

Health Affairs

At the Intersection of Health, Health Care and Policy

Cite this article as:
Marc L. Berk, Claudia L. Schur and Jacob Feldman
Twenty-Five Years Of Health Surveys: Does More Data Mean Better Data?
Health Affairs, 26, no.6 (2007):1599-1611

doi: 10.1377/hlthaff.26.6.1599

The online version of this article, along with updated information and services, is available at:

<http://content.healthaffairs.org/content/26/6/1599.full.html>

For Reprints, Links & Permissions:

http://healthaffairs.org/1340_reprints.php

E-mail Alerts : <http://content.healthaffairs.org/subscriptions/etoc.dtl>

To Subscribe: <http://content.healthaffairs.org/subscriptions/online.shtml>

Health Affairs is published monthly by Project HOPE at 7500 Old Georgetown Road, Suite 600, Bethesda, MD 20814-6133. Copyright © 2007 by Project HOPE - The People-to-People Health Foundation. As provided by United States copyright law (Title 17, U.S. Code), no part of *Health Affairs* may be reproduced, displayed, or transmitted in any form or by any means, electronic or mechanical, including photocopying or by information storage or retrieval systems, without prior written permission from the Publisher. All rights reserved.

Not for commercial use or unauthorized distribution

Twenty-Five Years Of Health Surveys: Does More Data Mean Better Data?

Federal investment in expanding surveys has not been matched by investment in survey methods research.

by **Marc L. Berk, Claudia L. Schur, and Jacob Feldman**

PROLOGUE: In a story regularly recounted in introductory statistics courses, the *Chicago Daily Tribune's* postelection headline of 3 November 1948 proclaimed that Republican Thomas Dewey had defeated incumbent Democrat Harry S. Truman in the 1948 presidential race. It was a public display of the importance of timely data. Unlike today, when polls are completed on election eve, the Gallup poll was completed ten days before election.

The task of collecting accurate and timely data confronts today's researchers with a growing array of problems along dimensions that could barely be imagined in 1948. In this paper, Marc Berk, Claudia Schur, and Jacob Feldman explore how these problems affect health-related data. Such data are vital to the formation of sound policy, yet increases in funding and advances in survey methodology have not led to demonstrable improvements in data. Among the countervailing factors that threaten the integrity of information used in policy making are declining response rates, technological changes, and an increasingly complex health care environment. The authors conclude with proposals to improve the quality of data used for policy making.

The authors have contributed much to the history described in this paper. Berk (berk-marc@norc.org), a sociologist, is senior vice president at NORC at the University of Chicago, located in the Washington, D.C., suburbs. Earlier in his career, Berk directed the pilot study for the National Medical Expenditure Survey (NMES—the forerunner to the Medical Expenditure Panel Survey) and the design of its provider components. Schur, a vice president at NORC, is an economist who worked with Berk on the design of the 1987 NMES, directing several components. Feldman, a sociologist and senior fellow at NORC, designed and implemented the first-ever survey of health care use and spending using a national probability sample. He spent nearly twenty-five years at the National Center for Health Statistics (NCHS), retiring as associate director for analysis, epidemiology, and health promotion; his program was responsible for the first twenty editions of the annual *Health, United States*, and he was intimately involved in numerous federal health data systems.

ABSTRACT: Major increases in the resources devoted to the collection of health-related data and advances in survey methodology may be offset by more nonresponse and coverage bias resulting from privacy concerns, technological changes, and an increasingly complex health care environment. Hence, it is unclear whether policymakers today are basing their decisions on data that are of higher or even the same quality as those collected twenty-five years ago. We offer several recommendations for improving data quality, including changes related to Office of Management and Budget review, broad reexamination of the federal health survey portfolio, and greater investment in survey methods research. [*Health Affairs* 26, no. 6 (2007): 1599–1611; 10.1377/hlthaff.26.6.1599]

SURVEYS EXAMINING THE U.S. POPULATION'S experiences with the health care system have been conducted for more than seventy-five years, with especially large investments made since the late 1970s. As more public funds are directed toward health care, the need to understand how those dollars are spent becomes more critical, as do the consequences of inaccurate data. Over the past twenty-five years, however, the growing demand for health-related data, coupled with new technologies, has precipitated changes in data collection procedures while also limiting data producers' ability to develop new methods to keep pace with changing analytical needs. Maintaining data quality has also been made more challenging by a host of other factors outside the control of survey designers, such as the growing complexity of health care delivery, increasing concerns about privacy, and changing attitudes toward government. Although there have been major increases in the level of resources devoted to the collection of health-related data as well as important advances in survey methodology, these may be offset by more nonresponse and coverage bias resulting from privacy concerns, technological changes, and an increasingly complex health care environment. Hence, it is unclear whether today's policymakers are basing their decisions on data that are of better or even the same quality as those collected a quarter-century ago.

