OBJECTIVE — To examine trends in the prevalence of diagnosed and undiagnosed diabetes and the proportion of total cases previously diagnosed, according to obesity status in the U.S. over the past 40 years.

RESEARCH DESIGN AND METHODS — We assembled data from five consecutive cross-sectional national surveys: National Health Examination Survey I (1960–1962), National Health and Nutrition Examination Survey (NHANES) I (1971–1974), NHANES II (1976–1980), NHANES III (1988–1994), and NHANES 1999–2000. Diagnosed diabetes was ascertained, and height and weight were measured in adults aged 20–74 years in all surveys. In NHANES II, NHANES III, and NHANES 1999–2000, a fasting glucose level ≥126 mg/dl was used to identify cases among individuals not reporting diabetes. Design-based analyses and Bayesian models estimate the probability that prevalence of diabetes increased within four BMI groups (<25, 25–29, 30–34, and ≥35 kg/m²).

RESULTS — In the U.S. population aged 20–74 years between 1976–1980 and 1999–2000, significant increases in the prevalence of diagnosed diabetes (3.3–5.8%, probability >99.9%) were accompanied by nonsignificant increases in undiagnosed diabetes (2.0–2.4%, 66.6%). This resulted in an increase in total diabetes (5.3–8.2%, >99.9%) and a modest nonsignificant increase in the proportion of cases that were diagnosed (62–70%, 62.4%). However, these trends varied considerably by BMI level. In individuals with BMI ≥35 kg/m², diagnosed diabetes increased markedly (from 4.9% in 1960, to 8.6% during 1976–1980, to 15.1% in 1999–2000; probability >99.9%), whereas undiagnosed diabetes declined considerably (12.5% during 1976–1980 to 3.2% in 1999–2000; probability of increase 4.3%). Therefore, the proportion of total diabetes cases that were diagnosed increased from 41 to 83% (probability 99.9%) among individuals with BMI ≥35 kg/m². By comparison, changes in prevalence within BMI strata <35 kg/m² were modest and there was no increase in the percent of total cases that were diagnosed.

CONCLUSIONS — National surveys over the last several decades have found large increases in diagnosed diabetes, particularly in overweight and obese individuals, but this has been accompanied by large decreases in undiagnosed diabetes only among individuals with BMI ≥35 kg/m². This suggests that improvements in diabetes awareness and detection are most prominent among this subgroup.
data from five consecutive nationally representative surveys. We present 40-year trends in the prevalence of diabetes and the proportion of total diabetes cases that are diagnosed according to levels of BMI.

RESEARCH DESIGN AND METHODS — The National Health Examination Survey (NHES) and National Health and Nutrition Examination Survey (NHANES) are a series of health surveys designed to be representative of the noninstitutionalized U.S. population, conducted from 1960 to 1962 (NHES), 1971 to 1975 (NHANES I), 1976 to 1980 (NHANES II), 1988 to 1994 (NHANES III), and 1999 to 2000 (NHANES 1999–2000) (16–20). Each of the surveys followed a stratified multistage probability design in which a sample of the U.S. civilian, noninstitutionalized population was selected (16–20).

To maximize comparability across the five surveys, we restricted analyses to nonpregnant adults aged 20–74 years who provided height and weight as part of a standardized examination and information about diabetes history as part of an interview. In the five surveys, there were 6,240, 12,900, 11,761, 14,301, and 3,598 participants, respectively. Overall response rates among adults in these surveys were 87, 70, 69, 77, and 76%, respectively (16–20).

BMI and demographic measures
Each NHANES consists of a detailed standardized medical examination in a mobile examination unit and an interview to obtain information on sociodemographic characteristics and cardiovascular risk factors (16–20). Weight and height were measured using a standard protocol and were used to calculate BMI (weight in kilograms divided by the square of the height in meters). Race and ethnicity were assessed differently in the latter two surveys than in the first three surveys. Therefore, we categorized participants as white or black for NHANES I and NHANES II and as non-Hispanic white, non-Hispanic black, Mexican American, or other for NHANES III and NHANES 1999–2000. We excluded individuals older than 74 years because patients in that age-group were not sampled in the first three surveys.

