The Dynamics of Diabetes among Birth Cohorts in the United States

Short running title: Cohort Dynamics of Diabetes

Ezra I. Fishman, Andrew Stokes, and Samuel H. Preston

All authors are with the Population Studies Center, University of Pennsylvania, Philadelphia, Pennsylvania.

Corresponding author:
Ezra I. Fishman
Population Studies Center, University of Pennsylvania
239 McNeil Building
3718 Locust Walk
Philadelphia, PA 19104-6298
Phone: 215-898-6441, Fax: 215-898-2124
fishmane@sas.upenn.edu

Abstract

Objective: Using a nationally representative sample of the civilian noninstitutionalized U.S. population, we estimated trends in diabetes prevalence across cohorts born 1910 to 1989 and provide the first estimates of age-specific diabetes incidence using nationally representative, measured data.

Research design and methods: Data were from 40,130 nonpregnant individuals aged 20-79 years who participated in the third National Health and Nutrition Examination Survey (NHANES III), 1988-1994, and the continuous 1999-2010 NHANES. We defined diabetes as hemoglobin $A_{1C} \geq$ 6.5% (48 mmol/mol) or taking diabetes medication. We estimated age-specific diabetes prevalence for the five-year age groups 20-24 through 75-79 for cohorts born 1910-1919 through 1980-1989 and calendar periods 1988-1994, 1999-2002, 2003-2006, and 2007-2010. We modeled diabetes prevalence as a function of age, calendar year, and birth cohort and used our cohort model to estimate age-specific diabetes incidence.

Results: Age-adjusted diabetes prevalence rose by a factor of 4.9 between the birth cohorts of 1910-19 and 1980-89. Diabetes prevalence rose with age within each birth cohort. Models based on birth cohorts show a steeper age-pattern of diabetes prevalence than those based on calendar years. Diabetes incidence peaks at ages 55 to 64.

Conclusions: Diabetes prevalence has risen across cohorts born through the 20th century. Changes across birth cohorts explain the majority of observed increases in prevalence over time. Incidence peaks between ages 55 and 64 and then declines at older ages.

Diabetes is a leading cause of death in the United States (1). A recent meta-analysis estimates that people with diabetes have a 50-80% increased risk of disability, including impaired mobility, activities of daily living, and instrumental activities of daily living, compared to people without diabetes (2). The prevalence of diabetes among adults is approximately 12%, corresponding to approximately 26.1 million adults with diabetes in 2005-10 (3).

The incidence and prevalence of type 2 diabetes, which accounts for over 90% of diabetes cases (4), are clearly related to factors in an individual's past. In particular, individuals' own histories of obesity and smoking (5,6) have been shown to affect the risk of developing diabetes. Of these risk factors, the relationship between obesity history and diabetes incidence has been studied more extensively. One study found a steep gradient in the lifetime risk of diabetes based on body mass index (BMI, measured in kilograms per meters squared) at age 18. Males in the optimal BMI range of 18.5 to 25 kg/m^2 at age 18 had a 19.8% lifetime risk of diabetes, while males with BMI in the obese range of 30 to 35 kg/m^2 at age 18 had a 57.0% lifetime risk of diabetes (7). A European cohort study found that the earlier in life that subjects gained weight, the more likely they were to develop diabetes (8). Among subjects in the Framingham Heart Study, each additional two years of obesity were associated with about a 12% increased odds of developing diabetes (9). In the National Longitudinal Study of Adolescent Health, persistent obesity was associated with twice the risk of diabetes prevalence compared to adult-onset obesity (10). In the CARDIA study, each additional year a person was obese increased their odds of developing diabetes by 4% (11). These and other studies indicate that obesity over the life course is an important predictor of diabetes incidence.

In this paper, we investigate the rise in diabetes in the United States through the lens of birth cohorts. Previous studies examining changes in diabetes prevalence over time have

compared one calendar-year period to another (3,12). However, like other chronic diseases, type 2 diabetes is the result of cumulative processes that develop over a lifetime. A full understanding of the prevalence of diabetes at a moment in time requires reference to the past, a past that is embodied in the birth cohorts alive during that period. Because histories in a birth cohort are persistent – characteristics of a birth cohort established at age 25 remain the age-25 characteristics of that cohort as it ages – we expect to find "cohort effects" that differentiate one birth cohort from another as they age.

Birth cohorts not only embody a history of exposures, they are also the appropriate vehicle for calculating disease incidence. We take advantage of this opportunity to present new estimates of the age-pattern of diabetes incidence in the United States. These are the first estimates of incidence that use measured data in a nationally representative sample. Previous national estimates of diabetes incidence used retrospective reports of individuals rather than biological indicators and provided little age detail (13,14).

RESEARCH DESIGN AND METHODS

Population and data collection

In order to investigate the dynamics of diabetes in the United States, we use data from the National Health and Nutrition Examination Surveys (NHANES). We employ data from NHANES III, conducted in two phases, 1988 to 1991 and 1991 to 1994; and from the Continuous NHANES that began in 1999, for which data are released in two-year cycles. We pool adjacent data-release cycles of Continuous NHANES to obtain three observation periods from Continuous NHANES: 1999 to 2002, 2003 to 2006, and 2007 to 2010. NHANES is a complex, multi-stage probability sample of the U.S. civilian non-institutionalized population. Participants complete a home interview and are then examined in a mobile examination center,

which includes sampling participants' blood for laboratory tests. Participants are randomized into morning or afternoon examinations, and the morning examinees are asked to fast for at least nine hours prior to the examination. Whenever possible, NHANES uses consistent laboratory procedures over time to facilitate analysis of trends in population health. The National Center for Health Statistics (NCHS) provides extensive documentation of NHANES survey, examination, and laboratory procedures on its website (15). The characteristics of the NHANES study sample are reported elsewhere (3,12).

There were 88,224 individuals examined during our study periods. We exclude individuals below age 20 (n=40,899), above age 80 (n=3,558), or who were pregnant (n=1,510). We also exclude individuals who were exactly 20 years old when surveyed in 2010 (n=105) because these individuals would not comprise a complete birth cohort, as described below. We also exclude subjects with missing HbA1c values (n=2,022). The final analytic sample for HbA1c-based measures consists of 40,130 observations, with 7,011 observations from Phase 1 of NHANES III, 7,427 from Phase 2 of NHANES III, 7,778 from NHANES 1999-2002, 7,755 from NHANES 2003-2006, and 10,159 from NHANES 2007-2010.

Definition of diabetes

We rely on laboratory results, rather than self-reported diagnoses, because the latter fails to capture the considerable number of individuals in the US population with undiagnosed diabetes. A 2010 study estimated that 3.9 million individuals above age 20 had undiagnosed diabetes, representing 19% of the diabetic population (16). Furthermore, intertemporal comparisons based on self-reported diagnosis are complicated by the fact that criteria for diagnosing diabetes in the clinical setting have changed (17,18).

Our primary definition of diabetes is based on HbA1C, which was first measured in NHANES III. This measure reflects average glycemia over a prolonged period and thus has more intra-subject stability than the leading alternative, a measure of fasting plasma glucose (FPG) (19). Furthermore, HbA1c-based measures of diabetes are more strongly associated with cardiovascular disease and death than are FPG-based measures (20). Finally, only 54% as many observations of diabetes status are available in NHANES using FPG as using HbA1c.

Several changes in laboratory measurement of HbA1C occurred over the course of Continuous NHANES (detailed elsewhere (12)), but we follow the NCHS recommendation and the methods of recent studies and used HbA1C data without any corrections or adjustments (3,12). Individuals are considered diabetic if they had HbA1c $\geq$ 6.5% (48 mmol/mol) (4). Because diabetes medication is expected to reduce glycemia, the HbA1c values of medicated persons might not capture their diabetes status correctly; therefore, all individuals who reported taking diabetes medication are considered diabetic. In our sample, there were 4,678 individuals who met our definition of having diabetes. There were 896 individuals, or 19.2% of the group with diabetes, who reported taking diabetes medication and who had HbA1c < 6.5%.

Cohort assignment

Birth cohorts must be constructed from repeated cross-sections because NHANES does not repeatedly sample the same individuals over time. We calculate each individual's birth year using the equation *Birth cohort = Period - Age*. For the purpose of calculating birth cohorts, *Period* is defined as the midpoint of the NHANES wave or phase: April 21, 1990 for Phase 1 of NHANES III, April 23, 1992 for Phase 2 of NHANES III, and January 1 of the second year of each data release cycle of Continuous NHANES. In a recent study of cohort obesity patterns that used NHANES data and the same procedure for calculating birth years, results were robust to

alternative specifications of *period* (21). *Age* is the age of the individual, in completed years, at the time of the survey. To ensure large enough age-cohort cells, we analyze cohorts born in ten-year-wide intervals (1910 to 1919, 1920 to 1929, etc.). Using this approach, we obtain a total of 8 ten-year birth cohorts between 1910-1919 and 1980-1989. This method involves assuming that upon reaching age 20, diabetes prevalence is not affected by migration. We test the sensitivity of our results to this assumption by excluding foreign-born individuals from the sample.

Statistical methods

Prevalence is calculated as the proportion of individuals in the given age-period or age-cohort cell with diabetes as defined above. Calculations are adjusted for complex survey design using strata and primary sampling units provided by the National Center for Health Statistics (NCHS), along with survey weights. For HbA1c, we use the final examination weight provided by NCHS; because we pool adjacent data release cycles of Continuous NHANES, we divide the examination weights in Continuous NHANES by 2, as recommended by NCHS (22).

We then used OLS regression to model the age-, cohort- and period-patterns of diabetes prevalence in the U.S. population. We regressed the log of the prevalence estimate on a series of age and cohort or age and period indicators, with each prevalence estimate weighted by the number of observations that gave rise to it. Then, in an age-period-cohort model, we regressed the log of the prevalence estimate on age and period indicators, plus a continuous variable equal to the prevalence of obesity at age 25 in the corresponding birth cohort. We use age 25 because NHANES inquired about weight at that specific age. Obesity at age 25 serves as an measure of a cohort's history of obesity. The use of a continuous variable to represent birth cohort influences avoids the identification problem that any two of age, cohort, and period indicators can be linearly combined to produce the third (23).