Health Surveys: The Early Years

■ **The NHIS and before.** The first attempt to create a national survey to support analyses of health care access and spending was instituted in the late 1920s, followed in 1935 by the Public Health Service's development and implementation of the National Health Survey.¹ In July 1957, the nation's oldest ongoing national health survey—the National Health Interview Survey (NHIS)—began to collect a wide range of data on demographic characteristics, health status and disability indicators, and use of physician and hospital services. By this time, the use of probability sampling was well established and served as the basis for the NHIS, supporting estimates of the national population as well as key demographic groups.

■ **Adding data on health spending.** Although the NHIS has traditionally provided important information about the state of the nation's health, it did not include data on health spending. By the 1950s, serious debate was emerging about a greater federal role in financing health insurance for vulnerable populations, including the

poor and elderly. A 1953 survey, conducted by the Health Information Foundation and the National Opinion Research Center (NORC), provided the first relevant data for researchers to examine the topic and marked “the beginning of refinement in social survey methodology for use and expenditures for health services.”²

The need for such data became more evident in the 1960s, when Medicare and Medicaid legislation was debated and eventually passed. In the absence of relevant data, the costs associated with these programs could not be examined; nor could the need for the program’s benefits be assessed. The federal government, however, was slow to react, and data collection for a 1963 NORC survey focusing on health spending was funded by foundations; NORC’s 1971 spending survey was the first to be sponsored by the federal government.

Building A More Systematic Foundation For Data Collection

■ **Development of NMCES.** It was not until the early 1970s, when most experts believed that some sort of national health insurance would be implemented, that the need for collection of spending data grew into the development of what became the 1977 National Medical Care Expenditure Survey (NMCES). An investigation specifically designed to inform the development of NMCES was the first major effort to undertake wide-ranging methodological work focused on collecting health spending data. Findings from this methods work confirmed that the new study would be complex and costly, calling for major reliance on in-person rather than telephone interviewing and also demonstrating the need to supplement household reporting of expenditures with data obtained from respondents’ medical care providers.³

As NMCES was pretested, the need for methodological innovations to support the collection of high-quality data became even more evident. The original contract for \$6.5 million was modified, and the survey eventually cost more than \$22 million. The additional cost reflected enhancements to support more-accurate data, including (1) an expanded provider verification component to augment household reporting of use and spending, (2) a separate survey of providers identified by household respondents to collect data on practice arrangements, and (3) a survey of employers and insurers to gather information on households’ health coverage.

■ **Expansion into NMES.** By the time data from the 1977 survey were processed, the push for national health insurance had slowed; instead, cuts in federal spending on health care were envisioned. The value of the data, however, did not diminish: NMCES demonstrated that the collection of good data is important regardless of the political agenda. NMCES was used in dozens of important studies that examined tax policies, as well as access to and use of care; and spending for the poor, the elderly, minorities, people living in physician shortage areas, and people in poor health. The Reagan administration supported NMCES and approved funding for an expanded study—the slightly renamed National Medical Expenditure Survey (NMES), implemented in 1987.⁴

Explosion In Health Data Initiatives

Throughout the 1980s and 1990s, new health-related data initiatives flourished. Along with the recognition of the need to learn more about the workings of the health care sector, there was a shift in the emphases for data collection. The NMCES effort stressed solidifying the underlying methodological approaches to collecting data on health spending, use, access, and insurance; by the late 1980s there was a broader, more far-reaching focus on collecting these data through multiple vehicles, from a growing number of groups, and with increasing frequency.

