

# Sample and design considerations in post-disaster mental health needs assessment tracking surveys

RONALD C. KESSLER,<sup>1</sup> TERENCE M. KEANE,<sup>2,†</sup> ROBERT J. URSANO,<sup>3,†</sup> ALI MOKDAD,<sup>4</sup>  
ALAN M. ZASLAVSKY<sup>1</sup>

1 Department of Health Care Policy, Harvard Medical School, Boston, MA, USA

2 National Center for PTSD, VA Boston Healthcare System, Boston, MA, USA

3 Department of Psychiatry and Center for the Study of Traumatic Stress, Uniformed Services University, Bethesda, MD, USA

4 Institute for Health Metrics and Evaluation, University of Washington, Seattle, WA, USA

---

## Abstract

*Although needs assessment surveys are carried out after many large natural and man-made disasters, synthesis of findings across these surveys and disaster situations about patterns and correlates of need is hampered by inconsistencies in study designs and measures. Recognizing this problem, the US Substance Abuse and Mental Health Services Administration (SAMHSA) assembled a task force in 2004 to develop a model study design and interview schedule for use in post-disaster needs assessment surveys. The US National Institute of Mental Health subsequently approved a plan to establish a center to implement post-disaster mental health needs assessment surveys in the future using an integrated series of measures and designs of the sort proposed by the SAMHSA task force. A wide range of measurement, design, and analysis issues will arise in developing this center. Given that the least widely discussed of these issues concerns study design, the current report focuses on the most important sampling and design issues proposed for this center based on our experiences with the SAMHSA task force, subsequent Katrina surveys, and earlier work in other disaster situations. Copyright © 2008 John Wiley & Sons, Ltd.*

**Key words:** disaster, epidemiology, needs assessment survey, PTSD

---

## Introduction

Although mental health needs assessment surveys are carried out after many large natural (Ironson et al., 1997; Kohn et al., 2005) and man-made (Gidron, 2002; North et al., 2004) disasters, synthesis of findings about patterns and correlates of post-disaster psychopathology is hampered by inconsistencies in study design and measures (Brewin et al., 2000; Galea et al., 2005; Norris, 2005). Recognizing this problem, the US Substance Abuse and Mental Health Services Administration

(SAMHSA) in 2004 assembled a task force to develop a model study design and interview schedule for use in post-disaster mental health needs assessment surveys. It was thought that such a protocol would both lead to greater consistency than currently exists across such surveys and reduce the sometimes substantial delays due to instrument development that occur in launching these surveys.

The interview schedule developed by this task force was pre-tested among victims of the Florida hurricanes of 2004. After revisions based on the results of cognitive interviews carried out with these pre-test respondents, a pilot survey of the revised interview schedule was carried out in late 2005 in three samples of people

---

<sup>†</sup>The contributions of Terence M. Keane and Robert J. Ursano were prepared as part of their official duties as US Government employees.

who were exposed to a natural or man-made disaster in the previous two years (a train crash and resulting toxic chemical spill in a small town in South Carolina; a plant explosion in a small town in Illinois; and a series of tornados in several small towns in the Midwest). Further instrument revisions were made based on debriefing interviews with a sub-sample of the respondents in this pilot survey, a clinical validation study of the post-traumatic stress disorder (PTSD) screening questions in the pilot survey, and quantitative analyses of survey responses to confirm that the interview schedule generated substantively plausible results.

An expanded version of this revised instrument was then used in a series of mental health needs assessment tracking surveys among victims of Hurricane Katrina in the US Gulf Coast (Kessler et al., 2006). These surveys posed a number of sampling and design challenges related to the special circumstances of Hurricane Katrina that are discussed later. They also highlighted the fact that the subsequently much-discussed deficiencies in federal disaster preparedness apply as much to disaster needs assessment surveys as to other areas of disaster response. Based on this realization, the US National Institute of Mental Health (NIMH) has established a center that will implement post-disaster mental health needs assessment surveys in the future using an integrated series of measures and designs. A wide range of measurement, design, and analysis issues will arise in implementing these surveys. Given that the least widely discussed of these issues concerns study design, the current paper focuses on the most important sampling and design issues likely to be faced by this center based on our experiences with the SAMHSA task force, subsequent Katrina surveys, and earlier work in other disaster situations. The issues considered are those that apply to surveys carried out in the US and, by extension, in other developed countries. Many of the considerations discussed here would be rather different in less developed countries.