Birth-cohort obesity prevalence is estimated using age-25 weight and height recall data in Continuous NHANES waves 1999-2008. Height recall was only asked of participants aged 50 and over; for younger individuals we used self-reported current height. We identify birth cohorts by subtracting age from survey year, using the beginning of the second year of each of the waves (e.g., 2000.0 for 1999-2000) and aggregate them into five-year wide intervals. The earliest and most recent birth cohorts for whom cohort obesity is calculated are the 1920-1924 and 1975-1979 birth cohorts, respectively. Thus, the age-period-cohort model excludes prevalence estimates that drew exclusively from the oldest or youngest birth cohorts (born 1910-1919 and 1980-1989). Appendix 1 shows a table of the obesity prevalence values used in this study.

The examination of diabetes prevalence within birth cohorts allows us to estimate the age-specific incidence of diabetes. In essence, this estimate is made by dividing the prevalence of non-diabetes in a birth cohort at one age interval (e.g. 50 to 54) by the prevalence of non-diabetes in the same birth cohort in the adjacent, younger age interval (e.g. 45 to 49) and adjusting for the fact that people without diabetes die at lower rates than the general population. The prevalence estimates used in this calculation are based upon the age coefficients estimated from the age/cohort model, presented in Figure 3B. These summarize the age-pattern of prevalence revealed within eight birth cohorts, adjusting for cohort-specific effects. Life tables for individuals without diabetes and for the general population are estimated using pooled data from NHANES III and Continuous NHANES (1999-2004 waves) cohorts linked to deaths in the National Death Index through 2006 (24). A discrete hazards model on a person-month file is employed to generate the underlying risks for predicting mortality rates. The model is implemented on baseline ages 20-74. There were 2,903 deaths among 25,971 respondents.

Derivation of the formula for estimating incidence is shown in Appendix 2. In deriving the formula, we assume that the prevalence of diabetes is not affected by migration beyond age 20. Furthermore, we assume that, once one becomes diabetic, diabetes is never cured. To smooth the incidence series, we employ a three-term moving average. The use of a moving average to infer incidence is appropriate because of the likelihood of offsetting errors in adjacent age intervals (see Appendix 2).

All statistical analysis was performed using Stata version 11 (StataCorp, College Station, TX). Standard errors were estimated using Taylor series linearization.

RESULTS

Prevalence Estimates and Modeled Age and Cohort Patterns

Figure 1A plots estimates of age-specific diabetes prevalence during the four observation periods under study. The underlying values and their standard errors are reported in Appendix Tables 3a and 3b. As reported elsewhere (3), there is a general upward trend in prevalence at each age.

Figure 1A shows a pattern in which the prevalence of diabetes declines at some set of ages above 60-64 in each of the four periods. Such a decline could be produced by higher mortality rates among those with diabetes than among those without. However, we show below that this pattern of decline with age is not present when prevalence rates are arrayed by birth cohort. In other words, the declines in prevalence with age in Figure 1A result from the increasing prevalence of diabetes among later-born cohorts.

Figure 1B presents estimates of diabetes prevalence among birth cohorts. It is clear that prevalence is rising from one birth cohort to the next, even at younger ages where prevalence is low. Furthermore, prevalence continues to rise even at the oldest ages, which is consistent with a

continued positive incidence of diabetes as cohorts age. Declining prevalence with age, a pattern suggested by period data, is not observed among real birth cohorts as they age.

The age-pattern of diabetes, as well as changes in diabetes prevalence from birth cohort to birth cohort, are summarized by our statistical model. Figure 2 plots the coefficients for each birth cohort in the age/cohort regression model. That the coefficients are monotonically increasing shows that more recent birth cohorts have higher diabetes prevalence than older cohorts. The increase is exceptionally rapid among cohorts born after 1950-59. The implication of the cohort coefficients is that the prevalence of diabetes at any age for the cohort born 1980-89 will be nearly triple that of the cohort born in 1950-59 and 4.9 times that of the cohort born in 1910-19 (derived from Appendix Table 4a).

Just as the age/cohort model produces rapidly increasing cohort effects, the age/period model produces rapidly rising period effects. This nearly straight-line increase in prevalence across periods is shown in Figure 3A (see Appendix Table 4b for actual values). By themselves, there is nothing in Figures 2 and 3A that would indicate which model is preferred. Both models produce R^2 values above 0.94. But when we add a cohort variable to the age/period model, the prevalence of obesity at age 25, the period effects nearly disappear, as shown in Figure 3A (Appendix Table 4c). They also become statistically insignificant.

Figure 3B compares the age-patterns of diabetes prevalence that are produced by the age/cohort model, the age/period model, and the age/period/cohort model. By far the most level age pattern is produced by the age/period model. As argued earlier, that age pattern is misleading because it fails to account for the rise in diabetes prevalence from one birth cohort to the next. As was suggested by a comparison of Figures 1A and 1B, the age pattern of diabetes prevalence in a birth cohort is steeper than that in a period. The age-pattern in the age/period model becomes

much steeper when birth-cohort obesity is introduced, as shown in Figure 3B. The age-pattern identified in the age/period/cohort model is very similar to that in the age/cohort model.

Incidence estimates

Based on the formula presented in Appendix 2, Figure 4 shows the age pattern of diabetes incidence that is implied by the age pattern of prevalence that we have uncovered. The values on the graph apply to the cohort born 1950-1959, but the shape of the curve is the same for all birth cohorts. The age-pattern of incidence rises to a peak in the age interval 55 to 64 (centered at age 60) and then declines slowly. At its peak from ages 55 to 64, for the cohort born 1950-1959, approximately 1.1% of the diabetes-free population will develop diabetes each year. Appendix 5 presents numerical details of our incidence estimates.

Sensitivity analysis

To examine the sensitivity of results to the choice of the HbA1c threshold, we adopt a threshold of HbA1c levels $\geq 6.0\%$. Recent guidelines from the American Diabetes Association consider individuals at this level to be at "very high risk" of incident diabetes (4). See Appendix 6 for a discussion of this choice of threshold. Using this lower threshold, we estimate the prevalence of being "at least at high risk" of diabetes over time and across birth cohorts, as shown in Appendix Figures 6a and 6b. 7,370 individuals in our sample met the more inclusive criterion. A comparison of Figure 1B to Appendix Figure 6b shows that the increase across birth cohorts in age-specific prevalence of "at least high-risk" is even more striking than that using the higher cut-off. In particular, the higher prevalence seen in more recent birth cohorts appears at earlier ages in "at least high-risk" than it does in diabetes itself.

We also estimate age/period, age/cohort, and age/period/cohort models of "at least high-risk" prevalence. The patterns described above are largely replicated using the lower cut-off.

Consistent with the higher level of prevalence, the rise in prevalence across ages and birth cohorts is greater when HbA1c ≥ 6.0% is used. However the introduction of obesity at age 25 into the age/period model has much the same effect as when HbA1c ≥ 6.5% is used; it steepens the age effects and reduces the period effects, though a significant period effect remains in the most recent period (see Appendix Figures 6c-6e and Figure 3B). Once again, this result places the spotlight on birth cohort influences in the rise of diabetes in the United States. Appendix Tables 6a-6c present numerical details of the results of our modeling of the prevalence of HbA1c ≥ 6.0%.

CONCLUSIONS

Birth cohorts are an attractive vehicle for investigating changes in the prevalence of diabetes because prevalence at any age is a cumulative product of influences in the past. These influences manifest themselves over the lifetime of birth cohorts, creating close associations in the prevalence of diabetes across age within a cohort.

We show that the prevalence of diabetes in the United States is rapidly increasing from one birth cohort to the next. We demonstrate this increase graphically and by means of an age/cohort model. The increase is especially rapid across cohorts born after 1950-59.

Our results also reveal that the pattern of increase with age in the prevalence of diabetes is considerably faster within a birth cohort than it is across ages in a particular period. The increase with age during any particular period is too mild, or even negative, because it does not account for the higher levels of diabetes evident among more recent birth cohorts.

An additional suggestion of the importance of birth cohort influences on diabetes prevalence is supplied by our age/period/cohort model. While an age/period model shows sharply increasing period effects, the addition of a term measuring birth cohort obesity at age 25

renders the period effects small and insignificant. This result indicates that birth cohort influences – in particular, birth cohort obesity levels – are important determinants of diabetes prevalence.

An innovation of our approach is that we convert estimates of birth cohort diabetes prevalence to estimates of incidence. Such estimates cannot be made using period data alone without the extreme assumption that no population rates are changing (25). This assumption is clearly not warranted in the case of diabetes, as shown in Figure 1. But such calculations of incidence can be made by comparing prevalence at different ages for the same birth cohort since any changes in prevalence within a birth cohort must be attributed to some combination of new diagnoses (incidence), differential mortality by diabetes status, and recovery (if any). To estimate incidence, we use the age effect coefficients from the age/cohort model, which is based on observations across eight birth cohorts. We demonstrate that the incidence of diabetes among diabetes-free persons rises steadily to a peak at ages 55 to 64 and then declines slowly.

To the best of our knowledge, these are the first estimates of the age-pattern of diabetes incidence that are based on measured data in a nationally-representative sample. Other estimates of age-patterns of diabetes incidence are few and inconsistent. Age patterns of diabetes incidence that peak and then decline are found in some populations (26–29). Other studies find that incidence continues to rise with age (30,31) or levels off at older ages (13,32). Annual estimates of incidence in the U.S. from the Centers for Disease Control and Prevention (CDC), which are based on retrospective self reports, show a peak in the age interval 45-64 in some years and at ages 65-79 in other years (33). Experimental evidence suggests a biological mechanism for increasing incidence with age at the individual level (34). One possible explanation for the peak and decline in diabetes incidence in a birth cohort is population heterogeneity in vulnerability to

diabetes, with the most vulnerable individuals being successively selected out of the diabetes-free population as birth cohorts age.

Our study has several limitations. We assume that migration does not affect the prevalence of diabetes in birth cohorts. When we removed foreign-born respondents from the sample, however, the pattern of our results for both prevalence and incidence was essentially unchanged (results available upon request). We also assume no age-cohort interactions. We tested this assumption by including interactions between a continuous variable for age and indicators for the three birth cohorts that provided the most prevalence estimates; coefficients on these interaction terms were not statistically significant ($p>.15$ in all cases).

