Diagnosed and Undiagnosed Diabetes Prevalence by County in the U.S., 1999–2012

Laura Dwyer-Lindgren,[1]
Johan P. Mackenbach,[2]
Frank J. van Lenthe,[2]
Abraham D. Flaxman,[1] and
Ali H. Mokdad[1]

OBJECTIVE

Previous analyses of diabetes prevalence in the U.S. have considered either only large geographic regions or only individuals in whom diabetes had been diagnosed. We estimated county-level trends in the prevalence of diagnosed, undiagnosed, and total diabetes as well as rates of diagnosis and effective treatment from 1999 to 2012.

RESEARCH DESIGN AND METHODS

We used a two-stage modeling procedure. In the first stage, self-reported and biomarker data from the National Health and Nutrition Examination Survey (NHANES) were used to build models for predicting true diabetes status, which were applied to impute true diabetes status for respondents in the Behavioral Risk Factor Surveillance System (BRFSS). In the second stage, small area models were fit to imputed BRFSS data to derive county-level estimates of diagnosed, undiagnosed, and total diabetes prevalence, as well as rates of diabetes diagnosis and effective treatment.

RESULTS

In 2012, total diabetes prevalence ranged from 8.8% to 26.4% among counties, whereas the proportion of the total number of cases that had been diagnosed ranged from 59.1% to 79.8%, and the proportion of successfully treated individuals ranged from 19.4% to 31.0%. Total diabetes prevalence increased in all counties between 1999 and 2012; however, the rate of increase varied widely. Over the same period, rates of diagnosis increased in all counties, while rates of effective treatment stagnated.

CONCLUSIONS

Our findings demonstrate substantial disparities in diabetes prevalence, rates of diagnosis, and rates of effective treatment within the U.S. These findings should be used to target high-burden areas and select the right mix of public health strategies.

[1]Institute for Health Metrics and Evaluation, University of Washington, Seattle, WA
[2]Department of Public Health, Erasmus MC, University Medical Centre Rotterdam, Rotterdam, the Netherlands

Corresponding author: Laura Dwyer-Lindgren, ladwyer@uw.edu.

Diabetes mellitus is a leading cause of death and poor health in the U.S. In 2013, diabetes was responsible for 74.9 thousand deaths (the seventh leading cause of death) and 1.85 million years lived with disability (the eighth leading cause of disability) (1,2). Diabetes also exerts a large and rapidly increasing burden on the U.S. economy, with total costs in 2012 estimated at $245 billion (3).

In addition to medical strategies for identifying and managing diabetes, there are a number of evidence-based public health strategies aimed at primary prevention, screening, and improved disease management (4,5). Effectively and efficiently

deploying these strategies, especially given financial constraints and competing priorities, requires detailed local information about diabetes burden. This information can be used to define the scope of the problem as well as to identify high-need areas. In particular, information about both diagnosed and undiagnosed cases is essential in order to fully appreciate the population that is in need of services. Similarly, local information about rates of diagnosis and effective treatment are important inputs for determining the right mix of strategies to address the diabetes burden of a particular community.

National trends in diabetes prevalence are typically based on the National Health and Nutrition Examination Survey (NHANES) (6). The NHANES comprises both an interview and a laboratory component, which includes collecting biomarkers for diabetes. This allows researchers to use NHANES data to describe trends in diagnosed and undiagnosed diabetes, as well as rates of diagnosis and effective treatment, but only at the national level. State and local trends (7–9), in contrast, are typically derived from the Behavioral Risk Factor Surveillance System (BRFSS) (10), which has a much larger sample size and more comprehensive geographic coverage than the NHANES. The BRFSS does not include any biomarkers, however, and can only be used to track diagnosed diabetes prevalence.

Most local health departments are organized by county or groups of counties (11); however; only trends in diagnosed diabetes are available at this level (7–9). We combined NHANES and BRFSS data in order to estimate county-level prevalence of both diagnosed and undiagnosed diabetes in adults ≥20 years of age for each year from 1999 to 2012. We also calculated several derived measures, including the proportion of diabetes case patients who have received a diagnosis and the proportion of case patients who have been effectively treated.

RESEARCH DESIGN AND METHODS

Overview

For this analysis we used a two-stage approach to estimate five measures of diabetes prevalence (Table 1). In the first stage, we used NHANES data to fit a model for predicting high fasting plasma glucose (FPG) levels (≥126 mg/dL) and/or A1C levels (≥6.5% [48 mmol/mol]) (12) on the basis of self-reported demographic and behavioral characteristics. We then applied this model to BRFSS data to impute high FPG and/or A1C status for each BRFSS respondent. In the second stage, we used the imputed BRFSS data to fit a series of small area models, which were used to predict the county-level prevalence of each of the five diabetes-related outcomes.