### Challenges and opportunities

The major challenges in designing disaster-related needs assessment surveys concern implementation. With regard to sampling, it is usually necessary to create an appropriate sampling frame very quickly so that survey results can be used to make timely planning decisions. A complicating factor in many disaster situations even in developed countries is that infrastructure damage creates logistical problems that hamper

implementation of conventional telephone surveys and that impedes the travel of field interviewers to carry out face-to-face surveys. In the case of Hurricane Katrina, there was the additional complication that a massive flood led to the evacuation and wide geographic dispersion of the population of New Orleans.

The existence of a center for disaster surveys will create opportunities to address these practical challenges in the US as well as to expand the conventional role of needs assessment surveys by developing ongoing collaborations with government disaster-preparedness agencies and relief agencies. Several such opportunities for expansion exist. For example, the US federal government uses the mass media to disseminate information aimed at increasing knowledge and changing attitudes and behavior (KAB) of populations both before disasters (i.e. disaster preparedness) and after disasters (i.e. disaster response). Needs assessment tracking surveys can be used to provide feedback to the message development teams involved in these KAB social marketing public information campaigns (Flay et al., 1989). This kind of collaboration would require coordination, as the survey team needs to be aware of the messages being disseminated by the message development team in order to build relevant questions about these messages into the needs assessment surveys.

A good example of such coordination is the current collaboration between our Hurricane Katrina Community Advisory Group (CAG; [www.HurricaneKatrina.med.harvard.edu](http://www.HurricaneKatrina.med.harvard.edu)) and the American Red Cross (ARC) in tracking awareness and response to the new ARC Access to Care (ATC) Program, a program designed to help low-income victims of Hurricane Katrina pay for emotional support services, such as mental health treatment and substance abuse treatment. The ongoing CAG tracking surveys are monitoring awareness of the ATC Program, attitudes about the program, and barriers to taking advantage of the program. Analysis of the CAG data is providing information to the ARC about population segments with low awareness of the ATC Program, media habits of these population segments that might be useful in developing new program dissemination strategies, and information about barriers to using the program among eligible community members who are aware of the program but have not used it to pinpoint potentially useful expansions of ARC outreach efforts.

This example is a rather obvious one, as public health marketing campaigns often use market tracking

surveys to evaluate campaign success (e.g. Subar et al., 1995). The only innovation is the use of mental health needs assessment surveys to take on the conventional role of market tracking surveys. Other opportunities to expand the conventional role of needs assessment surveys are less obvious, though, although equally important. A number of these are discussed later in the section on design consideration. Before this, though, we turn to the important matter of sampling.

### Sampling considerations

#### *General population sampling*

The difficulties associated with selecting a representative sample of disaster survivors differ depending on whether the disaster is or is not defined in terms of geography. In the case of natural disasters (e.g. tornadoes, hurricanes) or man-made disasters that have a geographic epicenter (e.g. the Oklahoma City bombing), it makes most sense to think in terms of area probability household sampling as the main basis for sample selection. There are inevitable practical problems with this form of sampling that can be exacerbated in situations of mass evacuation. As described later; multiple-frame sampling (Skinner and Rao, 1996) can be used to decrease coverage problems in situations of this sort. In the case of disasters that do not have a geographic epicenter (e.g. a plane crash), in comparison, the use of list samples is a necessity unless the researchers have the resources to engage in large-scale mass screening, using multiplicity sampling (Kalton and Anderson, 1986) whenever possible to increase the efficiency of the screening exercise. In any of these cases, frame biases have to be taken into consideration. Land line telephone frames, in particular, might under-represent the most disadvantaged segments of the population (Brick et al., 2006), making it particularly useful to implement a multiple-frame sampling approach that enriches the less restrictive frame for high-risk cases, possibly by over-sampling Census blocks with low rates of land line telephone penetration or high rates of poverty.

Some studies will involve both geographically clustered and dispersed cases. For example, the workers in a government building exposed to a terrorist attack with anthrax would be geographically dispersed during the initial time period when the building was evacuated and workers were sent home prior to a thorough evaluation of building contamination. The most feasible way

to evaluate need for mental health treatment of these workers and their families during that time period would be from an administrative list sampling frame with home contact information for all such workers. Once the Environmental Protection Agency makes an evaluation that the building is safe for workers to return, though, the affected workers (although not their families) would become highly clustered geographically (i.e. at their place of work), making it possible efficiently to carry out mental health needs assessment surveys on site.

