐
Osteoporosis	☐	☐	☐	☐☐☐☐
Cancer	☐	☐	☐	☐☐☐☐
(Option to enter site(s) only appears if answer to "Any cancer" question is "Yes")				
Lung	☐	☐	☐	☐☐☐☐
Prostate **(only appears if 0.3="male")**	☐	☐	☐	☐☐☐☐
Cervical **(only appears if 0.3="female")**	☐	☐	☐	☐☐☐☐
Breast **(only appears if 0.3="female")**	☐	☐	☐	☐☐☐☐
Oesophageal, stomach or intestinal	☐	☐	☐	☐☐☐☐
Oral, nasal, or throat	☐	☐	☐	☐☐☐☐
Other site	☐	☐	☐	☐☐☐☐

[If any of heart attack, angina or stroke are ticked "Yes" in 4.1, then ask question 4.2.1]:

4.2.1 Are you taking any drugs for your angina, heart attack, or stroke?

☐ Yes ☐ No ☐ Don't know / no answer

Name of medication (if yes above)

1. ……..

2. **……..**
3. **……..**
4. **etc,…**

[If breast cancer is ticked "Yes" in 4.1, then ask questions 4.3.1, 4.3.2 and 4.3.3]:

4.3.1 Are you taking tamoxifen (or Nolvadex)?

- ☐ Yes
- ☐ No
- ☐ Don't know / no answer

4.3.2 Are you taking an aromatase inhibitor (eg Arimidex)?

- ☐ Yes
- ☐ No
- ☐ Don't know / no answer

4.3.3 Was your breast cancer hormone receptor positive (ER+)?

- ☐ Yes
- ☐ No
- ☐ Don't know / no answer

[If diabetes is ticked "Yes" in 4.1, then ask questions 4.4.1 to 4.4.9]:

4.4.1 Was your diagnosis of diabetes during pregnancy only? (only asked if subject is female)

- ☐ Yes
- ☐ No
- ☐ Don't know / no answer

4.4.2 How frequently do you visit your doctor to control your diabetes?

- ☐ Daily
- ☐ Weekly
- ☐ Monthly
- ☐ A few times a year
- ☐ About once a year
- ☐ Never
- ☐ Don't know / no answer

4.4.3 How do you currently control your diabetes?

- ☐ Insulin only
- ☐ Pills only → Go to 4.4.3b
- ☐ Both insulin and pills → Go to 4.4.3b
- ☐ Neither insulin or pills
- ☐ Don't know / no answer

4.4.4 What pills do you take to control your diabetes?

1. **……..**
2. **……..**
3. **……..**
4. **etc…**

4.4.5 Do you have any damage to your eye's retina or have you lost your sight in either eye?

- ☐ Yes
- ☐ No
- ☐ Don't know / no answer

4.4.6 Have you ever been in a diabetic coma?

- ☐ Yes
- ☐ No
- ☐ Don't know / no answer

4.4.7 Have you ever had a limb or other part of your body amputed?

- ☐ Yes
- ☐ No
- ☐ Don't know / no answer

4.4.8 Have you ever received dialysis?

- ☐ Yes, but not now
- ☐ Yes, currently
- ☐ No
- ☐ Don't know / no answer

4.4.9 Have you ever received a kidney transplant?

- ☐ Yes
- ☐ No
- ☐ Don't know / no answer

4.5 Mini mental state questionnaire (asked to all participants)

4.5.1 What year is it?

4.5.2 What season is it?

4.5.3 What is today's date?

4.5.4 What day of the week is it?

4.5.5 What month is it?

4.5.6 What country are we in?

4.5.7 What city are we in?

4.5.8 What delegation are we in?

4.5.9 What street are we in?

4.5.10 What house number are we?

4.5.11 Name three objects at intervals of one second and then ask the interviewee to repeat the three

4.5.12 Spell "world " backwards.

4.5.13 Repeat the three objects learned in question 4.5.11.

4.5.14 Point to a pen and then a watch and have the participant tell you what they are.

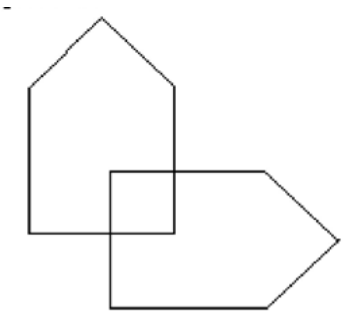
4.5.15 Have the participant repeat the phrase "No ifs, ands or buts "