■ **SIPP: income and employment surveys.** One of the first of the new wave of surveys was the Survey of Income and Program Participation (SIPP), which, during 1984–1993, interviewed a panel of respondents to collect information on their income, labor, program participation, eligibility, and demographics. As health insurance was recognized as an important employment benefit, items were added to the income supplement of the Current Population Survey (CPS) to collect information on the previous year's health insurance status.⁵

■ **Expanded NMES.** The mid-1980s also witnessed the planning for the 1987 NMES, with several expansions from the 1977 design, including a component to gather data from residents of nursing and personal care homes and facilities for the mentally retarded and an independent household survey of Native Americans and Alaska Natives living on or near reservations. As part of the household component, information was also collected from Medicare beneficiaries, to allow linkage of survey data to Medicare claims records.⁶

■ **Medicare's own survey.** The expanded focus of NMES on residents of nursing homes and the linkage of survey data to Medicare administrative data might have signaled a wider recognition of the potential value of studying beneficiaries' health care. In 1991, HCFA (now the Centers for Medicare and Medicaid Services, or CMS) decided to field its own survey—the Medicare Current Beneficiary Survey (MCBS)—collecting a range of data items similar to those included in NMES and linking the household survey data to Medicare administrative claims records. The survey was designed as a rotating panel, allowing for individuals to be followed for two and a half years and supporting the addition of topical modules. This signaled recognition of the importance of accommodating emerging issues, although the lead time required to add a module was still long.

■ **Private-sector surveys.** The growth of federal survey initiatives continued unabated into the 1990s. Nonetheless, the private sector saw a need for additional data collection, ushering in two of the largest foundation-funded health-related surveys: the Community Tracking Study (funded by the Robert Wood Johnson Foundation) and the National Survey of America's Families (sponsored by the Annie E. Casey Foundation). The former targeted the local health care structure of twelve specific communities, and the latter was intended to allow examination of the impact of welfare reform at the state level. Yet some of the resulting analyses seemed to duplicate those done with better-funded federal initiatives.

■ **Federal survey expansion.** Perhaps the most important expansion in the 1990s was the morphing of NMES into the Medical Expenditure Panel Survey (MEPS), stepping up the frequency of the survey from once every ten years to continuous data collection. Serious consideration also was given to an augmentation of sample size that would support state estimates, but the cost implications were so high that the ability to make state estimates has been limited to the largest states.

Post-Explosion Fallout: Too Much Of A Good Thing?

■ **From in-person to telephone surveys.** The 1980s and 1990s were the golden years for those who believed in the value of health survey data. However, rapid growth brought difficult challenges. As costs of surveys rose, there was a general movement from in-person to telephone interviewing. Although largely adopted among foundation-funded or other private-sector efforts and generally resisted in the federal surveys, this change has had a far-reaching impact on data collection.

■ **Bigger—and better?** The need for political and funding support of surveys also led to longer questionnaires. Most major surveys attempted to develop both technical and financial constituencies to support a survey's continuation and expansion. Incorporating the advice and interests of additional sponsors and advisers usually resulted in recommendations or mandates to increase sample sizes and augment substantive content, but rarely in efforts to eliminate survey components (or to eliminate a survey altogether). Decisions to expand survey content were often made with the view that the marginal cost of collecting additional data items was almost zero, without a full understanding of the implications for data quality.

The 1987 NMES—designed after careful consultation with experts from other federal agencies—is an example of the “too many cooks” approach to questionnaire design. A comparison of the 1977 and 1987 surveys showed that the only major deletion between the two efforts was the elimination of data on waiting time for each reported physician visit, with a more global question on waiting time (applying to all visits) used in its place.⁷ There were, however, many expansions to the content of the household survey—questions on health opinions, more detailed data on employment and health status, questions on informal caregivers, and more information about family structure. With respect to the NHIS, although it is difficult to measure changes in questionnaire administration time because of the restructuring of the survey and the transfer of items from the core questionnaire to supplemental questionnaires, in 1990 it was noted that “total administration time for the supplements has increased considerably...to the point where they now account for a substantial part of the total interview time.”⁸

Although distinct from the surveys focused on here, a notable example of the relationship between questionnaire length and data quality was the 1996 CMS-funded physician practice cost survey. In attempting to obtain an enormous amount of detailed data on practice costs from physicians, the survey clocked in at more than three hours per case; with little to no consideration of the implications

for respondents' burden and response rates, the disastrous results were a response rate of 27 percent and a major survey effort cancelled, ending with few useful data and a public relations fiasco in the medical community.