Definition of diabetes
Diagnosed diabetes was defined as individuals’ positive response to the question of whether they had ever been told by a doctor or other health professional that they had diabetes. In the NHANES III and NHANES 1999–2000, women who reported being diagnosed only during pregnancy were not considered to have diabetes. This information was not available for the earlier three surveys.

For the latter three surveys, undiagnosed diabetes was assessed in a subsample (n = 3,786, 5,800, and 1,441 in NHANES II, III, and 1999–2000, respectively) of nondiabetic persons who were randomly assigned to a morning fasting examination (2,4,21). Procedures for blood collection and processing have been described previously (2,4,20,21). We used American Diabetes Association (ADA) diagnostic criteria for undiagnosed diabetes (fasting blood glucose level \( \geq 126 \text{ mg/dl after} \geq 9 \text{ h fasting}. \))

Gestational diabetes values were not assessed in the first two surveys (NHES and NHANES I).

Data analysis
Statistical analyses were conducted using SAS for Windows software (SAS Institute, Cary, NC) (22) for data management, SUDAAN software (Research Triangle Institute, Research Triangle Park, NC) (23) to obtain point estimates and standard errors, and WinBUGS (Institute of Public Health, Cambridge, U.K.; Imperial College School of Medicine, London, U.K.) (24) to fit models. Survey weights were used that indicated the inverse of the probability of being sampled and adjust for noncoverage and nonresponse. Taylor series linearization was used to calculate SEs for NHES, NHANES I, NHANES II, and NHANES III, and the jackknife method was used for NHANES 1999–2000. For comparisons across surveys, data were age adjusted by the direct method to the U.S. Census 2000 population using the following age ranges: 20–34, 35–44, 45–54, 55–64, and 65–74 years.

Prevalence of diagnosed diabetes was calculated based on responses from individuals who attended the medical examination. Prevalence of undiagnosed diabetes was determined based on laboratory findings from those who were in the fasting morning subsample. The prevalence was adjusted to account for the fact that individuals with diagnosed diabetes were not part of the fasting morning sample denominator (2,25). Calculation of SEs assumed independence between the morning and examination samples. Prevalence of total diabetes was determined by adding the prevalences of diagnosed and adjusted undiagnosed diabetes. Our analyses examined trends in prevalence of diabetes (diagnosed, undiagnosed, and total) and the proportion of total cases that were diagnosed across survey years, by four BMI strata (<25, 25–29.9, 30–34.9, or \( \geq 35 \text{ kg/m}^2 \)). These categories correspond to normal, overweight, obese class I, and obese class II as suggested by the National Heart Lung and Blood Institute Expert Panel on Overweight and Obesity (26).

Trends in the age-adjusted estimates of diabetes prevalence were assessed by fitting a normal hierarchical Bayesian model using noninformative prior distributions (27). Using results of the Bayesian models, we report the probability that the prevalence has increased across the selected time frame for the overall population and according to BMI groups. Probability >50% indicates that prevalence is more likely to have increased rather than decreased, whereas a probability <50% suggests the prevalence is more likely to have decreased. Probability cut points of 97.5 and 2.5% may be used to assess significant increases or decreases. In other words, probability >97.5% may be interpreted as a significantly high probability that the prevalence increased, whereas probability <2.5% indicates a significantly high probability that the prevalence decreased. A more detailed description of the statistical approaches used is provided in the STATISTICAL APPENDIX.

