Volume Title: Insights in the Economics of Aging

Volume Author/Editor: David A. Wise, editor

Volume Publisher: University of Chicago Press

Volume ISBNs: 0-226-42667-X; 978-0-226-42667-9 (cloth); 978-0-226-42670-9 (e-ISBN)

Volume URL: http://www.nber.org/books/wise-21

Conference Date: April 30-May 2, 2015

Publication Date: March 2017

Chapter Title: Comment on "Measuring Disease Prevalence in Surveys: A Comparison of Diabetes Self-Reports, Biomarkers, and Linked Insurance Claims"

Chapter Author(s): James P. Smith

Chapter URL: http://www.nber.org/chapters/c13638

Chapter pages in book: (p. 252 – 257)

Reynolds, Timothy, Stuart Smellie, and Patrick Twomey. 2006. "Glycated Haemo-
    globin (HbA1c) Monitoring." *British Medical Journal* 333 (7568): 586–88.
Rohlfing, C. L., R. R. Little, H. H. Wiedmeyer, J. D. England, R. Madsen, M. I.
    Harris, K. M. Flegal, M. S. Eberhardt, and D. E. Goldstein. 2000. "Use of GhB
    (HbA1c) in Screening for Undiagnosed Diabetes in the US Population." *Diabetes
    Care* 23:187–91.
Sakshaug, Joseph W., David R. Weir, and Lauren H. Nicholas. 2014. "Identifying
    Diabetics in Medicare Claims and Survey Data: Implications for Health Services
    Research." *BMC Health Services Research* 14 (150): 1–6.
Smith, James P. 2007. "Nature and Causes of Trends in Male Diabetes Prevalence,
    Undiagnosed Diabetes, and the Socioeconomic Status Health Gradient." *Pro-
    ceedings of the National Academy of Sciences of the United States of America*
    204:13225–31.
Wansbeek, Tom, and Erik Meijer. 2000. *Measurement Error and Latent Variables in
    Econometrics*. Amsterdam: Elsevier.
Wolinsky, Frederik D., Michael P. Jones, Fred Ullrich, Yiyue Lou, and George L.
    Wehby. 2014. "The Concordance of Survey Reports and Medicare Claims in a
    Nationally Representative Longitudinal Cohort of Older Adults." *Medical Care*
    52 (5): 462–68.
Yasaitis, Laura, Lisa Berkman, and Amitabh Chandra. 2015. "Comparison of Self-
    Reported and Medicare Claims-Identified Acute Myocardial Infarction." *Circula-
   tion* 131:1477–85.

Comment   James P. Smith

In a thought-provoking chapter, Heiss et al. raise several important ques-
tions about the appropriate way to measure diabetes prevalence in house-
hold surveys. While diabetes is the disease at issue in the chapter, the same
questions would arise with many other disease outcomes. Three common
measures of diabetes prevalence are used and compared in their analysis—
self-reports of ever being diagnosed by a doctor, the common HbA1c diabe-
tes biomarker being above the standard American threshold of 6.5 percent,
and a diabetes diagnosis mentioned in Medicare claims data. The question
the authors ask is whether the three measures are "consistent" and which
one is "correct."

    Figure 7C.1, derived from the chapter, illustrates the central finding of the
chapter by showing diabetes prevalence rates for a sample of HRS respon-
dents who had their diabetes measured in all three ways in 2006 and in
2008. Rates of diabetes prevalence are clearly quite different using the three
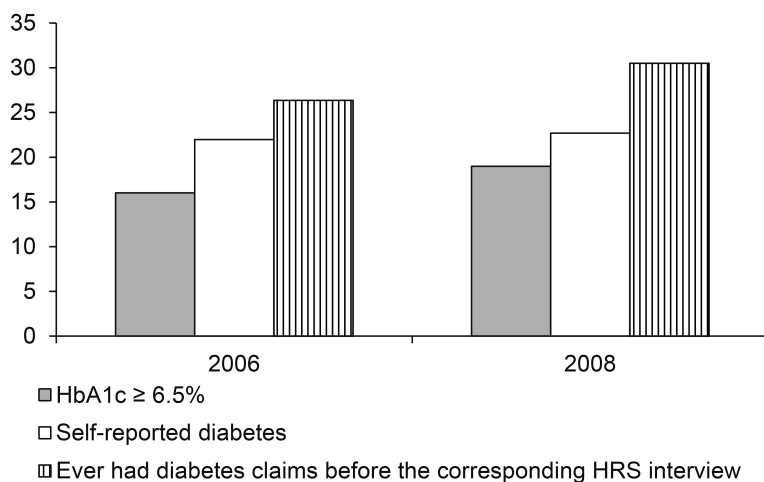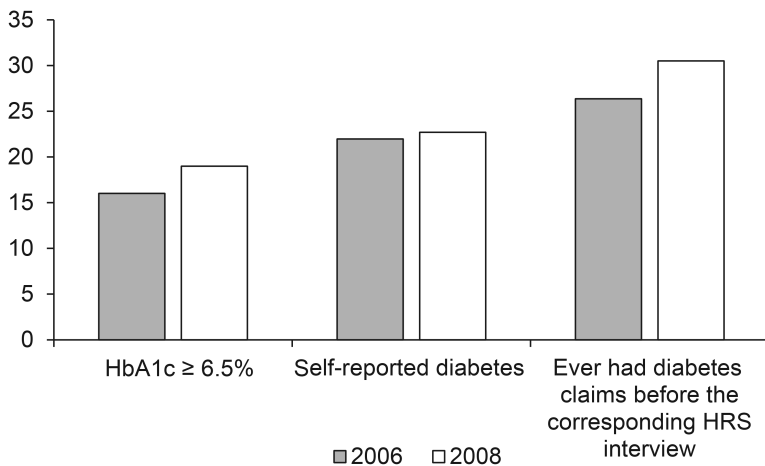
Fig. 7C.1   Diabetes indicators