Not only in situations such as the practice cost survey are response rates declining, but they appear to be falling regardless of interview mode or type of effort. The 1977 NMCES had a final response rate of 82 percent, while the 1996 survey obtained a 71 percent full-year response and the 2004 MEPS obtained 63 percent. The NHIS rate fell from 94 percent in 1996 to 88 percent in 2005. Nevertheless, both surveys have high response rates compared to other complex surveys.

■ **Declining trust in government.** Ironically, the expansion of health surveys overlaps an era when surveys became more difficult to implement for reasons outside the control of survey designers. Public opinion polls have demonstrated a continuing decline in trust in government, which could be affecting people's decision to participate.⁹ Despite the exemption of social surveys from the "do not call" legislation, the publicity surrounding its implementation—as well as that of the Health Insurance Portability and Accountability Act (HIPAA) privacy provisions—might have raised the public's expectations and concerns about confidentiality. The rise in the use of mobile phones is an additional and growing complication that may further depress coverage rates, contributing to more serious difficulties in maintaining adequate representation in survey samples.¹⁰

■ **Changes in the health care environment.** Also, numerous changes within the health care environment have made data collection more complex. One issue is the changing way in which health care dollars are billed and reimbursed. In the 1977 survey, the term "expenditure" was relatively unambiguous. The advent of managed care and capitation in the 1980s made the concept more complicated, since there was no longer necessarily a payment tied to each service. The current environment of preferred provider networks in which discounts (and therefore "expenditures") vary by patient and by health plan makes it unlikely that patients know the total charge for the service and, more to the point, makes reimbursement not at all reflective of resource use, so that summing across patients with different coverage becomes an exercise of adding apples and oranges.

Dramatic changes have also occurred in the continuum of care available to patients and the number of different types of providers and facilities. Although patients used to be either at home, in a short-term hospital, or in a nursing home, survey sampling frames must now include assisted living facilities and other living arrangements. From a sampling perspective, these arrangements are likely to be considered community residences, not institutions; moreover, people residing there have higher-than-average health care spending, and their inclusion is thus of importance. Surveys must also collect information about patients' use of a growing number of types of health care services (for example, home health, acupuncture, imaging) delivered in an array of outpatient settings such as urgent care centers, big-box retail health care outlets, and adult day care.

Assessing The Damage: Impact On Data Collection And Quality

■ **Overlapping surveys.** In this era of growth, survey initiatives began to compete for limited funds, and there was considerable overlap if not outright duplication. Federal agencies involved in data collection worked to create better integration across different efforts, and some important progress was made. Survey integration was an attempt to rationalize multiple related survey efforts, reduce redundancies, and coordinate across agencies, but it occurred with mixed results. The focus on survey integration became more evident during the Reinventing Government (REGO) initiative in 1995 intended to make government more efficient. The U.S. Department of Health and Human Services (HHS) took an extensive look at ongoing data collection activities; although there was little effort toward eliminating or reducing the size or frequency of any data collection efforts, there was a move toward survey coordination. Perhaps most importantly, an agreement was reached between the National Center for Health Statistics (NCHS) and the Agency for Healthcare Research and Quality (AHRQ) to use the NHIS as a sampling frame for MEPS, allowing greater sampling of groups of policy interest without the cost of additional screening.¹¹

■ **Declining survey methods funding.** During roughly this same period, federal spending on survey methods research—with the potential to make spending on surveys more cost-effective and improve data quality—seems to have declined. A special panel was convened in 1977 by the National Committee on Vital and Health Statistics (NCVHS). The committee made the potentially important recommendation that “not less than 10 percent of the NHIS budget be allocated to methodological and development research.”¹² This target was never reached or even approached.

Despite the committee’s support for methods work, the growth in the number and size of HHS surveys has not been accompanied by an equivalent increase in resources for methods development. By the mid-1990s it was clear that “it has been impossible to maintain even our historic level of funding for such research and our measurement tools suffer as a result.”¹³ It is hard to estimate funds devoted to methods research, since they are not always separated out in agency budgets. However, at a 2004 major data conference, a general consensus was that “research on survey methodology has not received sufficient funding priority.”¹⁴

Instead of expanding methods research, the budget pressures to collect more data have curtailed spending on methods research. The presence of validation components makes MEPS a valuable resource for survey methods studies, and the MEPS staff continue to produce highly respected methods studies. However, the level of investment in actual field experiments that preceded the 1977, 1987, and 1996 surveys has not been sustained. The highly regarded NCHS Series 2 publications have sharply declined as well. Between 1971 and 1980, forty-five reports were issued as part of this series; in the ten-year period 1996–2005, that number dropped to fifteen.¹⁵

■ **Effects on data quality.** All of the changes undergone by data collection efforts—the move to telephone interviewing, longer questionnaires, lower response rates, and more data collection overall—have had perhaps unforeseen effects on data quality and perceptions of data quality.