RESULTS — Table 1 describes the general characteristics of the sample populations across the five surveys. There were only slight changes in the age distributions; the tendency in recent surveys was toward a larger proportion of younger people (20–44 years) relative to middle-aged people (45–64 years) and a slightly smaller proportion of women. A change in the survey questionnaires makes it impossible to directly compare changes in the racial and ethnic profile across all surveys, but an increase in the proportion of Mexican Americans and individuals of other races or ethnicities is evident in the most recent survey. The proportion of participants with less than a high school education decreased over time. As has been previously published,
there were notable increases in mean weight and BMI (from 25.4 kg/m² during 1960–1962 to 28.0 kg/m² during 1999–2000) and prevalence of obesity (from 14.5% during 1960–1962 to 26.7% during 1999–2000) (7).

Trends in diagnosed and undiagnosed diabetes prevalence
Between 1960–1962 and 1999–2000, the prevalence of diagnosed diabetes in the U.S. population increased from 1.8 to 5.8%, an average increase of one percentage point per decade (Table 2). The probability that prevalence increased was 99.9%. The rate of increase was greatest for those with BMI $\geq 35$ kg/m², among whom the prevalence tripled (e.g., from 4.9 to 15.1%, probability $>99.9$%) and whose per-decade increase was almost twice that of the population overall. Prevalence more than doubled among persons in the middle two BMI groups (25 to $<30$; 30 to $<35$), although the absolute magnitude of change was less in these groups (probability >99.5 for each). The prevalence of diagnosed diabetes doubled among individuals with normal BMI ($<25$) (from 1.5 to 3%), although the statistical probability of this change was weaker (probability 87.6%). We found similar trends when we limited our analysis to the latter three surveys (from 1976–1980 to 1999–2000); again, the increase in the prevalence of diagnosed diabetes was greatest among those who were the most obese (from 8.6 to 15.1%, probability 88.7%).

There were modest, nonsignificant increases in the prevalence of undiagnosed diabetes in the overall population and across all BMI strata, with the exception of individuals with BMI $\geq 35$ kg/m² (Table 2). Among those individuals, the prevalence of undiagnosed diabetes decreased from 4.9 to 3.2%. The probability of an increase was only 4.5%, meaning that the probability of a decrease was 95.5%. During the same time frame, the prevalence of total diabetes (diagnosed and undiagnosed combined) increased from 5.3 to 8.2% in the overall population (probability >99.9%) and remained relatively stable within BMI groups. This prevalence increased slightly among those with BMI $<35$ kg/m² (for example, from 4.3 to 4.9% among those with BMI 25.0–29.9 kg/m²) and decreased slightly among those with BMI $\geq 35$ kg/m² (from 21.1 to 18.3%), but the probabilities associated with these changes were modest.

Trends in the ratio of diagnosed to total diabetes
Between 1976 and 2000, the proportion of total diabetes cases that were diagnosed (versus undiagnosed) (Fig. 1) increased slightly in the overall population (from 62.3% to 70.7%), but this change was also of modest statistical significance (probability increase 62.4%). However, this increase seemed to be concentrated