Data

This analysis used NHANES and BRFSS data from 1999 to 2012. Over this period, the NHANES subsample that contains FPG measurement included 17,375 respondents ≥20 years of age; 15,600 of these respondents (89.8%) had no missing values for any of the relevant variables and were incorporated into this analysis. Over the same period, the BRFSS included 4,620,693 respondents ≥20 years of age; of these, 4,107,972 respondents (88.9%) had no missing values for any relevant variable and were included in this analysis. Several additional data sources were used, either as covariates in the small area models or for poststratification of estimates, as described below. Further details on all data sources are provided in the Supplementary Data.

High FPG/A1C Models

Following Danaei et al. (13) and Olives et al. (14), we developed respondent-level logistic regression models for predicting high FPG and/or A1C status (referred to hereafter as "high FPG/A1C"). Using NHANES data, the following model was fit separately for males and females, and for individuals who had previously received a diagnosis and had not received a diagnosis:

$$Y_i \sim \text{Bernoulli}(p_i)$$

$$\begin{aligned}
\text{logit}(p_i) = {} & \beta_0 + \beta_{1,a_i} + \beta_{2,r_i} + \beta_{3,e_i} \\
& + \beta_{4,m_i} + \beta_5 \cdot BMI_i \\
& + \beta_6 \cdot BMI_i^2 + \beta_7 \cdot H_i + \beta_8 \cdot S_i
\end{aligned}$$
$$\text{(Eq. 1)}$$

where Y_i is 1 if individual i has high FPG/A1C and 0 otherwise; a_i, r_i, e_i, m_i, BMI_i, and BMI_i^2 are individual i's age group (20–29, 30–39, 40–49, 50–59, 60–69, 70+ years), race/ethnicity (white, black, Hispanic, other), education status (less than high school, high school graduate, some college, college graduate), marital status (currently married, formerly married, never married), BMI, and squared BMI, respectively; and H_i and S_i are indicators for whether or not individual i has health insurance and is a current smoker, respectively.

The fitted logistic regression models were used to impute current (at the time of survey) high FPG/A1C status for each BRFSS respondent. Ten separate imputed data sets were created using simulation methods (15) to reflect the uncertainty

Table 1—Outcome measure definitions	
Measure	Definition
Diagnosed diabetes prevalence	The proportion of adults ≥20 years of age who report a previous diabetes diagnosis
Undiagnosed diabetes prevalence	The proportion of adults ≥20 years of age who do not report a previous diabetes diagnosis and who have high FPG/A1C*
Total diabetes prevalence	The proportion of adults ≥20 years of age who report a previous diabetes diagnosis and/or have high FPG/A1C*; total diabetes prevalence is equal to the sum of diagnosed and undiagnosed diabetes prevalence
Diabetes awareness	The proportion of adults ≥20 years of age with a previous diabetes diagnosis and/or high FPG/A1C* who have received a diagnosis; diabetes awareness is equal to the ratio of diagnosed to total diabetes prevalence
Diabetes control	The proportion of adults ≥20 years of age with a previous diabetes diagnosis and/or high FPG/A1C* who currently do not have high FPG/A1C*

*FPG ≥126 mg/dL and/or A1C ≥6.5% (48 mmol/mol).

in each BRFSS respondent's true high FPG/A1C status.

The predictive accuracy of this model was assessed using cross-validation, as described in the Supplementary Data. The model was found to have high concordance overall—it correctly predicted high FPG/A1C status for ~9 of 10 respondents—however, the sensitivity (i.e., the proportion of true case patients identified) was relatively low (11.2–13.2%, depending on sex and previous diagnosis).

Small Area Models

Small area models were developed to estimate county-level diagnosed diabetes prevalence, undiagnosed diabetes prevalence, and uncontrolled (diagnosed and with high FPG/A1C) diabetes prevalence based on imputed BRFSS data. These models are designed to borrow strength across space and time, and from external information in the form of covariates in order to generate more precise estimates than those calculated directly from the small samples available in most counties.

Each of these models was specified as follows:

$$Y_{j,t,a,r,m,e} \sim \text{Binomial}\left(p_{j,t,a,r,m,e}, \ N_{j,t,a,r,m,e}\right)$$

$$\begin{aligned}\text{logit}\left(p_{j,t,a,r,m,e}\right) = {}& \beta_0 + \beta_{1,a} + \beta_{2,r} + \beta_{3,m} \\ & + \beta_{4,e} + \boldsymbol{\beta}_5 \cdot \boldsymbol{X}_{j,t} + u_j \\ & + w_t + d_{j,t}\end{aligned}$$

(Eq. 2)

where $N_{j,t,a,r,m,e}$, $Y_{j,t,a,r,m,e}$, and $p_{j,t,a,r,m,e}$ are the number of individuals sampled, the number of case patients among those sampled, and the true prevalence, respectively, in county j, year t, age group a, race/ethnicity group r, marital status group m, and education group e; β_0 is the global intercept; $\beta_{1,a}$ values are age group effects (20–29, 30–39, 40–49, 50–59, 60–69, and 70+ years); $\beta_{2,r}$ values are race/ethnicity effects (Hispanic, white non-Hispanic, black non-Hispanic, native non-Hispanic, and other non-Hispanic); $\beta_{3,m}$ values are marital status effects (currently married, formerly married, and never married); $\beta_{4,e}$ values are education effects (less than high school, high school graduate, some college, and college graduate); and $\boldsymbol{\beta}_5$ is a vector of effects for three county-year-level covariates ($\boldsymbol{X}_{j,t}$) (percentage of individuals

living in poverty, percentage of rural households, and the number of doctors per capita). u_j and w_t are county- and year-level random effects, respectively, both of which are assumed to follow a conditional autoregressive distribution (16). $d_{j,t}$ is a county-year-level random effect that is also assumed to follow a conditional autoregressive distribution (17). Separate models were fit for males

and females, and the procedure described by Dwyer-Lindgren et al. (18) was used to correct for noncoverage bias in BRFSS data prior to 2011 when a cell phone sample was introduced (19).