Another mixed case is the situation where a man-made disaster occurs at a place that involves both residents of the area in which the disaster occurred and people who were passing through the area at the time of the disaster. A good example is the 2005 train crash at a depot in the middle of the small town of Graniteville, South Carolina that released toxic chemicals into the local environment, leading to injury, death, and toxic exposure among the passengers and crew of the train and to risk of toxic exposure, evacuation, and community disruption among residents of the community in which the crash occurred (US Environmental Protection Agency, 2005). In a situation of this sort, the residents of the community would be geographically clustered while the surviving passengers and crew of the train would not be geographically clustered.

We faced an especially complex situation with regard to sampling in assembling the Hurricane Katrina CAG. A small proportion of the population, presumably representing the most high-risk pre-hurricane residents of the areas most hard hit by the storm and resulting flood in New Orleans, were living in evacuation centers (ECs) and later Federal Emergency Management Agency (FEMA)-sponsored hotel rooms, trailers, and even luxury liners. Many other pre-hurricane residents of the New Orleans Metropolitan Area were scattered throughout the country, largely living with relatives, but also in communities that had established ECs and subsequently created community living situations in which a certain number of needy families from New Orleans were, in effect, adopted by the community. The vast majority of pre-hurricane residents of the other areas in Alabama, Louisiana, and Mississippi that were affected by the hurricane remained living either in their pre-hurricane households or in the surrounding community in which they lived before the hurricane as they went about repairing the damage caused to their homes and communities. Telephone lines were down

in many parts of the affected areas for a considerably longer time than is typical in US natural disasters. In addition, physical movement was made difficult by infrastructure damage and difficulty finding gasoline for cars. Conventional household enumeration was made difficult in some areas by the fact that many pre-hurricane homes no longer existed.

At the same time, we had several important resources available to us that we used in building a multiple-frame sampling strategy that combined information from a number of restricted frames to assemble the sample of people who participated in the CAG. One rather unexpected resource was the use of random digit dialing (RDD). It seems counterintuitive that RDD could be used to study Katrina survivors in light of the fact that the vast majority of the New Orleans population was forced to evacuate their homes after the storm and the fact that many people who lived in other areas affected by the hurricane had non-working land lines because of damage to telephone infrastructure. However, the main telephone provider in the hurricane area, Bell South, forwarded phone calls made into the hurricane area to new numbers (either land line numbers or cell phone numbers) outside the area that were registered by the owners of the pre-hurricane numbers. As a result of this service, we were able to call an RDD sample of phone numbers selected from 1+ telephone banks working in New Orleans prior to the hurricane and to connect with many displaced pre-hurricane New Orleans residents in temporary residences all across the country.

A second important resource was the availability of extensive ARC and FEMA lists of people who registered for assistance. Of the over four million adult residents of the area defined by FEMA as affected by Katrina (4 137 000 adult residents in the 2000 Census), a majority applied for assistance to one or both of the two major agencies that maintained comprehensive applicant lists. We were in the fortunate position of having access to both of these lists. In order to reduce overlap with the RDD frame, we restricted our use of these lists to cell phone exchanges and to land line exchanges in areas outside of the RDD sampling area. Over 1.4 million families representing more than 2.3 million adults applied to the ARC for assistance and provided post-hurricane contact information that included new residential addresses, telephone numbers (often cell phones), and email addresses. An even larger number of families (roughly 2.4 million) applied to

FEMA for assistance and also provided post-hurricane contact information comparable to the ARC list information. As one would predict, considerable overlap existed in the entries on these two lists, but the more surprising finding was that a substantial number of families applied only to one of the two. There were also a number of families that fraudulently applied on multiple occasions and at different locations to the same agency. We corrected for these multiple counts in sampling from these lists.

It is also noteworthy that a great many hurricane evacuees registered with one or more of the "safe lists" set up on the internet by the television news channels, CNN and MSNBC, the ARC, and others. These lists allowed people separated from their loved ones during the hurricane or aftermath to let it be known that they were alive and to record their whereabouts in the hopes of reconnecting with their loved ones. Google subsequently integrated all the names recorded on all the internet safe lists into a single consolidated list that contained over 400 000 names. We made extensive use of this consolidated list in piloting the baseline CAG interview. However, this pilot testing led to the discovery that virtually all people on the safe lists were also on the more inclusive ARC and FEMA lists of people who applied for assistance. As a result, we did not use the safe lists in our final sample selection for the CAG.