| Table 1—Characteristics of participants aged 20–74 years, by survey* |
|--------------------------------|------------------|------------------|------------------|------------------|------------------|
| n                               | 6,240            | 12,900            | 11,761            | 14,301            | 3,598            |
| Age distribution (%)            |                  |                  |                  |                  |                  |
| 20–44 years                     | 54.4             | 54.2             | 56.2             | 60.4             | 57.6             |
| 45–64 years                     | 34.9             | 35.3             | 32.8             | 28.4             | 32.2             |
| 65–74 years                     | 10.8             | 10.5             | 11.0             | 11.2             | 10.2             |
| Women (%)                       | 52.5             | 52.2             | 51.9             | 50.7             | 50.4             |
| Race/ethnicity (%)              |                  |                  |                  |                  |                  |
| White                           | 88.1             | 89.0             | 87.3             | —                | —                |
| Black                           | 10.2             | 10.0             | 10.2             | —                | —                |
| Non-Hispanic white              | —                | —                | —                | 75.5             | 70.3             |
| Non-Hispanic black              | —                | —                | —                | 11.1             | 11.1             |
| Mexican American                | —                | —                | —                | 5.2              | 6.6              |
| Other                           | 1.7              | 1.0              | 2.5              | 8.1              | 12.0             |
| Less than high school education (%) | 31.8           | 34.6             | 32.0             | 23.1             | 23.3             |
| Height (mean cm)                | 166.4            | 168.0            | 168.3            | 168.9            | 169.1            |
| Weight (mean kg)                | 70.3             | 71.6             | 71.9             | 76.1             | 80.4             |
| BMI (kg/m²)†                    | 25.4             | 25.3             | 25.3             | 26.6             | 28.0             |
| Percentage in each BMI category |                  |                  |                  |                  |                  |
| <25 kg/m²                       | 52.1             | 53.4             | 54.0             | 45.4             | 36.3             |
| 25–29.9 kg/m²                   | 33.4             | 32.3             | 31.5             | 32.1             | 33.4             |
| 30–34.9 kg/m²                   | 10.9             | 10.2             | 10.1             | 14.4             | 16.9             |
| ≥35 kg/m²                       | 3.6              | 4.1              | 4.4              | 8.1              | 13.5             |
| Mean BMI in each category       |                  |                  |                  |                  |                  |
| <25 kg/m²                       | 21.9             | 21.8             | 21.9             | 22.1             | 22.2             |
| 25–29.9 kg/m²                   | 27.2             | 27.1             | 27.1             | 27.2             | 27.4             |
| 30–34.9 kg/m²                   | 31.9             | 31.9             | 32.0             | 32.0             | 32.3             |
| ≥35 kg/m²                       | 38.7             | 39.5             | 39.2             | 39.9             | 40.1             |

*Denominator is based on all participants who had BMI data obtained from the mobile examination center and responded to questions about previously diagnosed diabetes. †Data are SE.
among individuals with BMI $\geq 35$ kg/m$^2$, in whom the proportion of cases that were diagnosed increased from 40.8 to 82.5\% (probability 99.8\%). Therefore, persons with BMI $>35$ kg/m$^2$ went from having the lowest ratio of diagnosed to total diabetes cases in 1976–1980 to the highest ratio during 1976–1980 to ~83\% of cases in 1999–2000. Within other BMI strata, this proportion was either stable or decreased slightly.

### CONCLUSIONS

In this examination of five consecutive national surveys from 1960 to 2000, we found increasing levels of total and diagnosed diabetes in the U.S., consistent with previous analyses of national data (1–3,14). However, previous studies have not examined the relationship of obesity status to trends in diagnosed or undiagnosed diabetes or to the proportion of total diabetes cases that were diagnosed. We found that the trends in these outcomes varied by level of obesity. The prevalence of diagnosed diabetes increased in all groups, but the most dramatic change was seen among the most obese individuals (BMI $\geq 35$ kg/m$^2$), prevalence was three times as high in 2000 as 1960. Underlying these findings in the most obese group was a dramatic increase in the proportion of diabetes cases that are diagnosed, from 41\% of total cases diagnosed during 1976–1980 to ~83\% in 1999–2000.

Our study has several implications for interpreting the diabetes epidemic in the U.S. First, our finding of a large increase in the proportion of diagnosed cases of diabetes in individuals with BMI $\geq 35$ kg/m$^2$ suggests that detection of diabetes has increased in this subgroup. Whereas 25 years ago this group had the highest percentage of diabetes cases that went undiagnosed, it now has the highest percentage of diagnosed cases. We speculate that better awareness among patients and providers about diabetes and its potential to be asymptomatic may have led to increased opportunistic screening of obese individuals in health care settings. More specifically, increasing attention to obesity and diabetes in both the lay and professional media may have conditioned health care providers to associate extreme obesity with a high risk of undiagnosed diabetes, making them more likely to question obese patients about symptoms and then to perform more testing. Despite this increase in the proportion of diagnosed cases among the most obese group, there was no increase in the proportion diagnosed in other subgroups. Roughly 30\% of the overall diabetic population is estimated to be undiagnosed, indicating a continuing need to improve detection of undiagnosed diabetes and to increase awareness of key diabetes risk factors.