Models were fit using the Template Model Builder package (20) in R version 3.2.4 (21). Simulation methods (15) were used generate 1,000 draws of diagnosed, undiagnosed, and uncontrolled
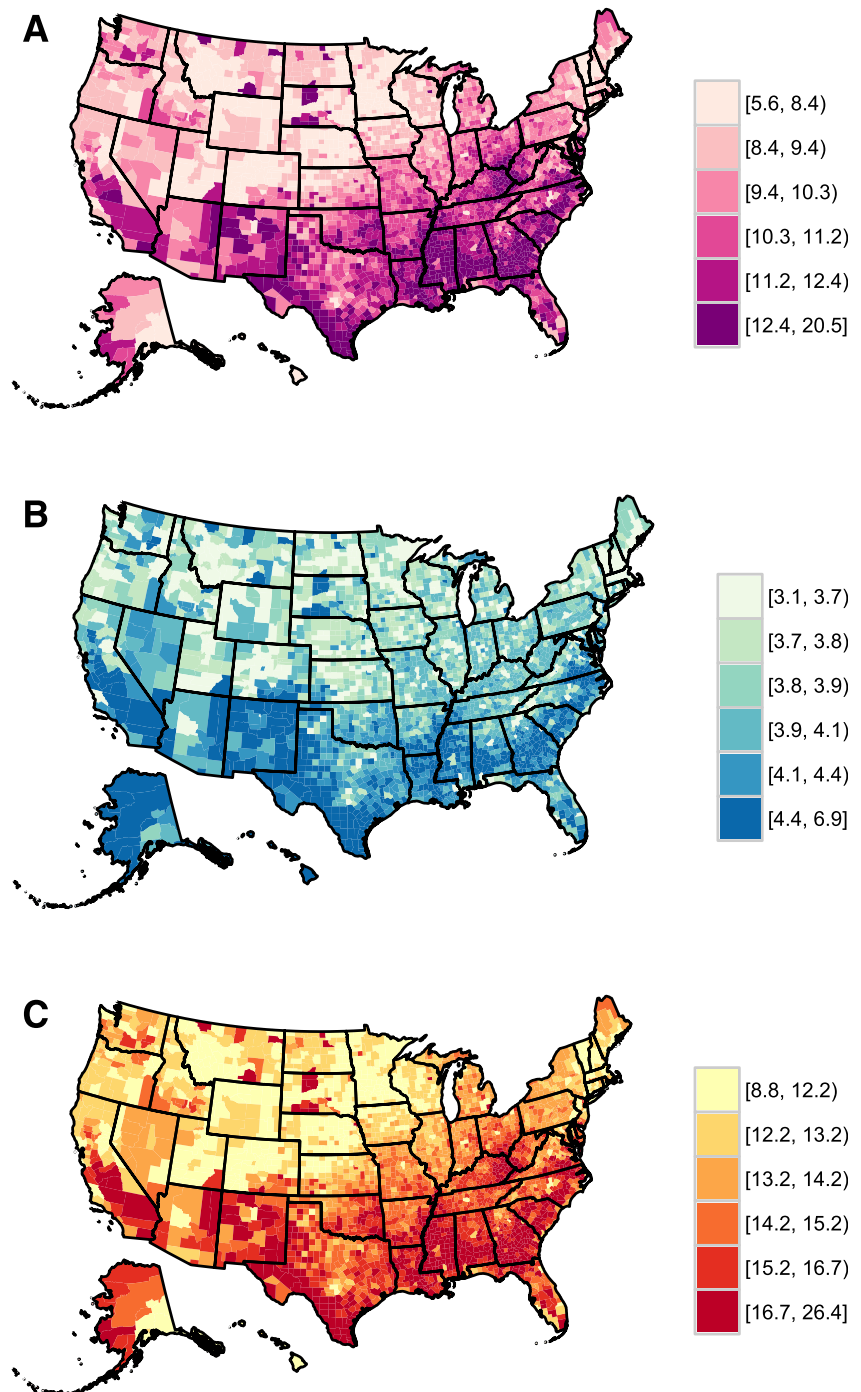
Figure 1—Age-standardized diabetes prevalence by county, 2012. *A*: Diagnosed diabetes prevalence. *B*: Undiagnosed diabetes prevalence. *C*: Total diabetes prevalence.

diabetes prevalence from the fitted small area models. These draws were poststratified by race, marital status, and education and then age standardized. Point estimates were calculated from the mean of these 1,000 draws, whereas 95% uncertainty intervals were calculated from the 2.5th and 97.5th percentiles. Estimates of total diabetes prevalence, diabetes awareness, and diabetes control were derived from the directly modeled quantities as follows: total = diagnosed + undiagnosed; awareness = diagnosed/total; control = 1 − uncontrolled/total. State- and national-level estimates of all quantities were derived by population weighting of county-level estimates.