By the time the baseline CAG survey was fielded, all the Katrina ECs had been closed and only a small number of evacuees were still housed in FEMA-supported hotel rooms. This made it relatively easy to screen a representative sample of hotels selected from the Donnelly commercial sampling frame to find hotels housing evacuees, to use information provided by hotel managers to select a sample of rooms with probabilities proportional to size from these hotels, and to include the respondents interviewed in this way as a supplemental sample. Not surprisingly, though, this exercise showed that virtually all hotel evacuees were included with valid contact information on the FEMA relief list that we were using as one of the main sample frames. As with respondents sampled from each of the other frames, information was included about this overlap and used in making weighting adjustments in the consolidated CAG sample.

The availability of these different frames allowed us to use relatively inexpensive telephone administration to reach the great majority of people who were living

in the areas affected by Katrina before the hurricane. As noted earlier, we reduced overlap between the two main frames by restricting our use of the ARC and FEMA lists to cell phone exchanges and to land line exchanges in areas outside of the RDD sampling area. In addition, we collected data from every respondent in the entire sample that allowed us to determine whether they had a non-zero probability of selection in each frame. For example, we asked respondents in the RDD sample if they applied to the ARC and to FEMA for assistance. This information made it possible for us to use capture–recapture methods (Fisher et al., 1994) to estimate the size of each population segment defined by the multivariate profiles of their existence or non-existence in each frame and to use these estimates of size to develop weights that were used to combine these segments into an equal-probability sample of the population.

Concerns could be raised about the under-representation of three population segments in the frames discussed up to now: evacuees who lived outside the hurricane area, were reachable by RDD, but who were not included on either the ARC or FEMA lists (either because they did not apply or because they did not provide traceable telephone contact information); other evacuees who lived outside the hurricane area who could not be reached by telephone (whether or not they applied for ARC or FEMA assistance); and residents of the affected area who remained in the area but could not be contacted by telephone (because they did not have a working land line that could be reached by RDD and they either did not have a cell phone or did not apply to the ARC or FEMA and provide a cell phone contact number). We attempted to reach the first of these three groups (i.e. evacuees who lived outside the hurricane area, were reachable by RDD, but who were not included in either the ARC or FEMA lists) by experimenting with the use of a national RDD sample that employed multiplicity methods (i.e. asking for evacuees among current household residents and among first-degree relatives of a randomly selected informant in each household) either with live telephone interviewers or interactive voice response (IVR) messages with follow-up live telephone interviewers. Based on data from the ARC and FEMA lists about geographic evacuation patterns, we anticipated that approximately one in every 500 households in the US outside of the hurricane area would contain one or more hurricane evacuees and that some additional number of

household informants would tell us about the whereabouts of such evacuees.

We screened a nationally representative sample of 20 000 listed telephone numbers to investigate the validity of these assumptions, a random half using IVR and the other half using live interviewers. We found a hit rate closer to one in 1000 in the households randomized to be screened by live interviewers, with the number of evacuees in these households typically quite large (4–7). This presumably reflects differential preferences for relocation destinations of evacuees with and without families. We found that the hit rate was much smaller in the households randomized to be screened by IVR. It is possible that this disadvantage of IVR could have been corrected if we had pursued additional iterations of alternative IVR scripts. We terminated the exercise before these iterations, though, based on the finding: that all evacuees in telephone households with listed phone numbers outside the hurricane area had applied either to the ARC or to FEMA for assistance with traceable contact information. This means that these people were already part of our primary sample frames, making it unnecessary to screen for them in a supplemental national RDD sample.

The most feasible way to reach the remaining groups that are under-represented in the frames discussed earlier (i.e. evacuees who could not be reached by telephone) using probability sampling would have been to use a survey field staff to carry out face-to-face interviews on an area probability sample of households and group quarters. We did not do this in our survey of Katrina survivors due to financial constraints. If we had done so, it would have been important to include information that allowed us to determine whether each respondent sampled from this frame also had a probability of selection in the list samples and the RDD sample. With regard to design considerations, a sample of this sort that focused on people living in the area affected by the hurricane would be based on a conventional multi-stage clustered area probability sampling design.

Logistical complications would exist in sample selection, as the Census measures of size used to select sampling segments (i.e. blocks in urbanized areas and block-equivalents in rural areas) would be much less accurate than normal because of housing destruction. Block listing would also be more complex than usual in that the normal landmarks used to define sample segments would in some cases be destroyed, possibly

making it necessary to work with knowledgeable local informants (e.g. mail delivery workers) to help define segment boundaries. It might also be efficient to select larger segments than in a usual household survey to allow for the likelihood of housing unit destruction and to invest more heavily in block listing than usual. Logistical complications would also exist in interviewer travel and housing and because of infrastructure damage. While making fieldwork more difficult, though, none of these problems would be insurmountable.