Second, our finding that the prevalence rates for total diabetes (i.e., diagnosed and undiagnosed cases combined) have remained relatively stable within BMI strata, even as the BMI levels of Americans have increased, is consistent with prior arguments that obesity is the predominant factor driving the growth of the diabetes epidemic. This is reflected in observational studies showing that the association of obesity and weight gain with diabetes is nonlinear, with disproportionate increases in diabetes incidence among individuals with BMI $\geq 35$ kg/m$^2$ and weight gain (9,28). Similarly, in the NHANES 1999–2000 sample, individuals with BMI $\geq 35$ kg/m$^2$ account for only 13.5\% of the overall population but for 36\% of the diabetic population. Other factors, such as levels of physical activity,
diet (e.g., variations in whole grain and fiber, dairy, saturated fat, and caffeine intake), or other unknown environmental factors may also be important (11,29–35). However, if changes in these factors were playing an important role in the diabetes epidemic above and beyond their effects on obesity, we would expect to see changes in total diabetes prevalence in groups with similar BMI.

A third implication of our findings is that because much diabetes surveillance is based on self-report of diabetes, a simultaneous increase in obesity prevalence and the detection of diabetes among the most obese could complicate our interpretation of historical trend data (1,3,8,13,14). In light of these observations, it is possible that the observed increases in self-reported diagnosed diabetes have slightly overestimated the true increase in total diabetes prevalence (1). We are also reminded that, although several studies have reported higher rates of diagnosed diabetes in recent decades, the NHANES data are the only national data examining changes in total diabetes prevalence (2).

A final implication of our findings is that greater detection of diabetes among the most obese will result in a higher prevalence of obesity among the population with diagnosed diabetes. This is reflected in an average increase in BMI among individuals with diagnosed diabetes from 27.3 kg/m² in NHANES 1 to 33.0 kg/m² in NHANES 1999–2000 (36). Therefore, controlling obesity may become an even greater focus in diabetes management and perhaps have implications for reducing the incidence of obesity-related complications, including hypertension, cardiovascular disease, and physical disability in the population with diabetes.

An important caveat to our study is that our analyses of total diabetes and the ratio of diagnosed to undiagnosed take the perspective of the current ADA diagnostic criteria (fasting glucose level >126 mg/dl) and assume these criteria were in use throughout the surveys. In reality, diagnostic standards changed considerably over this time, with no universal definitions before 1979 and encouragement of oral glucose tolerance tests in the 1980s and 1990s, followed by reliance on fasting glucose values and a lowering of the diagnostic threshold in 1997 (37). We applied the ADA criteria retrospectively in these analyses so that we could compare prevalences of undiagnosed and total diabetes across the latter three surveys. (Oral glucose tolerance test data were only available for two surveys on subsamples of persons aged 40–74 years.) Unfortunately, we lack medical record validation of diabetes care to examine how cases were truly diagnosed in the clinical setting or how this could have influenced our findings. However, the DECODE study has shown that using the ADA criteria is more likely to detect obese persons than World Health Organization criteria (38). Thus, it is conceivable that part of the greater increase in proportion diagnosed among the obese reflects clinicians shifting to the use of fasting glucose for diagnostic purposes. However, we are unaware of data to either support or deny this speculation.