Finally, in order to account for the uncertainty arising from using imputed data in the models for undiagnosed and uncontrolled diabetes, the entire procedure described above was repeated for each of 10 imputed data sets. Estimates were combined across data sets, and uncertainty intervals were recalculated to take into account the variation between the imputed data sets as well as the uncertainty from the small area models (22).

The predictive accuracy of this small area model was assessed with reference to diagnosed diabetes using empirical validation methods, as described in the Supplementary Data. In general, model predictions were found to have lower error and bias for counties with larger sample sizes. However, even for counties where only a single individual was sampled each year, the mean error (a measure of bias) was −0.3 percentage points, while the mean absolute error (a measure of precision) was 1 percentage point.

## RESULTS

Diabetes Prevalence in 2012
Age-standardized diagnosed diabetes prevalence for the U.S. as a whole was 10.2% (95% uncertainty interval 10.1%, 10.4%) in 2012, whereas undiagnosed diabetes prevalence was 4.1% (3.6%, 4.5%), resulting in a total diabetes prevalence of 14.3% (13.8%, 14.7%). Among counties, diagnosed diabetes prevalence ranged from 5.6% to 20.4%, undiagnosed diabetes prevalence ranged from 3.2% to 6.8%, and total diabetes prevalence ranged from 8.8% to 26.4%. Figure 1 shows age-standardized diagnosed, undiagnosed, and total diabetes prevalence by

county in 2012. Diagnosed diabetes prevalence was highest among counties in the deep South (excluding Florida), near the Texas-Mexico border, and in counties with Native American reservations in the four corners region of the Southwest and in North and South Dakota. In contrast, diagnosed diabetes prevalence was lowest among counties in the upper West and Midwest, parts of Alaska, and parts of New England. Undiagnosed diabetes prevalence similarly tended to be high among counties in the deep South, but also among counties in the Southwest and Alaska, whereas counties in New England and the upper West and Midwest tended to have lower undiagnosed diabetes prevalence. In both cases, there was significant variation among counties within as well as across states. At the county level, diagnosed and undiagnosed diabetes prevalence were positively correlated (Pearson correlation coefficient 0.77), but more so for women (0.73) than for men (0.57).

At the national level, diagnosed diabetes prevalence was marginally higher

among men (10.6% [10.4%, 10.8%]) than among women (9.9% [9.7%, 10.0%]), and undiagnosed diabetes prevalence was substantially higher among men (5.0% [4.5%, 5.5%]) than among women (3.2% [2.5%, 3.8%]). Consequently, at the national level total diabetes prevalence was also higher among men (15.6% [15.1%, 16.2%]) than among women (13.0% [12.3%, 13.7%]), a pattern that was reflected in 95.1% of counties.

Nationally, diabetes awareness was 71.6% (69.5%, 73.7%) in 2012, but varied by county, ranging from 59.1% to 79.8%. Similarly, whereas at the national level 26.9% (23.3%, 30.6%) of individuals who had previously received a diagnosis of diabetes had brought their diabetes under control (i.e., FPG <126 mg/dL and A1C <6.5% [48 mmol/mol]), this ranged from 19.4% to 31.0% at the county level. Figure 2 depicts age-standardized diabetes awareness and control at the county level. Awareness was highly correlated with total diabetes prevalence (Pearson correlation coefficient 0.77) and tended to be highest in
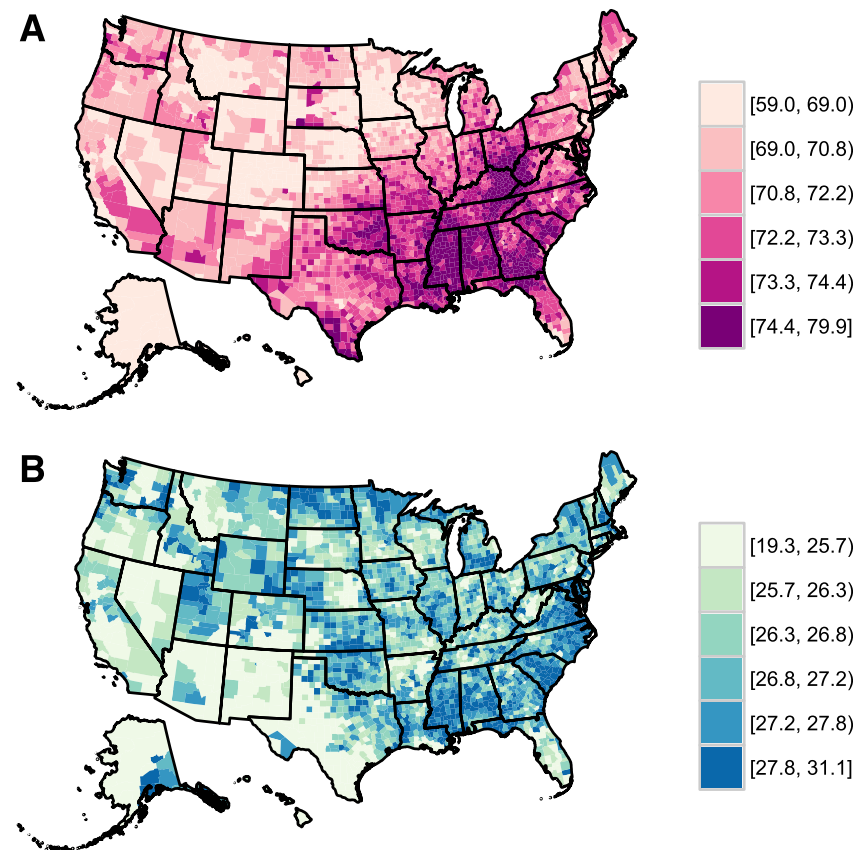
Figure 2—Age-standardized diabetes awareness and control by county, 2012. Diabetes awareness (A) and diabetes control (B).