The small sample sizes in NHANES required us to use ten-year wide birth cohorts and assume homogeneity within those birth cohorts. As a specification check, we divided the birth cohorts into different ten year intervals than reported in this paper (1915 to 1924, 1925 to 1934, etc.). Resulting patterns of prevalence were similar to the results presented here (results available upon request).

The NHANES data do not permit distinguishing between type 1 and type 2 diabetes. However, because type 2 accounts for about 90-95% of all diabetes cases (4), this was not a serious limitation.

We categorized as diabetic individuals below the 6.5% HbA1c threshold who reported taking medication for diabetes. On the other hand, we did not categorize as diabetic individuals below the 6.5% threshold with self-reported diabetes because we assume that the large majority of this group was assessed using alternative diagnostic criteria, such as FPG or Oral Glucose Tolerance Test. Prior research indicates that relative to these measures, the HbA1c test identifies

as diabetic a smaller group of high-risk individuals (16). For this reason, we did not assume that individuals with self-reported diabetes were ever above the HbA1c threshold for diabetes.

Finally, our method for estimating diabetes incidence assumes that mortality differences between people with and without diabetes have been constant and that remission rates are zero. The literature on the former is unresolved (17,35,36) and assuming zero remission is standard in projection models of diabetes prevalence (37,38). Appendix 2 provides more information on remission rates.

Two recent studies of individuals in NHANES found that secular changes in time-of-survey BMI explained some but not all of the secular increase in the prevalence of diabetes and prediabetes (3,12). Our findings also implicate the rise in obesity for increases in diabetes but we use aggregate data on birth cohorts and an historical rather than contemporary indicator of obesity. That both current and past levels of obesity affect an individual's risk of developing diabetes has been demonstrated in prior research (9). Thus, our results are consistent with other analyses that identify increases in the prevalence of obesity as an important factor in the rise in diabetes.

The prevalence of obesity has increased dramatically across recent US birth cohorts. We have shown that birth-cohort prevalence of diabetes is associated with birth-cohort levels of obesity at age 25. Because birth cohort effects persist as birth cohorts age, our results suggest that diabetes prevalence is likely to continue increasing despite an apparent plateauing of obesity in recent years (39). Additional analyses should investigate the implications of the birth cohort trends identified here for future diabetes prevalence in the United States.

Author contributions: EF and AS managed and analyzed the data. EF wrote the first draft of the manuscript. SP conceived of the analysis and oversaw the research. All authors edited the manuscript.
Ezra Fishman, Andrew Stokes, and Samuel H. Preston are the guarantors of this work and, as such, had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

ACKNOWLEDGEMENTS

The research has been supported by Grant R01AG040212 from the National Institute on Aging.

The authors acknowledge helpful comments from Irma Elo and Douglas Ewbank of the Population Studies Center, University of Pennsylvania, and Neil Mehta of Emory University Rollins School of Public Health. We thank the anonymous reviewers for this journal for their helpful comments.

Ezra Fishman, Andrew Stokes, and Samuel H. Preston take responsibility for the work as a whole.

No conflict of interest for any author.

References

1. Murphy SL, Xu J, Kochanek KD. *Deaths: Final Data for 2010*. Hyattsville, MD: National Center for Health Statistics; 2013.

2. Wong E, Backholer K, Gearon E, et al. Diabetes and risk of physical disability in adults: a systematic review and meta-analysis. *Lancet Diabetes Endocrinol.* 2013;1(2):106–114.

3. Cheng YJ, Imperatore G, Geiss LS, et al. Secular changes in the age-specific prevalence of diabetes among U.S. adults: 1988-2010. *Diabetes Care*. 2013;36(9):2690–6.

4. American Diabetes Association. Diagnosis and classification of diabetes mellitus. *Diabetes Care*. 2012;35 Suppl 1:S64–71.

5. Yeh H-C, Duncan BB, Schmidt MI, Wang N-Y, Brancati FL. Smoking, Smoking Cessation, and Risk for Type 2 Diabetes MellitusA Cohort Study. *Ann. Intern. Med.* 2010;152(1):10–17.

6. Luo J, Rossouw J, Tong E, et al. Smoking and diabetes: does the increased risk ever go away? *Am. J. Epidemiol.* 2013;178(6):937–45.

7. Narayan KM, Boyle JP, Thompson TJ, Gregg EW, Williamson DF. Effect of BMI on lifetime risk for diabetes in the U.S. *Diabetes Care*. 2007;30(6):1562–1566.

8. Schienkiewitz A, Schulze MB, Hoffmann K, Kroke A, Boeing H. Body mass index history and risk of type 2 diabetes: results from the European Prospective Investigation into Cancer and Nutrition (EPIC)–Potsdam Study. *Am. J. Clin. Nutr.* 2006;84(2):427–433.

9. Abdullah A, Stoelwinder J, Shortreed S, et al. The duration of obesity and the risk of type 2 diabetes. *Public Heal. Nutr*. 2011;14(1):119–126.

10. The NS, Richardson AS, Gordon-Larsen P. Timing and duration of obesity in relation to diabetes: findings from an ethnically diverse, nationally representative sample. *Diabetes Care*. 2013;36(4):865–872.

11. Reis JP, Hankinson AL, Loria CM, et al. Duration of abdominal obesity beginning in young adulthood and incident diabetes through middle age: the CARDIA study. *Diabetes Care*. 2013;36(5):1241–7.

12. Bullard KM, Saydah SH, Imperatore G, et al. Secular Changes in U.S. Prediabetes Prevalence Defined by Hemoglobin A1c and Fasting Plasma Glucose: National Health and Nutrition Examination Surveys, 1999-2010. *Diabetes Care*. 2013;36(8):2286–2293.

13. Geiss LS, Pan L, Cadwell B, Gregg EW, Benjamin SM, Engelgau MM. Changes in incidence of diabetes in US adults, 1997–2003. *Am J Prev Med*. 2006;30(5):371–377.

14. Narayan KMV, Boyle JP, Thompson TJ, Sorensen SW, Williamson DF. Lifetime risk for diabetes mellitus in the United States. *JAMA*. 2003;290(14):1884–90.

15. National Health and Nutrition Examination Survey. Survey questionnaires, datasets, and related documentation.http://www.cdc.gov/nchs/nhanes/nhanes_questionnaires.htm. Published 2012. Accessed May 1, 2013.

16. Cowie CC, Rust KF, Byrd-Holt DD, et al. Prevalence of diabetes and high risk for diabetes using A1C criteria in the U.S. population in 1988-2006. *Diabetes Care*. 2010;33(3):562–568.

17. Stokes A, Mehta NK. Mortality and excess risk in US adults with pre-diabetes and diabetes: a comparison of two nationally representative cohorts, 1988-2006. *Popul Heal. Metr*. 2013;11(1):7.

18. Gregg EW, Cadwell BL, Cheng YJ, et al. Trends in the prevalence and ratio of diagnosed to undiagnosed diabetes according to obesity levels in the US. *Diabetes Care*. 2004;27(12):2806–2812.

19. Bonora E, Tuomilehto J. The pros and cons of diagnosing diabetes with A1C. *Diabetes Care*. 2011;34 Suppl 2:S184–90.

20. Selvin E, Steffes MW, Zhu H, et al. Glycated hemoglobin, diabetes, and cardiovascular risk in nondiabetic adults. *N Engl J Med*. 2010;362(9):800–811.

21. Lee JM, Pilli S, Gebremariam A, et al. Getting heavier, younger: trajectories of obesity over the life course. *Int J Obes*. 2010;34(4):614–623.

22. National Health and Nutrition Examination Survey. *Analytic and reporting guidelines*. National Center for Health Statistics; 2006.

23. Mason WM, Fienberg SE. *Cohort Analysis in Social Research*. New York, NY: Springer-Verlag; 1985.

24. National Health Interview Survey (1986-2004). Linked Mortality Files, mortality follow-up through 2006: Matching Methodology, May 2009. 2009.

25. Greenland S, Rothman KJ. Measures of Occurrence. In: Rothman KJ, Greenland S, Lash TL, eds. *Modern Epidemiology*. 3rd ed. Philadelphia, PA: Lippincott Williams & Wilkins; 2008.

26. Pavkov ME, Hanson RL, Knowler WC, Bennett PH, Krakoff J, Nelson RG. Changing patterns of type 2 diabetes incidence among Pima Indians. *Diabetes Care*. 2007;30(7):1758–1763.

27. Khan NA, Wang H, Anand S, et al. Ethnicity and sex affect diabetes incidence and outcomes. *Diabetes Care*. 2011;34(1):96–101.

28. McDermott RA, Li M, Campbell SK. Incidence of type 2 diabetes in two indigenous Australian populations: a 6-year follow-up study. *Med. J. Aust.* 2010;192(10):4.

29. Berger B, Stenström G, Sundkvist G. Incidence, prevalence, and mortality of diabetes in a large population. A report from the Skaraborg Diabetes Registry. *Diabetes Care*. 1999;22(5):773–778.

30. Wilson PWF, Anderson KM, Kannel WB. Epidemiology of diabetes mellitus in the elderly: The Framingham study. *Am. J. Med.* 1986;80(5, Supplement 1):3–9.

31. Thunander M, Petersson C, Jonzon K, et al. Incidence of type 1 and type 2 diabetes in adults and children in Kronoberg, Sweden. *Diabetes Res Clin Pr.* 2008;82(2):247–255.

32. Monesi L, Baviera M, Marzona I, et al. Prevalence, incidence and mortality of diagnosed diabetes: evidence from an Italian population-based study. *Diabet. Med.* 2012;29(3):385–392.

33. Centers for Disease Control and Prevention. Crude and Age-Adjusted Incidence of Diagnosed Diabetes per 1,000 Population Aged 18-79 Years, United States, 1980-2007.http://www.cdc.gov/diabetes/statistics/incidence/fig2.htm. Published 2012. Accessed June 12, 2013.

34. Lee H-Y, Choi CS, Birkenfeld AL, et al. Targeted Expression of Catalase to Mitochondria Prevents Age-Associated Reductions in Mitochondrial Function and Insulin Resistance. *Cell Metab.* 2010;12(6):668–674.