4.5.16 Ask the participant to take a piece of paper with their right hand, fold it in half and put it on the floor.

4.5.17 Ask participant to read and obey a written command on the tablet stating "Close your eyes".

4.5.18 Have the participant write a sentence (any sentence).

4.5.19 Ask the participant to copy the following image onto the tablet (using the tablet pen);



Section 4.6: Further questions related to osteoporosis and fractures
[Should be asked to all participants].

4.6.1 Have you ever had a fracture?

☐ Yes
☐ No → **Go to 4.6.5**
☐ No answer/no response → **Go to 4.6.5**

4.6.2 How many fractures have you experienced during your whole life?

☐☐ Fractures ☐ No answer / no response

4.6.3 How long ago did you have your most recent fracture?

Either ☐☐ years or ☐☐ months

☐No response / no answer

4.6.4 Which of the following have you fractured?

☐ Hip
☐ Spine
☐ Femur
☐ Wrist
☐ Arm
☐ Other
☐ Don't know / no answer

4.6.5 In the past 12 months, have you experienced any serious fall?

☐ Yes
☐ No → Go to 4.6.7
☐ No answer/no response → Go to 4.6.7

4.6.6 How many times have you fallen in the past 12 months?

☐☐ times ☐ Don't know / no answer

4.6.7 Have any of your children, siblings or parents ever had a hip fracture?

☐ Yes
☐ No
☐ No answer/no response

Section 5: Reproductive history (only if participant is female)

5.1 Do you still have periods (menstruation)?

- ☐ Yes →***If ticked, Go to 5.6***
- ☐ No
- ☐ Don't know / no answer

5.2 Have you had surgery to remove your uterus?

- ☐ Yes
- ☐ No
- ☐ Don't know / no answer

5.3 Have you had surgery to remove your ovaries?

- ☐ Yes
- ☐ No
- ☐ Don't know / no answer

5.4 Have you ever received hormone replacement therapy for the menopause?

- ☐ Yes
- ☐ No
- ☐ Don't know / no answer

5.5 (If Yes to question 5.4) Are you currently taking hormone replacement therapy?

- ☐ Yes
- ☐ No
- ☐ Don't know / no answer

5.6 Which of the following contraceptives have you used most often?

- ☐ The pill
- ☐ Syringe hormonal
- ☐ Other contraceptive
- ☐ No contraceptive
- ☐ Don't know / no answer

Section 6: Diet

6.1 How many days a week do you normally eat fruit?

Either ☐ days or ☐ Don't know / no answer

6.2 How many days a week do you normally eat vegetables?

Either ☐ days or ☐ Don't know / no answer

6.3 How many days a week do you normally eat fried food (like spes, huaraches, fried beans, quesadillas)?

Either ☐ days or ☐ Don't know / no answer

6.4 What kind of oil/fat do you normally use for cooking?

- ☐ Olive oil
- ☐ Another vegetable oil (caramus, soya)
- ☐ Butter
- ☐ Margarine
- ☐ Animal fat
- ☐ None
- ☐ Don't know / no answer