counties in the South and in eastern Kentucky and West Virginia; and lowest in counties in the upper West and Midwest, Alaska, and parts of New England. In contrast, there was a small negative correlation between control and total diabetes prevalence (Pearson correlation coefficient −0.08). Control tended to be highest among counties in the deep South and along the Atlantic coast; and lowest among counties in the West, Southwest, and Alaska. At the national level, both awareness and control were higher for women than for men (75.7% [72.0%, 79.4%] vs. 68.0% [65.8%, 70.1%] for awareness; 30.7% [26.5%, 35.0%] vs. 23.2% [17.6%, 28.7%] for control), a pattern that was reflected in nearly all counties.

County-level estimates of all outcomes in all years are available from the authors upon request.
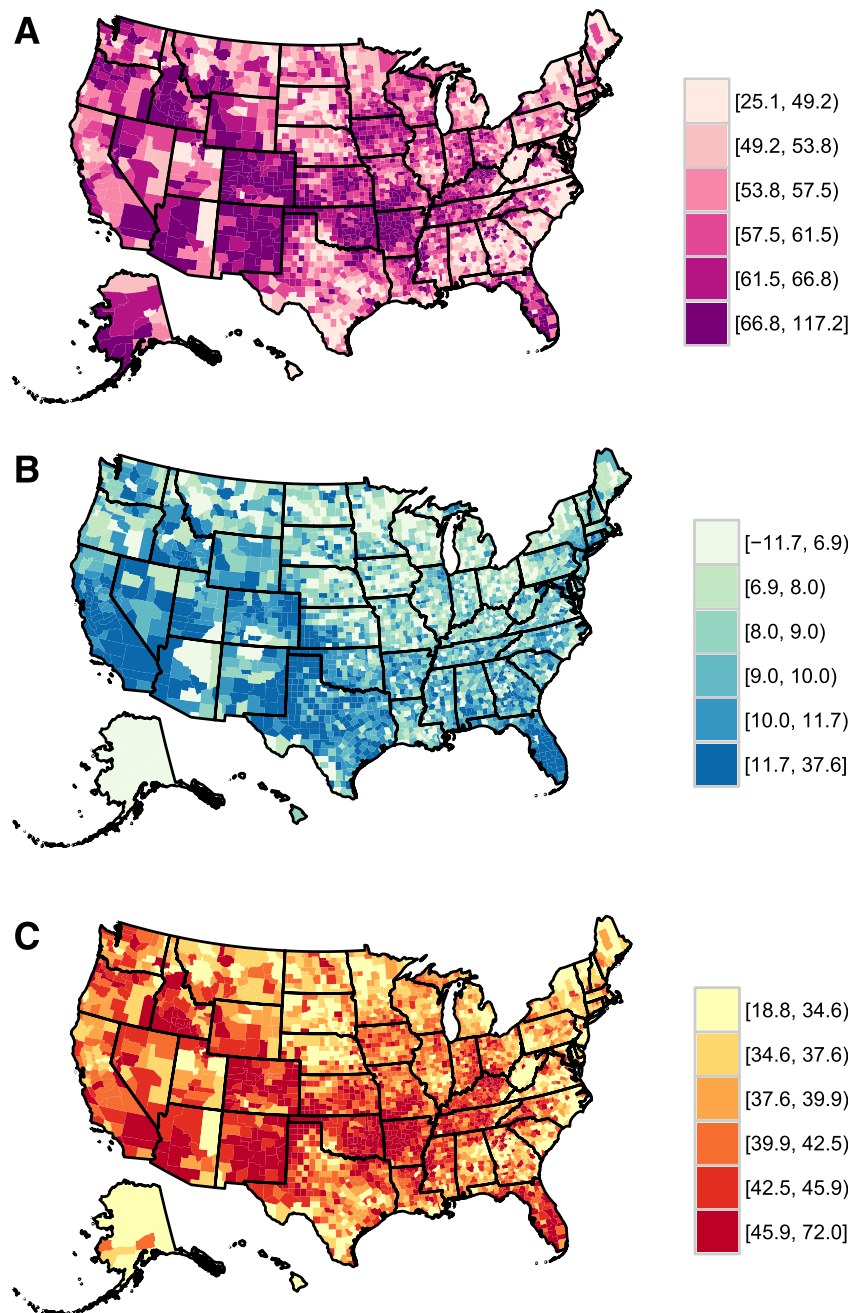
Change in Diabetes Prevalence From 1999 to 2012