35. Gregg EW, Gu Q, Cheng YJ, Narayan KMV, Cowie CC. Mortality trends in men and women with diabetes, 1971 to 2000. *Ann. Intern. Med.* 2007;147(3):149–155.

36. Gregg EW, Cheng YJ, Saydah S, et al. Trends in death rates among U.S. adults with and without diabetes between 1997 and 2006: findings from the National Health Interview Survey. *Diabetes Care*. 2012;35(6):1252–1257.

37. Boyle JP, Thompson TJ, Gregg EW, Barker LE, Williamson DF. Projection of the year 2050 burden of diabetes in the US adult population: dynamic modeling of incidence, mortality, and prediabetes prevalence. *Popul Heal. Metr*. 2010;8:29.

38. Narayan KM, Boyle JP, Geiss LS, Saaddine JB, Thompson TJ. Impact of recent increase in incidence on future diabetes burden: U.S., 2005-2050. *Diabetes Care*. 2006;29(9):2114–2116.

39. Flegal KM, Carroll MD, Kit BK, Ogden CL. Prevalence of obesity and trends in the distribution of body mass index among us adults, 1999-2010. *JAMA J. Am. Med. Assoc.* 2012;307(5):491–497.

Appendix 1: Cohort Obesity Prevalence

The prevalence of obesity at age 25 in successive cohorts, used in our age/period/cohort model as a continuous variable, is shown in the following table:

Percent Obese at Age 25 in Successive Cohorts

Birth Years	Percent
1920-1924	2.00%
1925-1929	2.58%
1930-1934	3.34%
1935-1939	3.46%
1940-1944	3.96%
1945-1949	4.85%
1950-1954	5.99%
1955-1959	6.99%
1960-1964	8.47%
1965-1969	11.32%
1970-1974	14.52%
1975-1979	15.92%

Sources: Calculated from NHANES continuous waves 1999-2008 using the interview sample.

One limitation of our study is that we used retrospective data on height and weight to estimate trends in cohort obesity at age 25 for subsequent US birth cohorts. Recall data may be subject to errors of misreporting. However, prior research using longitudinal data found a relatively high degree of correspondence between recall and contemporaneously reported data on BMI (A1). Validity of recall data over longer intervals of time has not been investigated. A second limitation is that we were not able to investigate cohort trends in obesity at younger ages because of lack of data on trends in childhood and adolescence.

Appendix 2: Derivation of Formula for Estimating Incidence of Diabetes from Prevalence of Diabetes in a Cohort

Suppose that 20% of a cohort has diabetes at age 30 and 25% of that cohort has diabetes at age 35. Then the incidence of diabetes (number of new cases per diabetes-free member of the population) between ages 30 and 35 is approximately .05/.80=.0625. That figure refers to incidence over a five-year period, whereas incidence is normally measured annually. An annualized rate would be .0625/5=.0125. This figure is based on the number who are free of diabetes at the beginning of the interval, whereas incidence is typically measured using a denominator measured at the middle of the interval. So the corrected figure is (.05/.775)/5=.0129.

This calculation makes three basic assumptions: (1) Migration does not affect birth cohort prevalence; (2) Those with diabetes at age 30 do not become diabetes-free by age 35, and (3) Those with diabetes at age 30 have the same probability of dying by age 35 as those who were diabetes-free at age 30. In constructing our estimates of the incidence of diabetes, we retain assumption 1 and 2, that migration does not affect prevalence and that those who enter the diabetic state leave it only by death (see discussion at end of this section). To check the sensitivity of our results to assumption 1, we excluded foreign-born individuals from our sample, and results were not substantially altered. However, assumption 3 is demonstrably untenable (A2). Accordingly, our estimates of diabetes incidence adjust for the higher mortality of those with diabetes.

We develop the estimation formula first by referring to the population at exact ages and then substituting equivalent formulas for the population at discrete age intervals. Under our assumptions, the diabetes-free population is subject to two sources of decrement, incident diabetes and death (A3).

(A2.1) $${}_5p_x^O = \exp[-5(\mu_x^O + \delta_x^O)]$$, where

${}_5p_x^O$ = probability of surviving in the disease-free state from age x to age x+5 for a person free of diabetes at age x

μ_x^O = death rate at age x for a person free of diabetes

δ_x^O = rate of acquiring diabetes (incidence rate) at age x for a person free of diabetes.

(A2.2) $${}_5p_x = \exp[-5\mu_x]$$, where

${}_5p_x$ = probability of surviving from age x to age x+5 for a randomly-chosen member of the population

μ_x = death rate at age x for a randomly-chosen member of the population.

Equations A2.1 and A2.2 assume that death rates and the incidence rate of diabetes are constant in the age interval x to x+5.

Call the non-diabetes population at age x N_x^O and the total population at age x N_x. Then the prevalence of non-diabetes at age x is

$$\Pi_x = \frac{N_x^O}{N_x}.$$

The prevalence of non-diabetes in the same cohort at age x+5 is

$$\Pi_{x+5} = \frac{N_{x+5}^O}{N_{x+5}} = \frac{N_x^O \exp[-5(\mu_x^O + \delta_x^O)]}{N_x \exp[-5\mu_x]}$$

(A2.3) $$= \Pi_x \exp[-5\delta_x^O]\exp[-5(\mu_x^O - \mu_x)].$$

Rewriting equation A2.3 gives

$$\exp[-5\delta_x^O] = \frac{\Pi_{x+5}}{\Pi_x}\frac{\exp[-5\mu_x]}{\exp[-5\delta_x^O]}, \text{ or}$$

(A2.4) $$\delta_x^O = -\frac{1}{5}\ln[\frac{\Pi_{x+5}}{\Pi_x}\frac{{}_5p_x}{{}_5^M p_x^O}], \text{ where}$$

${}_5^M p_x^O =$ probability of surviving the risk of death from x to x+5 for a diabetes-free person at age x.

Equation A2.4 shows that the incidence rate of diabetes between ages x and x+5 can be derived from the ratio of non-diabetes prevalence at x and x+5 and from differences in the survival probabilities between the entire population and the diabetes-free population over that age span. It also shows why a moving average of incidence estimates made using this equation is appropriate: errors in prevalence estimates at any particular age will appear in the numerator of one age-specific incidence estimate and in the denominator of the adjacent incidence estimate.

Substituting expressions for discrete five-year intervals into the equivalent terms in A2.4 gives

(A2.5) $${}_{10}\bar{\delta}_x^O = -\frac{1}{5}\ln[\frac{{}_5\Pi_{x+5}}{{}_5\Pi_x}\frac{{}_5L_{x+5}/{}_5L_x}{{}_5^M L_{x+5}^O/{}_5^M L_x^O}], \text{ where}$$

${}_{10}\bar{\delta}_x^O$ = rate of developing diabetes for a non-diabetic person in the age interval x to x+10,

${}_5\Pi_x =$ prevalence of non-diabetes at ages x to x+5

${}_5L_x =$ person-years lived between ages x and x+5 in a life table for the population

${}_5^M L_x^O =$ person-years lived between ages x and x+5 in a life table for persons free of diabetes.

We interpret ${}_{10}\bar{\delta}_x^O$ as pertaining to the age interval x+2.5 to x+7.5, i.e. the five-year age span at the middle of the ten-year age interval x to x+10. We use equation A2.5 for our incidence estimates in this study, assuming the incidence rate and differential mortality are constant within the five-year age intervals used. Values of ${}_5\Pi_x$ are calculated from fitted values in the age-cohort

model of prevalence. Values of ${}_{10}\bar{\delta}_x^O$ shown in Figure 4 come from fitted values of prevalence that use the coefficient for the cohort born in 1950-1959, but the shape of the graph in Figure 4 is robust to the use of other cohort coefficients. Values of ${}_5L_x$ and ${}_5L_x^{OM}$ come from the life tables as described in the Statistical Methods section.

Our calculations assume that there is no remission once the diabetes-defining threshold is reached. Remissions would offset new cases and produce an underestimate of the incidence rate. The principal source of remission of diabetes is bariatric surgery. According to the American Society for Metabolic and Bariatric Surgery (ASMBS), the number of procedures reached 103,000 in 2003 (A4). There were approximately 21,708,000 Americans aged 20+ with HbA1c values of 6.5% or greater in that year (A5,A6). Assuming that all those who had the surgery were diabetic, the annual rate of surgery among people with diabetes was .00497 in 2003. Two recent randomized clinical trials investigated the efficacy of bariatric surgery among those with diabetes. One found a one-year success rate in reducing HbA1c below 6.0% of 42% (A7) and the other a two-year rate of success of reducing HbA1c below 6.5% of 75% (A8). If we assume that the higher figure applies to the 5-year success rate required in our calculations, bariatric surgery would produce a remission rate of (.75)(.00497) = .00373 among people with diabetes in 2003. Since the ratio of people without diabetes to people with diabetes in that year was 9.12 (A6), the rate of flow into the non-diabetic population as a result of successful bariatric surgery was .00373/9.12 = .00041. This value compares to an incidence rate above age 50 of about .010 in our calculations. So the incidence rate above age 50 would be perhaps higher by the factor 1.04 if allowance were taken of remission from bariatric surgery. There are other sources of remission, of course, but in these two randomized clinical trials the remission rates for very intensive non-surgical medical treatment was only 12% (A7) and 0% (A8). Due to the intensive nature of the medical treatment, these findings can be considered an upper bound on remission rates in the diabetic population at large. It is worth noting that projections of future diabetes prevalence assume the cure rate for diabetes is zero (A9), and clinical guidelines imply that people who have been diagnosed with diabetes are considered diabetic even if their blood glucose is under control (A10).

Appendix 3: Prevalence Estimates and Confidence Intervals

Results displayed in Figures 1a and 1b are shown in more detail in **Appendix Tables 3a and 3b**.