Telephone interviewing. Although debate is still ongoing about the effect of telephone interviewing on data quality, there is cause for serious concern. Following analyses of the 1986 RWJF Access to Care Survey, researchers found that a switch from in-person sampling to a telephone frame resulted in estimates of the uninsured that were significantly lower.¹⁶ They concluded that the systematic relationship between having a phone and being uninsured was at the root of the problem. As the subsequent access survey was conducted for the RWJF, the decision was made to use the NHIS as a sampling frame, which allowed for the vast majority of the interviews to be conducted through less costly telephone interviews. But since the sample itself was a list of households, in-person interviews could be conducted with people known not to have telephones. Based on knowledge at that time, it was expected that this would eliminate any bias resulting from lack of phone coverage.

Resulting estimates of the uninsured from the 1994 survey, however, when compared to other sources, were higher than those from phone surveys but still lower than those from household surveys, which led to the conclusion that some additional phenomenon related to telephone interviewing was affecting the data collected. Although hypotheses have been put forward as to what may account for this other telephone effect, there has been no direct study and no hard evidence.

In fielding the NSAF, researchers decided they would avoid the phone coverage issue by providing cell phones to people without telephones; in addition, a final question was added to the insurance sequence to confirm uninsured status. Despite these refinements, estimates of the uninsured from the NSAF were still lower than those obtained from in-person surveys during the same period.

Administration time. In addition to increasing the burden on respondents, increases in survey administration time can have deleterious (and generally unforeseen) implications for the quality of the resulting data. In a study of household reports of medical visits, researchers at AHRQ found that the number of physician visits declined over the course of multiple interviews in a way that was not accounted for by any secular trend or seasonality.¹⁷ One possible conclusion is that NMCES (or NMES) respondents simply learned over time that if they reported another visit, they would be asked a series of additional questions on reasons for the visit, sources of payment for the visit, and so on; this is referred to as a “conditioning” effect. Facing these additional questions, respondents may have decided that it was easier to avoid reporting some of the care they received.

A similar effect was noted in an examination of household reporting of problems with access to care.¹⁸ When results from six surveys using the same basic access question were compared, researchers found an inverse relationship between

the length of the questionnaire and the proportion of the uninsured reporting an access problem. Here again, one might infer that respondents figured out the consequences of a “yes” response and chose the easier path of responding “no.”

“*Dueling data.*” Another outcome of the multitude of surveys is the issue of “dueling data”: when different surveys produce widely differing estimates of key variables related to access to care, insurance coverage, or other matters of interest, or when survey estimates are inconsistent with other data sources (for example, survey-produced spending estimates versus data from the National Health Accounts). There are legitimate reasons why different estimates may be obtained, and survey methodologists can point to many design factors including question wording, mode of interview, or reference period that might explain these differences.¹⁹ Although these explanations provide potentially satisfactory answers to methodologists, they are of little value to a policy expert who needs the “right” answer to examine the potential cost or other impact of new initiatives. Since ultimately it is policymakers and not researchers who decide the level of investment made in health survey research, these inconsistent estimates represent an obstacle to advocates of survey research.

One of the most troubling and visible examples of conflicting estimates in the health arena has to do with estimates of the number of uninsured—where six major surveys serve as sources of information on the number of uninsured Americans, and each provides a different figure. Because this issue has been so widely explored and written about, we do not belabor the point further but merely reiterate that the resulting confusion to policymakers likely lessens the role of data in decision making instead of increasing it.²⁰ This same phenomenon of conflicting estimates was found in the previously mentioned study comparing estimates of access to care from six national surveys.²¹ The range of estimates, among the uninsured, of those unable to obtain care, was substantial—at least a sevenfold differential—leaving the message to policymakers muddled.