Between 1999 and 2012, total diabetes prevalence nationally increased by 40.0% (35.3%, 44.8%), from 10.2% (9.7%, 10.7%) to 14.3% (13.8%, 14.7%). This reflects an increase in both diagnosed and undiagnosed diabetes, but the rate of increase was larger for diagnosed than undiagnosed diabetes: 56.8% (52.3%, 61.7%) compared with 10.3% (4.8%, 15.7%). Changes in diabetes prevalence varied at the county level, however, with increases ranging from 25.2% to 117.1% for diagnosed diabetes and from 18.9% to 72.0% for total diabetes. Changes in undiagnosed diabetes prevalence ranged from a decline of 11.6% to an increase of 37.5%. We estimated a decline in undiagnosed diabetes prevalence in 0.5% of counties; however, this decline was not statistically significant in any county (one-tailed test, α = 0.05). Figure 3 shows the percentage changes in age-standardized diagnosed, undiagnosed, and total diabetes at the county level. Counties with relatively small and relatively large increases in diagnosed diabetes are distributed throughout the country, although concentrations of counties with large increases are seen in the West, Southwest, and southern half of the Midwest, whereas a large number of counties with relatively small increases in diagnosed diabetes can be found along the Atlantic coast and parts of the deep South. Similarly, below and above average increases in undiagnosed diabetes

were realized throughout the country, although in general there is a higher concentration of counties with large increases in the South and West and in Florida, and a higher a concentration of counties with small increases in the North and East and in Alaska. The map for changes in total diabetes prevalence reflects the map for changes in diagnosed diabetes, because increases in total diabetes were in large part driven by changes in diagnosed rather than undiagnosed diabetes prevalence.

Nationally, awareness increased by 12.0% (10.4%, 13.6%) between 1999 and 2012, from 63.9% (61.6%, 66.3%) to 71.6% (69.5%, 73.7%). At the same time, control has held roughly constant, increasing by 1.5% (−5.9%, 9.1%). from 26.5% (22.3%, 30.8%) to 26.9% (23.3%, 30.6%). Over this same period, we found increases in awareness for all counties, ranging from 4.8% to

Figure 3—Percentage change in age-standardized diabetes prevalence by county, 1999–2012. A: Diagnosed diabetes prevalence. B: Undiagnosed diabetes prevalence. C: Total diabetes prevalence.

38.2%. Changes at the county level in diabetes control were more mixed, however, ranging from a 12.3% decline to a 31.1% increase. Figure 4 shows the percentage changes in awareness and control at the county level. The largest gains in awareness were realized in counties in the Midwest, Southwest, Pacific Northwest, and Alaska, whereas the smallest gains were typically observed in the East and the South. Counties that increased control most dramatically tended to be clustered in and around Virginia, Oklahoma, and North Dakota, whereas those where control declined are somewhat concentrated along the coasts but also are well represented throughout the interior.