Our study has other limitations. First, the analysis of trends is based on estimates of point prevalence. Estimates based on incidence would be an ideal way to assess shifts in diagnosed versus undiagnosed diabetes because there is likely to be less variation or confounding by differences in disease duration at onset or by improved survival in recent decades. A second limitation is that we only had data on undiagnosed diabetes in the most recent surveys. Therefore, although we observed consistent trends in diagnosed diabetes dating all the way back to 1960, we can generalize only about effects on undiagnosed diabetes and detection since 1976. Third, our analysis did not adjust for the effects of shifts in race and ethnicity over the survey years because of changes in the way race and ethnicity were assessed. We also lacked power to stratify our analysis by race and ethnicity to determine, for example, whether the growth in the Mexican-American population affected our findings. However, when we excluded Mexican Americans from our analyses, we found essentially no difference in our overall findings. Finally, although our estimates come from large, nationally representative surveys, our estimates of undiagnosed diabetes, particularly in the most recent survey, are based on relatively small samples. More data to evaluate undiagnosed diabetes will be forthcom-
ing from the National Center for Health Statistics to allow further evaluation of undiagnosed diabetes, but other investigators should explore these questions using other currently available data, especially where glucose measurements extend over several decades.

In summary, we demonstrate an increasing prevalence of diagnosed diabetes that is most prominent in overweight and obese persons and apparent increases in diabetes detection among the most obese. Our findings of continued high prevalence of undiagnosed cases indicates a need to increase detection of diabetes across the full spectrum of body weight. Given the dynamic nature of the diabetes epidemic, along with the potential to increase both prevention and detection of the disease, continuing to track trends in both diagnosed and undiagnosed diabetes will be increasingly important.

STATISTICAL APPENDIX

Use of Bayesian hierarchical models

Analyses for this study were conducted using normal hierarchical Bayesian model with noninformative prior distributions on model hyperparameters. A Bayesian approach was preferred because inferences do not depend on asymptotic assumptions. Bayesian approaches make use of prior information (or prior distributions) and synthesizes this information with the data through the use of Bayes theorem to produce an updated distribution (or posterior distribution). Bayesian probabilities have a direct interpretation as the probability that the hypotheses is true.

Our Bayesian models have two levels. At the first level, we specify a probability distribution for the observed data. Letting \( y_i \) denote the estimate for diabetes prevalence determined from survey \( i \) and \( \sigma_i^2 \) denote the variance for \( y_i \) estimated from SUDAAN software (Research Triangle Institute, Research Triangle Park, NC), we assume

\[
y_i \sim N(\theta_i, \sigma_i^2)
\]

The survey means, \( \theta_i \), are assumed to come from a linear model

\[
\theta_i = \alpha + \beta(t_i - t_1) + Z_i
\]

where \( t_i \) represents the midpoint year of survey \( i \), \( \alpha \) represents the mean value of the risk factor from the first survey (intercept), and \( \beta \) represents the annual change in the risk factor overtime. The random-effect \( Z_i \) represents the variability in \( \theta_i \) remaining after accounting for the linear trend, \( Z_i \) is assumed to be distributed \( N(0, \tau^2) \).

At the second level of our model, we assign probability distributions to each hyperparameter (\( \alpha \), \( \beta \), and \( \tau \)). Normal priors with variance of 10,000 are assigned to \( \alpha \) and \( \beta \), whereas \( \tau^{-1} \) is approximated with \( \gamma(0.001, 0.001) \). Combining the prior probabilities specified at the second level with the probability distribution for the observed data specified at the first level, we express knowledge about model parameters through the posterior distribution, derived using Bayes theorem. Random draws from this posterior distribution are generated through an iterative Markov Chain Monte Carlo simulation known as the Gibbs sampler.

We simultaneously ran three Gibbs samplers and saved 50,000 of the 60,000 iterations from each. Model convergence was assessed visually by inspecting trace plots of sampled values versus iteration number and inspecting the Gelman-Rubin statistic for each iteration. We used the mean of the posterior distribution of \( \beta \) as our point estimate and estimated the probability that \( \beta \) is negative. Significant decreases in diabetes prevalence were defined by probabilities >97.5%.

References

24. MRC Biostatistics Unit: WinBUGS. Cambridge, U.K., Institute of Public Health, and London, Imperial College School of Medicine, 2003