Appendix Table 3a: Age-Specific Prevalence Across Observation Periods

Period:	1988-1994		1999-2002		2003-2006		2007-2010	
Age	Estimate	95% CI	Estimate	95% CI	Estimate	95% CI	Estimate	95% CI
20-24	0.004397	(0,0.009)	0.004269	(-0.001,0.009)	0.006232	(0.002,0.011)	0.005288	(-0.002,0.012)
25-29	0.004864	(0.001,0.009)	0.017177	(0.007,0.027)	0.022937	(0.01,0.035)	0.021857	(0.011,0.033)
30-34	0.008792	(0.004,0.014)	0.026586	(0.008,0.045)	0.024158	(0.013,0.035)	0.028596	(0.018,0.039)
35-39	0.032287	(0.015,0.05)	0.029327	(0.016,0.043)	0.037264	(0.022,0.052)	0.039685	(0.028,0.051)
40-44	0.045436	(0.03,0.06)	0.046697	(0.033,0.061)	0.047103	(0.03,0.064)	0.052208	(0.036,0.069)
45-49	0.050596	(0.036,0.066)	0.070413	(0.045,0.096)	0.06769	(0.048,0.087)	0.085597	(0.061,0.11)
50-54	0.091078	(0.063,0.119)	0.096351	(0.073,0.12)	0.117324	(0.095,0.14)	0.1389	(0.113,0.165)
55-59	0.114349	(0.091,0.138)	0.114759	(0.087,0.142)	0.152951	(0.112,0.194)	0.156262	(0.12,0.192)
60-64	0.153817	(0.129,0.178)	0.178177	(0.146,0.21)	0.165387	(0.138,0.193)	0.192598	(0.151,0.234)
65-69	0.140782	(0.109,0.172)	0.193894	(0.158,0.23)	0.200292	(0.164,0.237)	0.264501	(0.204,0.325)
70-74	0.142179	(0.112,0.173)	0.160267	(0.126,0.194)	0.211401	(0.175,0.247)	0.248475	(0.218,0.278)
75-79	0.186321	(0.148,0.224)	0.160083	(0.113,0.207)	0.183436	(0.145,0.222)	0.230738	(0.182,0.279)

Appendix Table 3b: Age-Specific Prevalence Across Birth Cohorts

Cohort:	1910-1919		1920-1929		1930-1939		1940-1949	
Age	Estimate	95% CI	Estimate	95% CI	Estimate	95% CI	Estimate	95% CI
20-24								
25-29								
30-34								
35-39								
40-44							0.0613	(0.031,0.091)
45-49							0.0506	(0.036,0.066)
50-54					0.0743	(0.048,0.1)	0.1043	(0.075,0.134)
55-59					0.1143	(0.091,0.138)	0.1460	(0.119,0.173)
60-64			0.1596	(0.128,0.191)	0.1626	(0.131,0.194)	0.1814	(0.158,0.205)
65-69			0.1408	(0.109,0.172)	0.2059	(0.178,0.234)	0.2430	(0.191,0.295)
70-74	0.1337	(0.098,0.17)	0.1506	(0.119,0.182)	0.2190	(0.198,0.24)	0.2620	(0.13,0.394)
75-79	0.1863	(0.148,0.224)	0.1724	(0.14,0.204)	0.2246	(0.179,0.27)		

Cohort:	1950-1959		1960-1969		1970-1979		1980-1989	
Age	Estimate	95% CI	Estimate	95% CI	Estimate	95% CI	Estimate	95% CI
20-24			0.0079	(0,0.016)	0.0022	(-0.001,0.006)	0.0058	(0.002,0.009)
25-29			0.0049	(0.001,0.009)	0.0211	(0.013,0.03)	0.0201	(0.009,0.031)
30-34	0.0051	(0.001,0.009)	0.0203	(0.01,0.031)	0.0257	(0.017,0.034)	0.0364	(-0.009,0.081)
35-39	0.0323	(0.015,0.05)	0.0324	(0.023,0.042)	0.0403	(0.029,0.052)		
40-44	0.0364	(0.023,0.05)	0.0507	(0.04,0.061)	0.0497	(0.007,0.092)		
45-49	0.0733	(0.056,0.09)	0.0770	(0.057,0.097)				
50-54	0.1229	(0.106,0.14)	0.1054	(0.045,0.166)				
55-59	0.1406	(0.109,0.172)						
60-64	0.1721	(0.105,0.24)						
65-69								
70-74								
75-79								

Appendix 4: Results of Models of Prevalence of Diabetes

Results displayed in Figures 2 and 3 are shown in more detail in Appendix Tables 4a, 4b, and 4c.

Appendix Table 4a: Results of Age-Cohort Model

Dependent Variable = Log Diabetes Prevalence

Indicator	Coefficient	SE	t-statistic	p-value	95% CI	
Age 25-29	1.080695	0.292925	3.689325	0.001283	0.473206	1.688185
Age 30-34	1.704125	0.314012	5.426944	1.89E-05	1.052904	2.355346
Age 35-39	2.532282	0.328767	7.702351	1.1E-07	1.85046	3.214103
Age 40-44	2.96014	0.342128	8.652146	1.57E-08	2.25061	3.669669
Age 45-49	3.384705	0.367831	9.201791	5.36E-09	2.62187	4.14754
Age 50-54	3.988511	0.395827	10.0764	1.05E-09	3.167616	4.809406
Age 55-59	4.232604	0.430616	9.829178	1.65E-09	3.33956	5.125647
Age 60-64	4.553761	0.435592	10.45419	5.35E-10	3.650399	5.457124
Age 65-69	4.748697	0.46782	10.15069	9.19E-10	3.778497	5.718897
Age 70-74	4.791991	0.494981	9.68117	2.17E-09	3.765464	5.818518
Age 75-79	4.92001	0.536914	9.163504	5.77E-09	3.806519	6.0335
1920-29 cohort	0.0491136	0.446574	0.109979	0.913423	-0.87702	0.975251
1930-39 cohort	0.3153463	0.443356	0.711272	0.48439	-0.60412	1.234809
1940-49 cohort	0.4939475	0.492615	1.002704	0.326905	-0.52767	1.515569
1950-59 cohort	0.5045705	0.527574	0.956399	0.349263	-0.58955	1.598691
1960-69 cohort	0.809644	0.554556	1.459987	0.158427	-0.34043	1.959722
1970-79 cohort	1.161477	0.583122	1.991825	0.058947	-0.04784	2.370798
1980-89 cohort	1.588968	0.630111	2.521726	0.019422	0.282197	2.895739
constant	-6.687672	0.60934	-10.9753	2.16E-10	-7.95137	-5.42398
R-Squared	0.9486					
N	41					

Age/Cohort Model: $ln(Y_{ia}) = \alpha + \beta_a X_a + \beta_i X_i$, where Y_{ia} = the proportion of the population in cohort i at age a with diabetes, X_a is a dummy variable indicating that the observation pertains to age a, and X_i is a dummy variable indicating that the observation pertains to cohort i.

Appendix Table 4b: Results of Age-Period Model

Dependent Variable: Log Diabetes Prevalence

Indicator	Coefficient	SE	t-statistic	p-value	95% CI	
Age 25-29	0.8340861	0.177858	4.689627	4.59E-05	0.472232	1.19594
Age 30-34	1.229262	0.176308	6.972242	5.69E-08	0.87056	1.587963
Age 35-39	1.935714	0.176051	10.99518	1.43E-12	1.577535	2.293892
Age 40-44	2.256934	0.173813	12.98482	1.61E-14	1.903309	2.61056
Age 45-49	2.564451	0.181592	14.12202	1.52E-15	2.194998	2.933903
Age 50-54	3.056156	0.183954	16.61367	1.36E-17	2.681898	3.430414
Age 55-59	3.259276	0.195733	16.65169	1.27E-17	2.861055	3.657497
Age 60-64	3.512932	0.177514	19.78963	7.07E-20	3.151777	3.874086
Age 65-69	3.625526	0.187088	19.37873	1.34E-19	3.244892	4.006159
Age 70-74	3.586429	0.189392	18.93653	2.7E-19	3.201108	3.97175
Age 75-79	3.620085	0.210792	17.17377	5.08E-18	3.191226	4.048944
2001 NHANES	0.2922673	0.108457	2.694766	0.010993	0.071609	0.512926
2005 NHANES	0.4265295	0.108519	3.930467	0.00041	0.205746	0.647313
2009 NHANES	0.5343021	0.100049	5.34043	6.76E-06	0.330752	0.737852
constant	-5.571567	0.133038	-41.8794	3.46E-30	-5.84224	-5.3009
R-Squared	0.9686					
N	48					

Age/Period model: $ln(Y_{ia}) = \alpha + \beta_a X_a + \beta_p X_p$, where Y_{ia} = the proportion of the population in cohort i at age a with diabetes, X_a is a dummy variable indicating that the observation pertains to age a, and X_p is a dummy variable indicating that the observation pertains to period p.

Appendix Table 4c: Results of Age-Period-Cohort Obesity Model

Dependent Variable: Log Diabetes Prevalence

Indicator	Coefficient	SE	t-statistic	p-value	95% CI	
Age 25-29	0.9926029	0.155681	6.375895	5.71E-07	0.6742	1.311005
Age 30-34	1.655445	0.18084	9.154213	4.71E-10	1.285587	2.025304
Age 35-39	2.672352	0.22659	11.79378	1.38E-12	2.208924	3.13578
Age 40-44	3.191595	0.25972	12.28859	5.08E-13	2.660407	3.722782
Age 45-49	3.636119	0.288422	12.60695	2.71E-13	3.046231	4.226008
Age 50-54	4.259973	0.314103	13.56235	4.38E-14	3.61756	4.902385
Age 55-59	4.545092	0.333427	13.63146	3.86E-14	3.863158	5.227026
Age 60-64	4.892977	0.345447	14.16418	1.46E-14	4.186459	5.599496
Age 65-69	5.07219	0.360942	14.05265	1.79E-14	4.333981	5.810399
Age 70-74	5.107939	0.392845	13.00242	1.26E-13	4.304481	5.911398
Age 75-79	5.116603	0.416248	12.2922	5.04E-13	4.26528	5.967926
2001 NHANES	-0.0517601	0.119307	-0.43384	0.667616	-0.29577	0.192251
2005 NHANES	0.0782829	0.119932	0.652728	0.519073	-0.16701	0.323571
2009 NHANES	0.0561693	0.142336	0.394624	0.696008	-0.23494	0.347279
obesity	13.10013	2.828275	4.631844	7.05E-05	7.315658	18.8846
constant	-7.107672	0.35859	-19.8212	2.09E-18	-7.84107	-6.37427
R-Squared	0.9813					
N	45					

Age/Period/Cohort model: $ln(Y_{ia}) = \alpha + \beta_a X_a + \beta_p X_p + \gamma Coh\_ob$, where Y_{ia}, X_a, and X_p are defined as in the Age/Period model and *Coh_ob* is a continuous variable representing the prevalence of obesity at age 25 in the cohort corresponding to the given age and period.