CONCLUSIONS

The substantial and increasing health and financial burden of diabetes in the U.S. has been well documented (1,3). Existing estimates (8,23) of county-level diagnosed diabetes have previously highlighted a dramatic variation in prevalence within the U.S. Our findings on diagnosed diabetes are v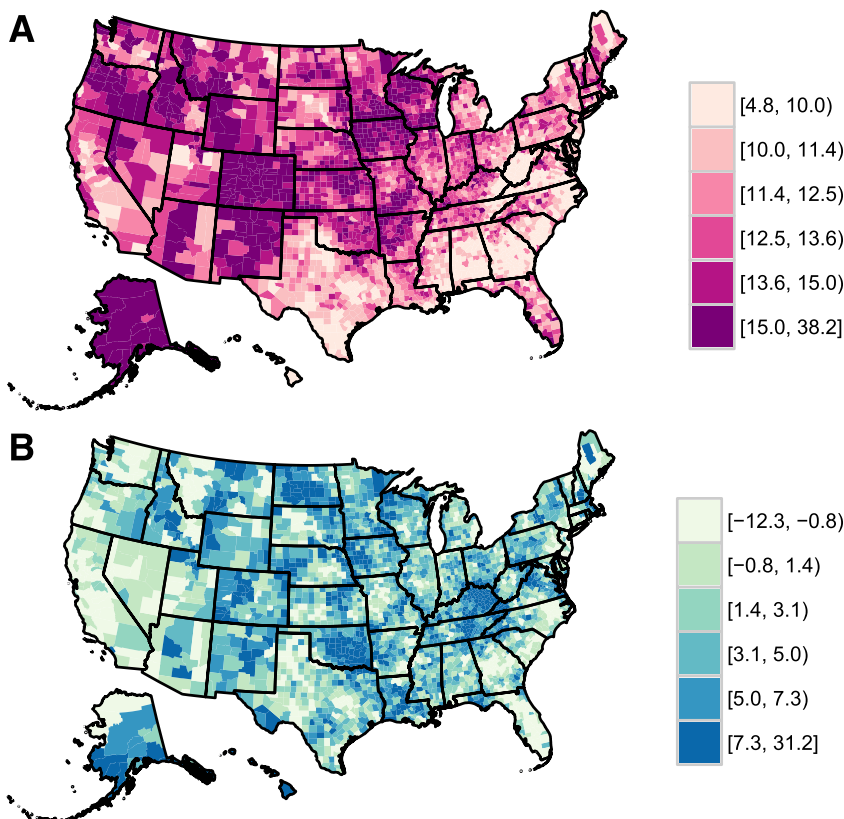ery similar (for 2012, the correlation between the two sets of estimates is 0.79 for men and 0.82 for women), but we were also able to report on undiagnosed and total diabetes prevalence, as well as on diabetes awareness and control at the county level. These results reveal significant variation within the U.S. and within states not only in undiagnosed and diagnosed diabetes prevalence, but also in local capacity to address the burden of diabetes through diagnosis and successful treatment. This type of local information is essential in order to identify the most impacted communities, and to enable public health officials to design targeted and effective intervention strategies.

This analysis is subject to a number of limitations. Most importantly, high FPG/A1C was imputed for BRFSS informants based on relevant variables shared between the BRFSS and NHANES rather than measured directly, and as a result the estimates of undiagnosed diabetes prevalence, total diabetes prevalence, diabetes awareness, and diabetes control are considerably less precise than the estimates of diagnosed diabetes prevalence, as evidenced by the much larger uncertainty intervals. Further, county-level estimates of undiagnosed and total diabetes, as well as the other measures derived from these, account for the variation in diagnosed diabetes, demographic features, BMI, smoking, and health insurance, but not for other factors. This is reflected by the relatively low sensitivity of the models for predicting high FPG/A1C—although the variables included in the model are certainly predictive of diabetes, they explain only a small portion of the individual-level variation in diabetes risk. As such, we are almost certainly underestimating the true variation in these outcomes and may be missing important outlier counties with unexpectedly high or low performance in terms of diagnosis and treatment. This analysis represents an important step forward in beginning to account for undiagnosed diabetes in addition to diagnosed diabetes, and also in exploring the variation in awareness and control, but further work on these topics is certainly needed, and will likely involve more substantial data collection at the county level.

The NHANES and BRFSS are both subject to nonresponse bias. We address this issue by explicitly incorporating many of the variables used to develop sample weights for both surveys into the small area model and poststratifying the results. Further, the BRFSS is also potentially subject to noncoverage bias because individuals without phones cannot be interviewed and a cell phone sample was only added in 2011. Previous research, however, suggests that the bias due to omission of cell phones is expected to be small for diabetes (19), and we explicitly correct for this bias. Nonetheless, it is possible that some bias due to nonresponse or noncoverage remains.

These limitations notwithstanding, this study also has a number of strengths. Most importantly, we made efficient use of the available data, capitalizing on the strengths of the BRFSS, namely its large sample size and broad geographic coverage, as well as on the strengths of the NHANES, in particular the collection of biomarker data. This allowed us to generate a significantly more detailed picture of diabetes prevalence at the county level than has previously been available. Further, we used sophisticated small area models, which simultaneously borrow strength spatially, temporally, and from external sources of information, allowing us to generate more precise estimates for each county than is

[4.8, 10.0)
[10.0, 11.4)
[11.4, 12.5)
[12.5, 13.6)
[13.6, 15.0)
[15.0, 38.2]

[−12.3, −0.8)
[−0.8, 1.4)
[1.4, 3.1)
[3.1, 5.0)
[5.0, 7.3)
[7.3, 31.2]

Figure 4—Percentage change in age-standardized diabetes awareness and control by county, 1999–2012. Diabetes awareness (A) and diabetes control (B).

possible in a strictly design-based setting. Finally, our methods explicitly accounted for uncertainty in all modeling stages, and the results are accompanied by 95% uncertainty intervals to convey the level of precision associated with each estimate.

The variation in total diabetes prevalence within the U.S. is staggering, with a threefold difference between the counties with the lowest prevalence and those with the highest prevalence. Some of this variation can be accounted for by socioeconomic and demographic factors, which are explicitly incorporated in our analysis of undiagnosed and total diabetes prevalence. However, our estimates of diagnosed diabetes, which are based on data directly observed at the county level, suggest that there is more variation in diabetes prevalence among counties than can be explained by socioeconomic and demographic differences alone. Further, the underlying factors driving differences between socioeconomic and demographic groups have not been entirely elucidated. Given the significant health and financial burden of high diabetes prevalence, this disparity demands further investigation into what underlying (and potentially modifiable) factors drive the exceedingly high diagnosed and total diabetes rates found in many communities.