Getting Back On Track: Directions For The Future

The issues noted above—more and broader data collection efforts, declining response rates, conflicting estimates—are of immediate concern to those who rely on high-quality data to support the policy-making process. Unfortunately, enumerating these problems is easier than solving them. Most of the issues discussed are well known to the leadership of HHS agencies: They want more methodological research, worry about response rates, and prefer that the data being collected not duplicate other efforts. In some ways, the survey designers are victims of their own success; as policymakers understand the value of survey data in assessing policy changes, growing demands for data force agency budgets to emphasize short-term efforts while postponing longer-term investments focused on data quality.

Improving the quality of federal health surveys is imperative if survey findings

are to be of continued value to policymakers, but changes will be neither fast nor inexpensive. Here we offer a few suggestions as small steps to move the process forward, focusing on the importance of reduced bureaucratic obstacles, survey integration, and, perhaps most important, regular and ongoing investment in the development and refinement of methodological approaches to collecting high-quality data in an increasingly difficult environment.

■ **Reduce bureaucratic obstacles.** The role of the Office of Management and Budget (OMB), as set forth in the Paperwork Reduction Act, is to provide oversight so that respondents are not burdened unnecessarily and taxpayers are not paying for unnecessary data collection. In practice, however, the review process often involves layers of review that rarely result in much change to surveys yet may adversely affect the timeliness of the data collected.

Another OMB issue, with perhaps greater implications for survey quality, is the low threshold of nine participants that makes a data collection subject to a complex review and public comment period. Based on this requirement, pre-testing—and therefore the assessment of critical design features—is often limited to only nine cases. Allowing the administrator of any agency involved in data collection to approve a pretest of up to fifty cases would offer many benefits, potentially including learning that certain questions don't work and should be eliminated, which would ultimately reduce the burden on respondents.

A careful assessment of the OMB process is called for—one that would systematically identify the extent of changes made subsequent to OMB review and how these modifications have affected the overall data collection effort. A major issue for such a review is whether the delays in survey implementation resulting from the approval process have brought with them corresponding reductions in respondent burden and the overall resources required for data collection. Depending upon the findings of such a study, appropriate modifications to the OMB's practices in survey review should be made to address burden on respondents and the timeliness of data in a more balanced manner.

■ **Integrate surveys.** Tackling survey integration is less straightforward, in part because it may be defined in different ways. Here we focus briefly on two issues—one narrow and one broad—that may represent the continuum of how survey integration is envisioned. The first issue is whether analytical concepts should be measured similarly across surveys. We believe that there is no general rule and that this should be approached case by case. Variables such as “access” or “quality” can be conceptualized and therefore measured in a variety of legitimate ways. Standardization would limit researchers' ability to study the different components of access or quality. Point-in-time health insurance, however, is not subject to different theoretical approaches: All surveys that collect this information are trying to estimate the number of people who lack coverage at a specific time. Differences in the estimates of insurance across surveys are more difficult to justify, and the standardization of questions appears to have clear advantages. Standardization of the approach to ask-

ing about current health insurance coverage, however, should not preclude efforts to collect information on lack of coverage over other time periods, as well. An approach to moving forward on standardization is offered by the Census Bureau; researchers there, faced with similar inconsistencies in poverty and income statistics, have conducted extensive analyses and fostered discussions on the potential for combining information from multiple data sources to an official single series of estimates.

The second issue relates to the kind of consolidation anticipated by REGO. In fact, few surveys or even sections of surveys have actually been eliminated through integration efforts. Although this might have been due in part to turf issues among agencies, there are legitimate reasons for separate efforts in many cases. One survey may focus on health status data, while still collecting data on health care use as an independent variable, and another survey might focus on gathering health spending data, while collecting data on health status as a contributing factor. Although these decisions might be sensible, the result is that there are two or more surveys essentially collecting many of the same data elements. The goal should not be so much to restrict any given survey from including a data element common to another survey, but to conduct a very high-level examination of a large number of somewhat related surveys to see if a major restructuring and redesign effort could produce a better trade-off between the amount of data collected and the overall level of resources expended. These surveys have largely grown in an ad hoc fashion—perhaps a more systematic design effort starting at ground zero could produce a more efficient overall effort. This suggestion is consistent with the Government Accountability Office's (GAO's) recommendation that the OMB director work with the Interagency Council on Statistical Policy to comprehensively reexamine the government's survey portfolio.²²

Several positive outcomes related to survey integration can be noted. The NHIS now serves as the sampling frame for MEPS, thereby eliminating the need to conduct an extra round of screening to determine eligibility. Similarly, the NCHS has used its National Immunization Survey to create the State and Local Area Integrated Telephone Survey (SLAITS), which has permitted national surveys of rare populations that would otherwise be too costly to field. These successes suggest that despite inherent obstacles, bold thinking about survey integration should continue to be encouraged.