measures—19.0 percent for the biomarker measure, 22.7 percent for the self-report of ever diagnosed, and 26.4 percent using Medicare claims data.

The discrepancy between the biomarker index and the ever self-report is easy to explain since they are actually measuring very different things. The self-report is obtained from a question to respondents about whether they had *ever* been diagnosed by a doctor while the biomarker is an indicator of whether the respondent *currently* exceeds the diabetes threshold of 6.5 percent. There is no real inconsistency between these two measures since they are in fact measuring very different things (Smith 2007). Biomarker values above the diagnosis threshold for those who claimed they were never diagnosed are a possible indication of undiagnosed disease, an important phenomenon in itself. Similarly, biomarker values below the diagnosis threshold for those who claimed they were diagnosed in the past may indicate that the respondent is managing the disease well. Thus, there is no fundamental inconsistency between the ever self-report of diagnosis and the current biomarker indicator. They should be used together as they are in the chapter to provide insights into diagnosis, undiagnosis, and good disease management.

The real puzzle centers on the claim data measure, which is much higher than either of the other two. If it was the "correct" measure, it would imply that we are seriously underestimating diabetes prevalence in the age fifty and older population in the United States.

Figure 7C.2 uses the same data as in figure 7C.1, but rearranges it to highlight changes over time between the 2006 and 2008 waves. Once again, the depiction of trends varies significantly depending on which measure is used. While self-reports of ever diagnosed show very little change between the two HRS waves (less than 1 percentage point), the change is almost 3

Fig. 7C.2   Diabetes indicators

percentage points using the biomarker and 4 percentage points using claims data. It would be useful as well for research to focus also on the reasons for the discrepancies in the three measures in measuring changes in diabetes prevalence over time. Secular trends in diabetes prevalence are equally as important as comparisons at a moment in time, the primary focus of this chapter.

My main suggestion at the conference to the authors is that they extend their analysis from simply doing cross-sectional comparisons between the three prevalence measures by using the panel nature of HRS data, which contain repeat measures of all three prevalence concepts. Not only would this address another central question of the nature of secular changes in diabetes prevalence, but multiple measures in the HRS panel for all three diabetes prevalence measures can go a long way to cleaning each measure of any reporting errors that may be present.

Table 7C.1 illustrates the potential contribution of the panel component of HRS to measurement of the ever self-report of disease for three diseases—cancer, diabetes (the relevant disease for this chapter), and hypertension (HBP). These data are derived from the first eight waves of the original HRS sample. In this table, a "no" answer to the diabetes question is translated into a zero while a "yes" answer is translated into a 1. The first two columns list the fraction of respondents who always answered "no" or always "yes" to the ever self-report question. For diabetes, for eight waves in a row, 74.6 percent of HRS respondents said "no"; 6.4 percent said "yes" to the ever-diagnosed diabetes question, so we should be very confident in this 81 percent subsample who is a diabetic and who is not.

The interesting case is when there is not complete consistency in responses

Table 7C.1   **HRS disease measurement on ever self-report (prevalence)**

Diseases	All 0s (%)	All ones (%)	Mixed (%)	All 1s after all 0s (%)	0s after 1s (%)	0s after 1s "corrected" (%)
Cancer	82.0	3.9	14.1	11.9	2.2	1.0
Diabetes	74.6	6.4	19.0	15.1	3.9	0.7
HBP	33.6	29.0	37.5	27.0	9.5	3.7

over the eight waves so that we translate respondents' answers into mixed values of 0s and 1s. For diabetes, this represents 19 percent of the cases. However, values of zeros (not a diabetic) followed by a series of 1s are fully consistent since that simply means that there was an onset of diabetes during the first eight HRS waves that was not contradicted in a future wave. That situation represents 15.1 percent of the cases, leaving only 3.9 percent of cases with an obvious inconsistency (a 0 that follows a 1). One in twenty-five is not so bad.