Appendix 5: Model Age Pattern of Diabetes Incidence

The model age pattern of diabetes incidence shown in Figure 4 is shown in detail in the following table:

Appendix Table 5: Estimates of Annual Incidence (New Cases Per Person-Year without Diabetes) for 1950-59 birth cohort

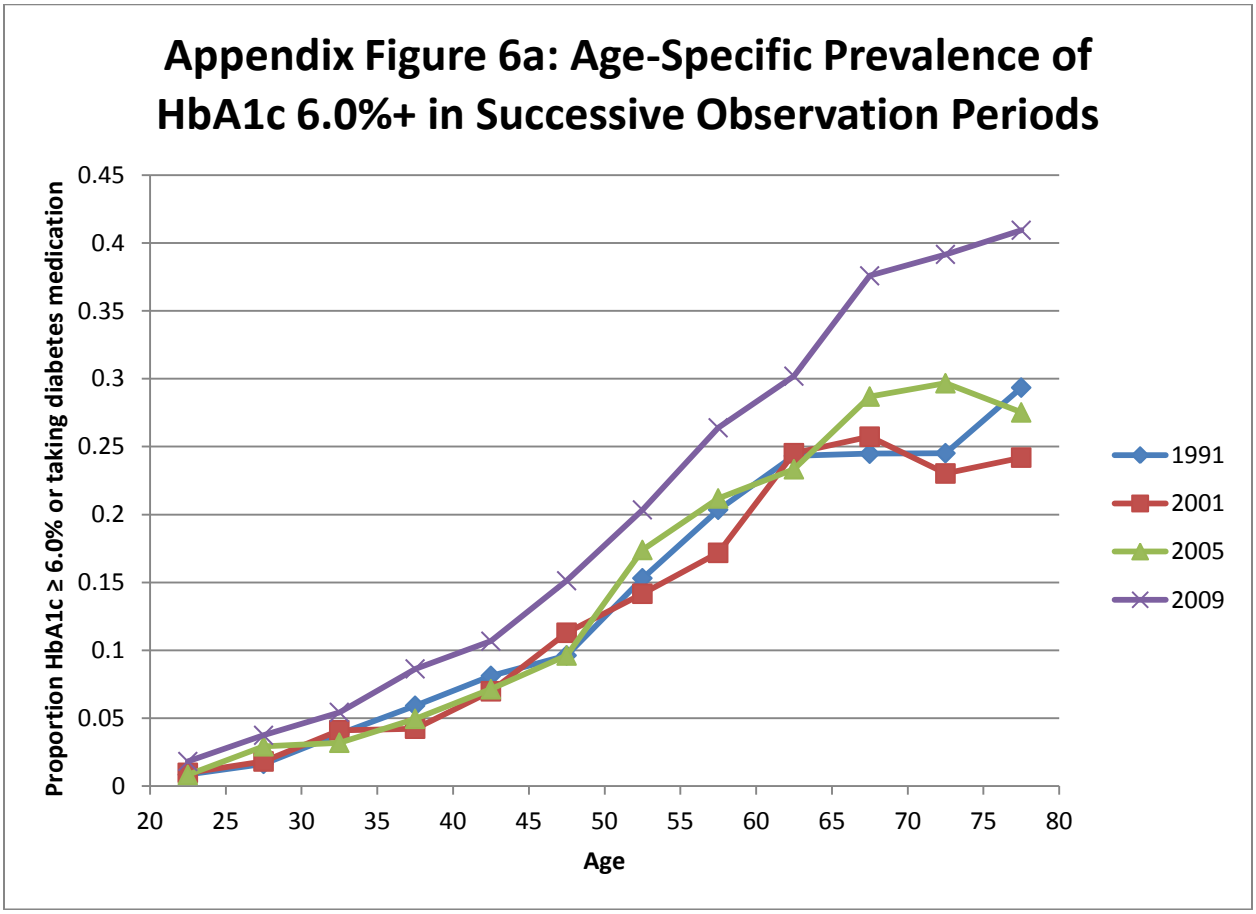
Age Interval	Incidence	Incidence with no differential mortality
20-24 to 25-29	0.000994	0.000934
25-29 to 30-34	0.001693	0.001617
30-34 to 35-39	0.002422	0.002303
35-39 to 40-44	0.003611	0.003429
40-44 to 45-49	0.006399	0.00612
45-49 to 50-54	0.007944	0.007516
50-54 to 55-59	0.011015	0.010359
55-59 to 60-64	0.011268	0.010269
60-64 to 65-69	0.010361	0.008846
65-69 to 70-74	0.009888	0.007609
70-74 to 75-79	0.008723	0.006022

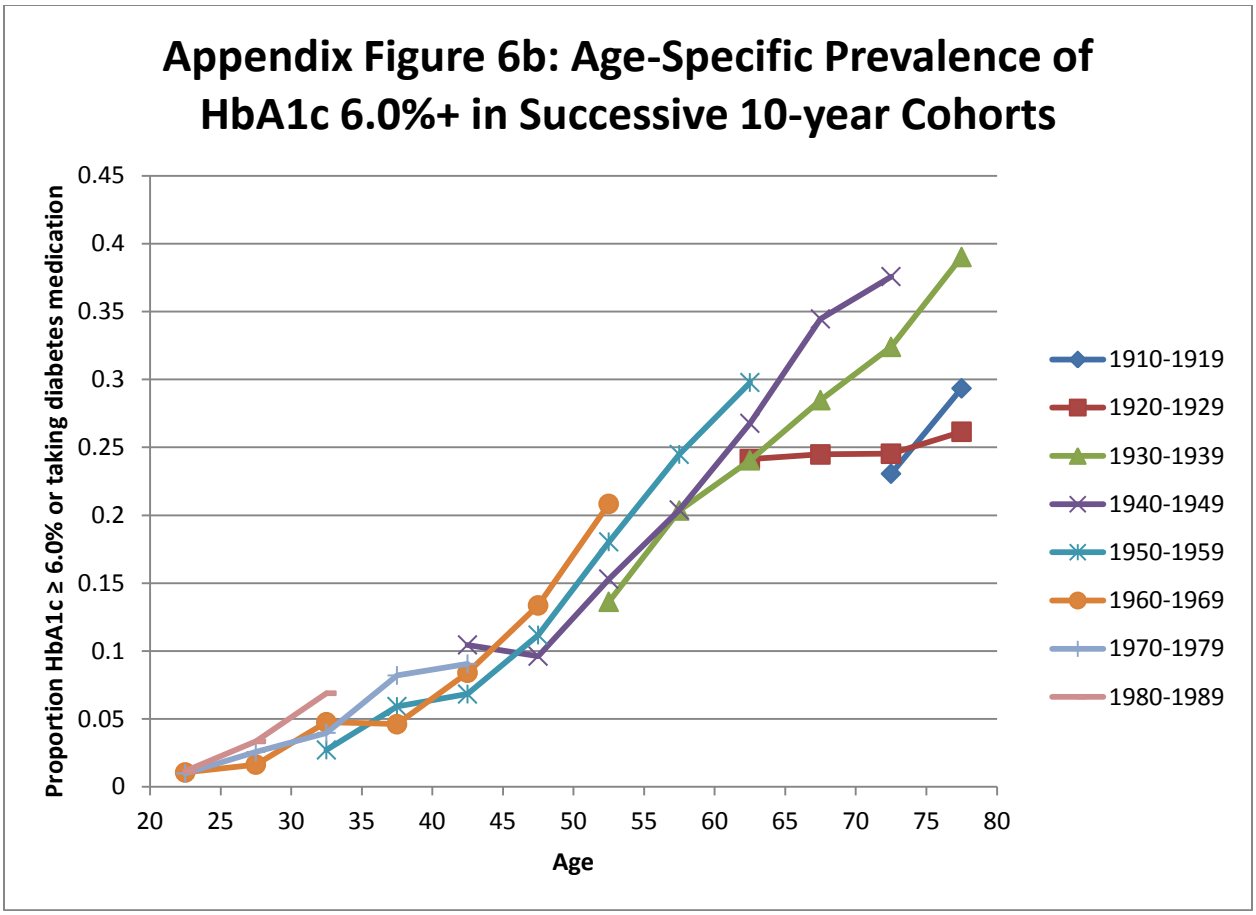
Incidence estimates are based on cohort prevalence estimates from age-cohort model and life-table values by diabetes status (nondiabetic versus entire population); see Methods section in text and Appendix 2 for details. Figure 4 plots the values in the "Incidence" column above. To demonstrate the effect of using mortality differences by diabetes status on the estimates, we present the estimates of incidence that would result if we had ignored mortality differences by diabetes status.