The recent experience of SIPP—the process used to eliminate the survey effort as well as to design a replacement that would gather the most essential data in a more cost-effective manner—might prove to be instructive as a model. Instead of a series of minor modifications, the decision was made to completely overhaul the survey by making a major investment in evaluation studies as well as by conducting numerous forums with research and policy constituencies to build a new and better mousetrap. If the reconstituted endeavor ultimately proves to be superior to SIPP, then the statistical community should carefully examine the process used.

■ **Invest in methods research.** With respect to survey evaluation, although the

1977 NCVHS resolution advocating that 10 percent of the NHIS budget be spent on development and methods might be arbitrary, more resources need to be devoted to exploring methods and assessing the quality of the data collected. Methods work should go beyond response rates in measuring survey quality; it should be rigorous and incorporate studies that look at interviewer and respondent error. Survey response rates are often used as a rule-of-thumb measure of survey quality; however, this is not because of any consensus that nonresponse is the major source of survey error but rather because it is the easiest measure to obtain. Cognitive interviews can be useful in designing questionnaires but are limited by respondents' ability to report accurately. Because of the well-known limitations of household reporting, validation studies are critical. If the respondent doesn't know or doesn't wish to reveal the answer, even the best of questions will not result in good data.

Given the low federal investments in methods, foundations could play an important strategic role by sponsoring survey evaluations that the government might not be able or willing to fund. In recent years, foundations have sponsored very large survey efforts, playing a role similar to the one they played in the era prior to the expanded government initiatives. Although such surveys have been useful, foundations are in a unique position to serve as a catalyst for better survey development. Moderate foundation investments would likely be followed by more federal work to improve on issues raised by the privately funded efforts.

THE SUGGESTIONS WE HAVE MADE IN THIS PAPER—while hardly radical—will not be easy to implement, nor will they be sufficient. But the need for reliable and timely survey data will remain essential to the development of sound social policy, even as a changing health care environment continues to mount new obstacles to the design and implementation of high-quality surveys.

.....
The opinions expressed are those of the authors and are not held by NORC. The authors thank Jyoti Gupta of NORC for her research assistance.

NOTES

1. I.S. Falk, M.C. Klem, and N. Sinai, *The Incidence of Illness and the Receipt and Costs of Medical Care among Representative Family Groups* (Chicago: University of Chicago Press, 1933).
2. O.W. Anderson, "Foreword," in *Total Survey Error: Applications to Improve Health Surveys*, ed. R. Andersen et al. (San Francisco: Jossey-Bass, 1979).
3. Health Services Research and Development Center, Johns Hopkins Medical Institutions, *Medical Economics Survey—Methods Study*, Final Report Conducted under Contract no. HRA 106-74-150 (Baltimore: Johns Hopkins, August 1977).
4. In the intervening period between NMCES and NMES, the NCHS implemented the National Medical Care Use and Expenditures Survey (NMCUES), which combined an area probability sample, a list sample from Medicaid files in four states, and an administrative records component to obtain information on program eligibility and payments for Medicare for people enrolled in Medicare and Medicaid.
5. U.S. Census Bureau, "Health Insurance," 28 June 2007, <http://www.census.gov/hhes/www/hlthins/hlthins.html> (accessed 2 August 2007).
6. As in 1977, there was also a Medical Provider Component and a Health Insurance Plan Survey. In addition, poor and low-income families and blacks and Hispanics were oversampled. Centers for Disease Control