6.5 How often do you normally drink sugary drinks (eg, sodas, aquas frescas)?

Either ☐ days or ☐ Don't know / no answer

If 0 days per week then → Question 6.7

6.6 On the days you drink sugary drinks, how many such drinks would you normally have?

Either ☐ drinks or ☐ Don't know / no answer

6.7 How often do you usually add salt to your food?

- ☐ Never
- ☐ Rarely
- ☐ Sometimes
- ☐ Always

6.8 How often do you normally eat meat (eg, pork, beef, chicken)?

Either ☐ days or ☐ Don't know / no answer

6.9 How often do you normally eat fish?

Either ☐ days or ☐ Don't know / no answer

6.10 How often do you normally eat between meals?

Either ☐ days or ☐ Don't know / no answer

6.11 How often do you normally eat desserts?

Either ☐ days or ☐ Don't know / no answer

6.12 Are you on any special diet?

- ☐ No
- ☐ Yes (low calorie, eg, for diabetes/weight loss)
- ☐ Yes (low salt, eg, for hypertension)
- ☐ Yes (low phosphate/potassium, eg, for kidney disease)
- ☐ Yes (other type)

Section 7: Current Medication

7.1 Are you currently taking any medication (eg, aspirin) that you have been taking regularly for a period <u>of more than one month</u>?

☐ Yes

☐ No →***If ticked, Go to <u>Section 8</u>***

☐ Don't know / no answer →***If ticked, Go to <u>Section 8</u>***

7.2 Please list these long-term medications.

Name of medication

1. ……..

2. ……..

3. ……..

4. ……..

5. ……..

etc.

[For question 7.2, the tablet should automatically populate previous medicines indicated by questions 4.2.1, 4.3.1, 4.3.2 and 4.4.4]

Section 8: Physical examination and blood sample

8.1 Blood pressure (to be measured after 3 minutes in the seated position): **data to be transferred to tablet automatically via Bluetooth**

	First	Second	Third	
SBP	☐☐☐	☐☐☐	☐☐☐	mmHg
DBP	☐☐☐	☐☐☐	☐☐☐	mmHg

8.2 Standing height (without shoes) to nearest cm. ☐☐☐ cm

8.3 Waist circumference to nearest cm …….…… ☐☐☐ cm

8.4 Hip circumference to nearest cm …….. ☐☐☐ cm

[If the participant is willing, take two photographs of them with their back against a wall, each from a distance of about 2 metres, one side facing and one front facing]

8.5 Weight (without shoes, but in light clothing), kgs
(data to be transferred to tablet automatically via USB cable)

8.6 Other measurements recorded by full body analyser (Tanita SC240-MA)
(Other data to be transferred to tablet automatically via USB: body type, preset tare, fat %, fat mass, FFM, muscle mass, muscle score, bone mass, body water, body water %, BMI, standard body weight, degree of obesity, visceral fat rating, BMR (kJ), BMR (kcal), BMR score, metabolic age, Rohrer's index, Impedance (50kHz))

8.7 Spirometry measurements
(Data to be transferred to tablet automatically via USB: FVC, FEV1, FEV6, FEV1/FVC%, PEF, FEF25-75%, ELA), quality grade (A, B, C, D, E), test completion (Yes, No, Refused, Medically excluded)

8.8 Did you take any drugs to lower blood pressure in the last 2 days? ☐ Yes ☐ No

8.9 Hours since last ate or drank anything (ignore non-caloric drinks such as water and tea)? _____ hours

8.10 Blood sample collected: Yes ☐ Partial ☐ Failed ☐ Refused ☐

8.11 Scanned barcode of the vacutainer (to be recorded automatically by the tablet)

8.12 Date and hour of the blood sample (to be recorded automatically by the tablet on completion of 8.10)

8.13 Interviewer's assessment of subject's cooperation

Excellent ☐ Good ☐ Fair ☐ Poor ☐

DERIVED BY TABLET

Result of interview

Refused to give information	☐
Full interview with measurements and blood sampling	☐
Full interview but without measurements or blood samping	☐
Full interview with blood simple but without measurements	☐
Full interview with measurements but without blood sampling	☐

Section 9: Auditing data (automatically captured on completion of 8.13)

9.1 Tablet to automatically capture date and time

9.2 Tablet to automatically capture name of fieldworker

9.3 Tablet to automatically capture GPS co-ordinates

END OF SURVEY