Appendix 6: Results based on "High Risk" of Diabetes, and Discussion of Threshold Choice

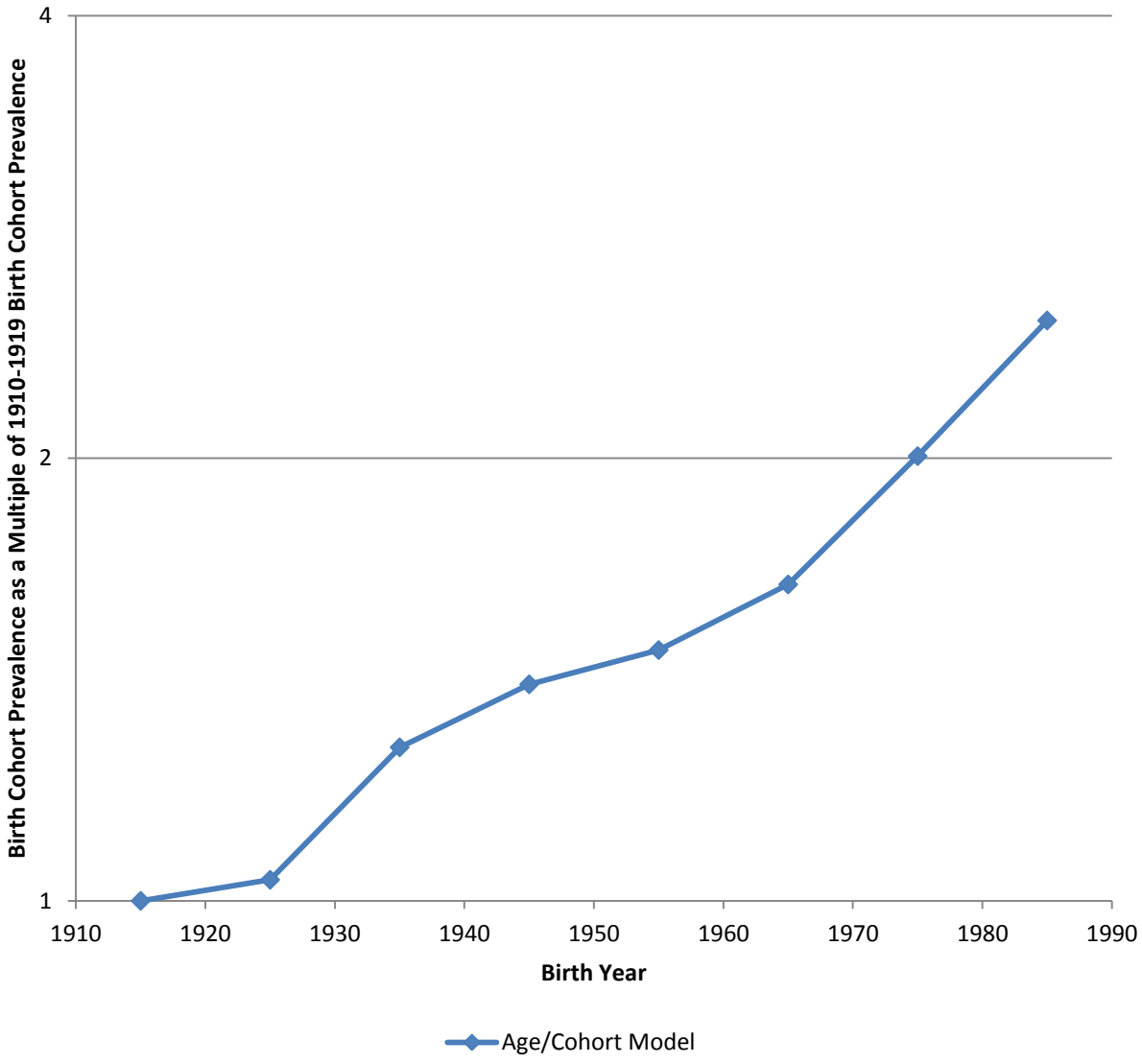
Appendix Figures 6a and 6b show estimates of the prevalence of "at least high risk" of diabetes, using HbA1c ≥ 6.0% (42 mmol/mol) (A11), by period and cohort. Appendix Figure 6c shows the cohort coefficients from the age/cohort model. Appendix Figure 6d shows the period effects in the age/period and age/period/cohort models discussed in the Statistical Methods section, as applied to the threshold HbA1c ≥ 6.0%. Appendix Figure 6e shows the modeled age-pattern of "at least high risk."

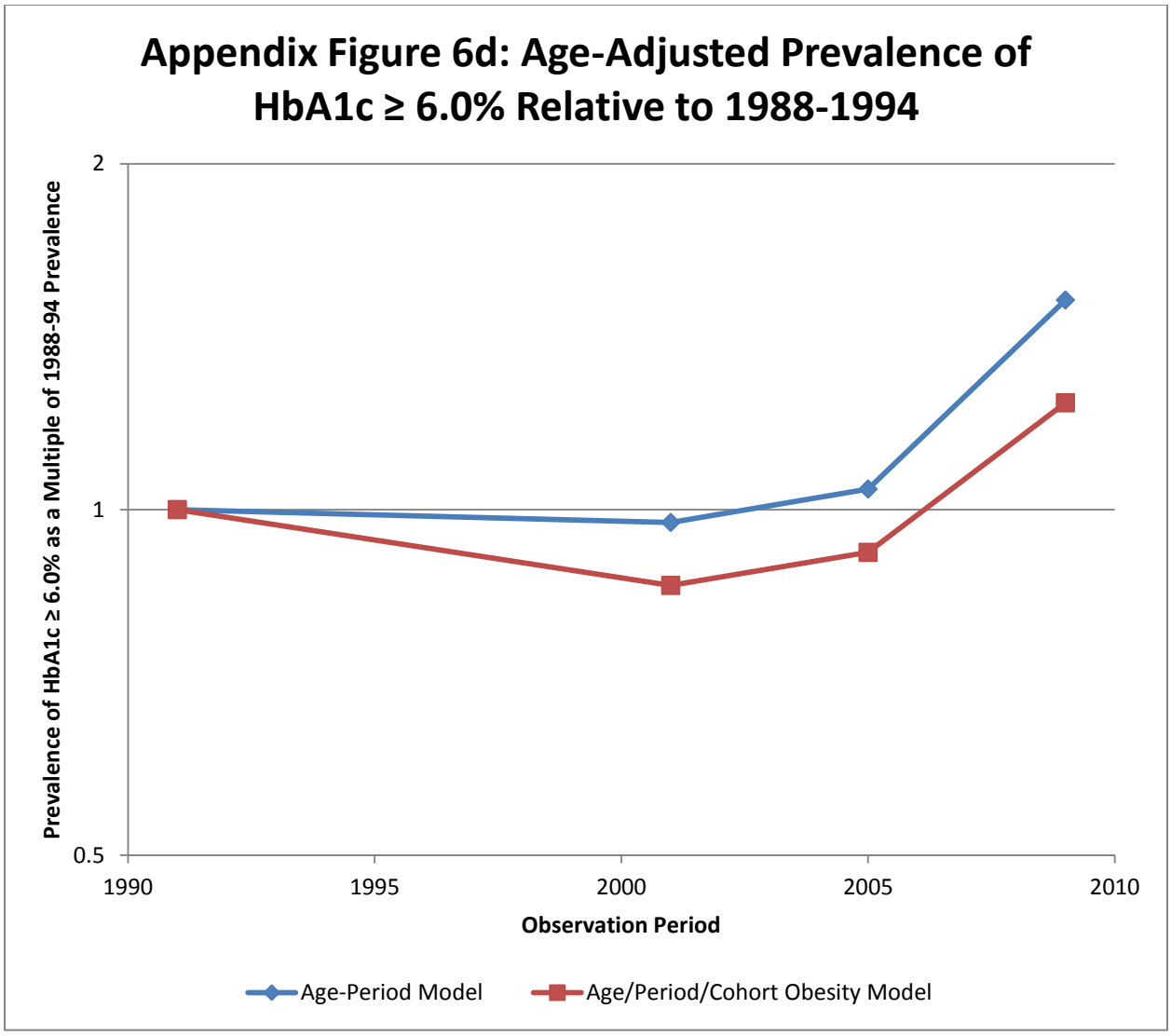
Although the recent ADA guidelines mention 6.0% as a possible threshold, they note that there is a "continuum of risk for diabetes with all glycemic measures" and did not formally identify 6.0% as a formal "high risk" threshold (A11). A recent meta-analysis indicated that there is no clear HbA1c-based threshold above which the risk of incident diabetes increases dramatically (A12). Nevertheless, using the 6.0% threshold is a useful way to test the sensitivity of our methods to the choice of threshold. The patterns we find using the 6.0% threshold are similar to the patterns we find using the 6.5% threshold.



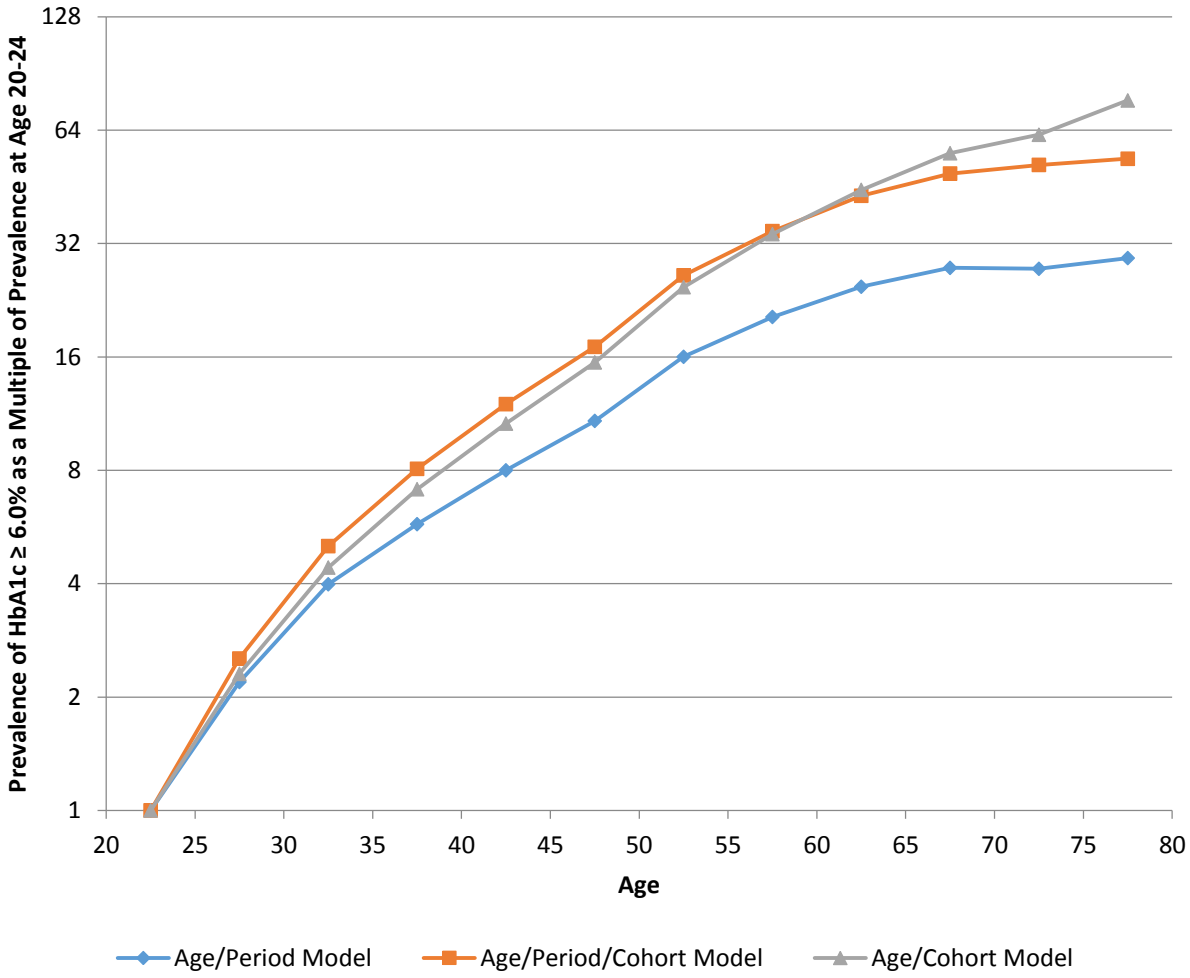

Appendix Figure 6b: Age-Specific Prevalence of HbA1c 6.0%+ in Successive 10-year Cohorts
Proportion HbA1c ≥ 6.0% or taking diabetes medication
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
20
25
30
35
40
45
50
55
60
65
70
75
80
Age
1910-1919
1920-1929
1930-1939
1940-1949
1950-1959
1960-1969
1970-1979
1980-1989