The situation is even not near that dire, since many of these inconsistences are easy to repair. For example, a zero followed by two 1s and then followed by five 0s should, in this author's opinion, be changed to all 0s since five times in a row in the more recent waves of HRS the respondent said he/she was not told he/she was a diabetic. Thus, the use of the panel data in HRS results in a big decrease in measurement error and a big increase in signal/noise for self-reports of prevalence, and especially, and even more importantly, for the incidence of diabetes.

As table 7C.1 shows, the situation is very similar for "self-reports of ever-diagnosed cancer," but there remain a larger fraction of uncertain diagnosis for hypertension, a far less serious disease. But for all the diseases in HRS the use of the full panel waves of responses helps a great deal in determining with good confidence whether a respondent was ever diagnosed with a particular disease.

Next consider the use of the biomarker index HbA1c, a measure of the percent of hemoglobin molecules bound to glucose (Goldman et al. 2003). Biomarker data are a very useful addition to population-based aging surveys, but they should not be treated as an uncontested gold standard. Biomarkers suffer from their own forms of measurement error and there is often inconsistency between alternative biomarkers meant to measure the same thing. For diabetes, HbA1c and fasting glucose would be an excellent example. Fasting glucose is typically the clinical measure used to diagnose in the doctor's office, while HbA1c is the standard survey population measure since it does not require respondents to fast. Incomplete fasting and momentary stress affect the accuracy of both measures.

As mentioned above the most important point is that self-reports of dia-

betes and an HbA1c threshold are measuring ever diagnosed, and now there is no real inconsistency in their values. Instead of diagnosis, biomarkers are most useful in combination with self-reports of ever diagnosed as an indicator of undiagnosed disease or well-managed disease. It is also an excellent indicator of disease severity by using the continuous measure of the biomarker outcome. For example respondents scoring above 8 in their HbA1c test not only are much more likely to be diabetics, they also are much more likely to suffer from a more serious level of the disease.

Since biomarkers are now measured in every other wave of HRS, I would make the same comment about biomarkers that I made about "ever" self-report above. They are most useful when the full set of panel data available across all waves is employed. Having a respondent above the diagnosis threshold in multiple waves of the survey should be a reliable indication that she is, in fact, a diabetic.

I finish with the third measure—the report of diabetes in Medicare claims data. The real puzzle of the chapter is why the prevalence report of diabetes is so much higher in claims data compared to the self-report of ever diabetes. This difference is not unique to anything about the HRS since similar levels of self-reports of ever diabetes have been found in NHANES for the same age group as HRS (Sakshaug, Weir, and Nicholas 2014). Similar to this chapter, Sakshaug et al. also report much higher rates of diabetes prevalence in claims data compared to ever self-reports.

There is a natural temptation to treat claims data as the real gold standard, but in my view that would be a mistake. It might not even rank as a bronze standard. There are several reasons for this. First, diabetes is often put on claims data to justify taking blood—the "rule out" hypothesis—or to give a warning sign and to have a talk with the patient. Sakshaug, Weir, and Nicholas (2014) also report that the higher rates of diabetes prevalence in claims data are due to false positives and may indicate intensive monitoring of prediabetes patients, especially those with other cardiovascular risk factors. Once again, researchers may be on safer ground if they use multiple waves of the claims data as well. It might even be better to use two or more visits to the doctor to confirm the diabetes diagnosis.

Conclusions

In this intriguing and stimulating chapter, Heiss et al. demonstrate that using three common alternative measures of disease prevalence—self-reports of past prevalence, a common biomarker (HBa1c), and being listed in Medicare claims data—over the same sample of respondents produces very different measures of diabetes prevalence. From highest to lowest prevalence, claims data are ranked the highest and the biomarker the lowest.

In my view, comparing the three measures and looking for a winner in which is the best measure of diabetes prevalence is not really the thing to

do. Rather the issue is how to use them together, especially using the panel aspect of the data to obtain better health measurement and not just for knowing diabetes prevalence.

References

Goldman, N., I. Lin, M. Weinstein, and Y. Lin. 2003. "Evaluating the Quality of Self-Reports of Hypertension and Diabetes." *Journal of Clinical Epidemiology* 56:148–54.

Sakshaug, J. W., D. R. Weir, and L. H. Nicholas. 2014. "Identifying Diabetics in Medicare Claims and Survey Data: Implications for Health Services Research." *BMC Health Services Research* 14:150.

Smith, J. P. 2007. "Nature and Causes of Male Diabetes Trends, Undiagnosed Diabetes, and the SES Health Gradient." *Proceedings of National Academy of Sciences* 104 (33): 13225–31.