Diabetes is both preventable and treatable. The public health system has a roll to play in increasing awareness of and screening for diabetes, connecting affected and high-risk individuals with appropriate medical care, and promoting community-level interventions that address known risk factors such as poor diet or lack of physical activity (4,5). The results of this analysis should be considered by state and local health officials aiming to increase early detection and improve the health of impacted communities.

References

1. GBD 2013 Mortality and Causes of Death Collaborators. Global, regional, and national age-sex specific all-cause and cause-specific mortality for 240 causes of death, 1990-2013: a systematic analysis for the Global Burden of Disease Study 2013. Lancet 2015;385:117–171

2. Vos T, Barber RM, Bell B, et al.; Global Burden of Disease Study 2013 Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990-2013: a systematic analysis for the Global Burden of Disease Study 2013. Lancet 2015;386:743–800

3. American Diabetes Association. Economic costs of diabetes in the U.S. in 2012. Diabetes Care 2013;36:1033–1046

4. Batan M, Jayapaul-Philip B, Gruss S, et al. *Effective Public Health Strategies to Prevent and Control Diabetes*. Atlanta, GA, Centers for Disease Control and Prevention, 2013

5. Centers for Disease Control and Prevention. The guide to community preventive services: diabetes prevention and control [Internet], 2016. Available from http://www.thecommunityguide.org/diabetes/index.html. Accessed 28 February 2016

6. Centers for Disease Control and Prevention. National Health and Nutrition Examination Survey [Internet], 2016. Available from http://www.cdc.gov/nchs/nhanes.htm. Accessed 11 October 2013

7. Srebotnjak T, Mokdad AH, Murray CJ. A novel framework for validating and applying standardized small area measurement strategies. Popul Health Metr 2010;8:26

8. Cadwell BL, Thompson TJ, Boyle JP, Barker LE. Bayesian small area estimates of diabetes prevalence by U.S. county, 2005. J Data Sci 2010;8:173–188

9. Congdon P, Lloyd P. Estimating small area diabetes prevalence in the US using the Behavioral Risk Factor Surveillance System. J Data Sci 2010;8:235–252

10. Centers for Disease Control and Prevention. Behavioral Risk Factor Surveillance System [Internet], 2016. Available from http://www.cdc.gov/brfss/. Accessed 5 August 2013

11. National Association of County and City Health Officials. *2013 National Profile of Local Health Departments*. Washington, DC, National Association of County and City Health Officials, 2014

12. American Diabetes Association. Standards of medical care in diabetes, 2016. Diabetes Care 2016;39(Suppl. 1):S1–S112

13. Danaei G, Friedman AB, Oza S, Murray CJ, Ezzati M. Diabetes prevalence and diagnosis in US states: analysis of health surveys. Popul Health Metr 2009;7:16

14. Olives C, Myerson R, Mokdad AH, Murray CJL, Lim SS. Prevalence, awareness, treatment, and control of hypertension in United States counties, 2001-2009. PLoS One 2013;8:e60308

15. King G, Tomz M, Wittenberg J. Making the most of statistical analyses: improving interpretation and presentation. Am J Pol Sci 2000;44:347–361

16. Leroux BG, Lei X, Breslow N. Estimation of disease rates in small areas: a new mixed model for spatial dependence. In *Statistical Models in Epidemiology, the Environment, and Clinical Trials*. Halloran ME, Berry D, Eds. New York, Springer-Verlag, 2000, p. 179–191

17. Knorr-Held L. Bayesian modelling of inseparable space-time variation in disease risk. Stat Med 2000;19:2555–2567

18. Dwyer-Lindgren L, Mokdad AH, Srebotnjak T, Flaxman AD, Hansen GM, Murray CJ. Cigarette smoking prevalence in US counties: 1996-2012. Popul Health Metr 2014;12:5

19. Centers for Disease Control and Prevention (CDC). Methodologic changes in the Behavioral Risk Factor Surveillance System in 2011 and potential effects on prevalence estimates. MMWR Morb Mortal Wkly Rep 2012;61:410–413

20. Kristensen K, Nielsen A, Berg CW, Skaug H, Bell B. TMB: automatic differentiation and Laplace approximation. J Stat Softw 2016;70:1–21

21. R Core Team. R: a language and environment for statistical computing [Internet], 2016. Vienna, Austria, R Foundation for Statistical Computing. Available from http://www.R-project.org/. Accessed 15 March 2016

22. Rubin DB, Ed. *Multiple Imputation for Nonresponse in Surveys*. Hoboken, NJ, Wiley, 1987

23. Centers for Disease Control and Prevention, Division of Diabetes Translation. County data indicators: diagnosed diabetes prevalence by sex [Internet], 2013. Available from http://www.cdc.gov/diabetes/atlas/countydata/atlas.html. Accessed 15 July 2015