Appendix Figure 6c: Age-Adjusted Prevalence of HbA1c ≥ 6.0% in Birth Cohorts Relative to 1910-1919 Birth Cohort




Appendix Figure 6d: Age-Adjusted Prevalence of HbA1c ≥ 6.0% Relative to 1988-1994
Prevalence of HbA1c ≥ 6.0% as a Multiple of 1988-94 Prevalence
2
1
0.5
1990
1995
2000
2005
2010
Observation Period
Age-Period Model
Age/Period/Cohort Obesity Model

Appendix Figure 6e: Age-Specific Prevalence of HbA1c ≥ 6.0% Relative to Age 20-24 Prevalence



Appendix Table 6a: Results of Age-Cohort Model

Dependent variable = Log Prevalence of (HbA1c ≥ 6.0%)

Indicator	Coefficient	SE	t-statistic	p-value	95% CI	
Age 25-29	0.8303195	0.121278	6.846402	7.07E-07	0.578804	1.081835
Age 30-34	1.474147	0.130009	11.33883	1.17E-10	1.204525	1.743768
Age 35-39	1.944194	0.136118	14.28316	1.31E-12	1.661903	2.226485
Age 40-44	2.33351	0.141649	16.47384	7.35E-14	2.039747	2.627273
Age 45-49	2.691181	0.152291	17.67128	1.74E-14	2.375348	3.007013
Age 50-54	3.129008	0.163882	19.09303	3.51E-15	2.789137	3.468879
Age 55-59	3.4235	0.178286	19.20231	3.11E-15	3.053757	3.793242
Age 60-64	3.654953	0.180346	20.26636	1.01E-15	3.280939	4.028967
Age 65-69	3.828174	0.193689	19.76451	1.71E-15	3.426487	4.229861
Age 70-74	3.876574	0.204934	18.91618	4.26E-15	3.451566	4.301582
Age 75-79	4.00686	0.222296	18.02492	1.16E-14	3.545847	4.467873
1920-29 cohort	0.0330163	0.184893	0.17857	0.85991	-0.35043	0.41646
1930-39 cohort	0.240484	0.18356	1.31011	0.203674	-0.1402	0.621165
1940-49 cohort	0.3391446	0.203955	1.66284	0.110527	-0.08383	0.762121
1950-59 cohort	0.3927891	0.218429	1.79825	0.085875	-0.0602	0.845782
1960-69 cohort	0.4955284	0.2296	2.158226	0.042089	0.019368	0.971689
1970-79 cohort	0.6965528	0.241427	2.885149	0.008593	0.195864	1.197242
1980-89 cohort	0.9089721	0.260882	3.48423	0.002102	0.367937	1.450008
constant	-5.280235	0.252282	-20.9299	5.14E-16	-5.80344	-4.75703
R-Squared	0.9878688					
N	41					

Appendix Table 6b: Results of Age-Period Model

Dependent variable = Log Prevalence of (HbA1c ≥ 6.0%)

Indicator	Coefficient	SE	t-statistic	p-value		
Age 25-29	0.785042	0.089818	8.740355	4.21E-10	0.602306	0.967778
Age 30-34	1.382502	0.089036	15.52754	9.9E-17	1.201358	1.563646
Age 35-39	1.750047	0.088906	19.68429	8.32E-20	1.569167	1.930927
Age 40-44	2.079389	0.087776	23.6898	2.73E-22	1.900808	2.25797
Age 45-49	2.381469	0.091704	25.96907	1.53E-23	2.194896	2.568043
Age 50-54	2.77329	0.092897	29.85342	1.85E-25	2.58429	2.962291
Age 55-59	3.016927	0.098845	30.52183	9.13E-26	2.815825	3.218028
Age 60-64	3.201748	0.089644	35.71608	5.92E-28	3.019365	3.384131
Age 65-69	3.317046	0.094479	35.10868	1.03E-27	3.124826	3.509265
Age 70-74	3.311332	0.095643	34.6218	1.61E-27	3.116745	3.505919
Age 75-79	3.376998	0.10645	31.7239	2.66E-26	3.160425	3.593572
2001 NHANES	-0.0258708	0.054771	-0.47235	0.639789	-0.1373	0.085562
2005 NHANES	0.0411817	0.054802	0.751463	0.4577	-0.07031	0.152677
2009 NHANES	0.4200199	0.050525	8.313196	1.33E-09	0.317227	0.522813
constant	-4.683064	0.067184	-69.7048	2.09E-37	-4.81975	-4.54638
R-Squared	0.9897704					
N	48					

Appendix Table 6c: Results of Age-Period-Cohort Obesity Model

Dependent variable = Log Prevalence of (HbA1c ≥ 6.0%)

Indicator	Coefficient	SE	t-statistic	p-value	95% CI	
Age 25-29	0.9286565	0.087923	10.5622	1.88E-11	0.748835	1.108479
Age 30-34	1.61597	0.102132	15.82242	8.41E-16	1.407087	1.824853
Age 35-39	2.088447	0.12797	16.31986	3.74E-16	1.82672	2.350175
Age 40-44	2.484058	0.146681	16.93516	1.41E-16	2.184062	2.784053
Age 45-49	2.834346	0.16289	17.40037	6.89E-17	2.501198	3.167493
Age 50-54	3.271163	0.177394	18.44012	1.46E-17	2.908352	3.633973
Age 55-59	3.541548	0.188307	18.8073	8.63E-18	3.156417	3.92668
Age 60-64	3.757913	0.195096	19.26188	4.54E-18	3.358897	4.156929
Age 65-69	3.893697	0.203847	19.1011	5.69E-18	3.476784	4.31061
Age 70-74	3.946568	0.221865	17.78819	3.83E-17	3.492804	4.400332
Age 75-79	3.984019	0.235082	16.94739	1.39E-16	3.503223	4.464814
2001 NHANES	-0.1520016	0.06738	-2.25588	0.031793	-0.28981	-0.01419
2005 NHANES	-0.0857865	0.067733	-1.26654	0.2154	-0.22432	0.052743
2009 NHANES	0.2139795	0.080386	2.661893	0.012539	0.049571	0.378388
obesity	4.388277	1.597306	2.747298	0.010222	1.121419	7.655136
constant	-5.274105	0.202519	-26.0426	1.15E-21	-5.6883	-4.85991
R-Squared	0.9925176					
N	45					

References

A1. Perry GS, Byers TE, Mokdad AH, Serdula MK, Williamson DF. The Validity of Self-Reports of Past Body Weights by U.S. Adults. *Epidemiology*. 1995;6(1):61–66.

A2. Stokes A, Mehta NK. Mortality and excess risk in US adults with pre-diabetes and diabetes: a comparison of two nationally representative cohorts, 1988-2006. *Popul Heal. Metr*. 2013;11(1):7.

A3. Preston SH, Guillot M, Heuveline P. *Demography*. Oxford: Blackwell; 2001.

A4. National Institute of Diabetes and Digestive and Kidney Diseases. Longitudinal Assessment of Bariatric Surgery (LABS).http://win.niddk.nih.gov/publications/labs.htm. Published 2011.

A5. U.S. Census Bureau. Statistical Abstract of the United States, 2004-2005.http://www.census.gov/prod/2004pubs/04statab/pop.pdf. Published 2005.

A6. Cheng YJ, Imperatore G, Geiss LS, et al. Secular changes in the age-specific prevalence of diabetes among U.S. adults: 1988-2010. *Diabetes Care*. 2013;36(9):2690–6.

A7. Schauer PR, Kashyap SR, Wolski K, et al. Bariatric surgery versus intensive medical therapy in obese patients with diabetes. *N Engl J Med*. 2012;366(17):1567–1576.

A8. Mingrone G, Panunzi S, De Gaetano A, et al. Bariatric surgery versus conventional medical therapy for type 2 diabetes. *N Engl J Med*. 2012;366(17):1577–1585.

A9. Boyle JP, Thompson TJ, Gregg EW, Barker LE, Williamson DF. Projection of the year 2050 burden of diabetes in the US adult population: dynamic modeling of incidence, mortality, and prediabetes prevalence. *Popul Heal. Metr*. 2010;8:29.

A10. Ali MK, Bullard KM, Saaddine JB, Cowie CC, Imperatore G, Gregg EW. Achievement of goals in U.S. diabetes care, 1999-2010. *N Engl J Med*. 2013;368(17):1613–1624.

A11. American Diabetes Association. Diagnosis and classification of diabetes mellitus. *Diabetes Care*. 2012;35 Suppl 1:S64–71.

A12. Gregg EW, Geiss L, Zhang P, Zhuo X, Williamson DF, Albright AL. Implications of risk stratification for diabetes prevention: the case of hemoglobin A1c. *Am J Prev Med*. 2013;44(4 Suppl 4):S375–80.