- and Prevention, *Scientific Data Documentation: Institutional Population Component, Baseline Data, 1987*, http://wonder.cdc.gov/wonder/sci_data/surveys/nmes/type_txt/nmesipc.asp (accessed 2 August 2007).
7. J. Kasper and M. Berk, "Sociological Gains and Losses: The Case of the National Health Care Use and Expenditure Surveys," *American Sociologist* 19, no. 3 (1988): 232–242.
 8. D. Judkins, J. Waksberg, and D. Northrup, "Cost Functions for NHIS and Implications for Survey Design," *Proceedings of the Survey Research Methods Section, American Statistical Association*, 1990, http://www.amstat.org/sections/SRMS/proceedings/papers/1990_006.pdf (accessed 6 August 2007).
 9. According to a study conducted in 1996 by the *Washington Post*, Henry J. Kaiser Family Foundation (KFF), and Harvard's Kennedy School of Government, only one-quarter of Americans surveyed trusted the federal government, down from three-quarters in 1964. Additionally, in 1964, 54 percent of respondents said that they trusted the government to do the right thing, while in 1995 that figure dropped to 35 percent. A study conducted in 2000 by National Public Radio, the KFF, and the Kennedy School found that one-quarter of Americans surveyed believed that the federal government is a major threat, and half believed that it is at least some threat, to their personal rights and freedom; 60 percent believed that the government does what is right only some of the time. See "National Public Radio/Kaiser Family Foundation/Kennedy School of Government Survey on Americans' Attitudes toward Government," 2000, <http://www.kff.org/kaiserpolls/3036-index.cfm> (accessed 6 August 2007). There is some racial and ethnic variation in distrust for the government: Latinos are the least likely to believe that the federal government is a threat to personal rights and freedom, whites are the most likely to believe this, and African Americans and blacks are the most likely to believe that state and local governments are a threat to personal rights and freedom. See R. Morin and D. Balz, "Reality Check: The Politics of Mistrust," 28 January 1996, <http://hackvan.com/pub/stig/etext/politics-of-mistrust/1-trust-nobody.html> (accessed 2 August 2007).
 10. "During the last 6 months of 2006, more than 3 of every 20 American homes (15.8%) did not have a landline telephone. Of those homes without a landline telephone, most had at least one working wireless telephone." S.J. Blumberg and J.V. Luke: "Wireless Substitution: Early Release of Estimates Based on Data from the National Health Interview Survey, July–December 2006" (Hyattsville, Md.: National Center for Health Statistics, 14 May 2007), 1.
 11. For a discussion of REGO and survey integration, see R.H. Arnett III, "Survey Integration and End-of-Century Constraints," in *Informing American Health Care Policy: The Dynamics of Medical Expenditure and Insurance Surveys, 1977–1996*, ed. A.C. Monheit, R. Wilson, and R.H. Arnett (San Francisco: Jossey-Bass, 1999), 201–212.
 12. B.C. Greenberg, "The National Health Interview Survey—Recommendations by a Technical Consultant Panel," in *NCHSR Research Proceedings Series: Health Survey Research Methods*, Pub. no. (PHS)81-3268 (Washington: U.S. Department of Health and Human Services, 16–18 May 1979), 18–23.
 13. E.J. Sondik and E.L. Hunter, "National Data Experts on Access Measures," *Health Affairs* 17, no. 1 (1998): 189–190.
 14. AcademyHealth, *Improving Federal Health Data for Coverage and Access Policy Development Needs* (Washington: AcademyHealth, October 2004).
 15. A list of all NCHS Series 2 publications (Data Evaluation and Methods Research) is available at NCHS, "Series 2. Data Evaluation and Methods Research," 3 August 2007, <http://www.cdc.gov/nchs/products/pubs/pubd/series/sr02/ser2.htm> (accessed 6 August 2007).
 16. C.R. Corey and H.E. Freeman, "Use of Telephone Interviewing in Health Care Research," *Health Services Research* 25, no. 1, Part 1 (1990): 129–144.
 17. S.B. Cohen and V.L. Burt, "Data Collection Frequency Effect in the National Medical Care Expenditure Survey," *Journal of Economic and Social Measurement* 13, no. 2 (1985): 125–151.
 18. M.L. Berk and C.L. Schur, "Measuring Access to Care: Improving Information for Policymakers," *Health Affairs* 17, no. 1 (1998): 180–186.
 19. *Ibid.*
 20. HHS, Office of the Assistant Secretary for Planning and Evaluation, "Understanding Estimates of the Uninsured: Putting the Differences in Context," ASPE Issue Brief, September 2005, <http://aspe.hhs.gov/health/reports/hiestimates.htm> (accessed 17 September 2007).
 21. Berk and Schur, "Measuring Access to Care."
 22. U.S. Government Accountability Office, *Federal Information Collection: A Reexamination of the Portfolio of Major Federal Household Surveys Is Needed*, Pub. no. GAO-07-62 (Washington: GAO, November 2006).