An argument could be made that even non-probability sampling would be useful situations where probability sampling is prohibitively expensive so long as the sampling was based on characteristics identified as reflecting high exposure to disaster-related stressors (e.g. areas that were directly hit by a tornado or areas that were not reconnected to services after a natural disaster), as such an approach could provide useful information about the range of exposures and psychological reactions to the disaster. Quotas on the basis of a cross-classification of basic socio-demographic variables could be imposed in such a case in order to guarantee breadth of coverage.

#### *High-risk population sampling*

Initial needs assessment surveys of Hurricane Katrina survivors focused on high-risk populations, including pre-hurricane residents of New Orleans who remained in their homes shortly after the hurricane (Centers for Disease Control and Prevention, 2006a), people staying in evacuation centers (Centers for Disease Control and Prevention, 2006b), and people residing in FEMA-sponsored trailers or hotel rooms (Abramson and Garfield, 2006). First responders also are a high-risk population of importance that has been the focus of considerable research attention (Ben-Ezra et al., 2006; Fullerton et al., 2004). Although these populations make up only a small percentage of all the people who were affected by Katrina, their distinct geographic characteristics and their presumably high level of exposure to hurricane-related stressors make them important targets for needs assessment.

Such high-risk populations can be expected to vary widely across disaster situations. The workers in a government office building that was the target of an anthrax attack along with their families might be a high-risk group in one disaster situation, while the residents of a geographic area close to a toxic chemical spill might be a high-risk group in another disaster situation.

Geographic propinquity need not be a defining feature of these groups. The families and close friends of the people killed in an airplane crash, for example, would be a high-risk group for needs assessment that is widely dispersed in terms of geography. In the case of natural disasters, there are some other high-risk groups that might be expected to be more consistent across situations, such as residents of nursing homes and people with physical disabilities who would have a difficult time evacuating.

One of the most important of these high-risk groups after Hurricane Katrina consisted of people with pre-hurricane severe-persistent mental illness (SPMI) whose medical records were temporarily lost in the storm, whose local pharmacies were destroyed, and who were unable to refill their antipsychotic medications. This group represents an extreme case of the much larger group of people with pre-existing chronic conditions who were found in assessments of EC residents often to have unmet need for maintenance medications to treat their chronic conditions (Brodie et al., 2006). An exacerbating factor is that the Strategic National Stockpile of emergency medications (Centers for Disease Control and Prevention, 2005) and short-term deployments of emergency medical personnel in the Public Health Security and Bioterrorism Preparedness and Response Act (Rosenbaum, 2006) both failed to anticipate this problem by providing ready access to desperately-needed medications for SPMI and other extreme chronic conditions. Once the problem was recognized, emergency mental health service planners made special efforts to obtain psychotropic medications for emergency medical clinics as well as to recruit psychopharmacology experts to provide appropriate medications to people with SPMI who sought care in these clinics.

In the course of these planning activities, questions arose about the magnitude and distribution of unmet needs for services of the pre-hurricane SPMI population. Needless to say, people with SPMI make up such a small part of the general population that we were unable to make reliable statements about the special needs of people with SPMI based on the CAG sample. Assessments could, of course, be made of unmet demand for treatment of SPMI based on systematic epidemiologic surveillance systems set up in emergency health clinics. However, we know that information on demand for services often fails to give an accurate assessment of need for services, which is why general population needs assessment surveys are of such great importance.

In the case of comparatively rare high-risk populations, the only practical option for needs assessment is to gain access to a list sample that can be used as a sampling frame for tracing. It might sometimes be possible to merge multiple list samples to refine sampling or to answer certain critical policy questions regarding high-risk populations. For example, a comprehensive list existed of all nursing home residents in the areas affected by Katrina that could be linked to the National Death Index (NDI) in order to address concerns that the relocation was associated with a substantial increase in mortality of nursing home residents, although this would involve substantial delays in light of the fact that posting in the NDI sometimes does not occur until as much as a year after death. Linkage of this sort could be done across multiple administrative data systems to generate very useful data, especially when done in conjunction with follow-up surveys. It would be possible, for example, to use linked income tax records and mortality records to track the mortality experience of pre-hurricane residents of the areas affected by Hurricane Katrina who either subsequently returned to their pre-hurricane residence or who moved to a different part of the country.

Similarly, it would be possible to link pre-disaster medical-pharmacy claims data of members of large health plans in areas affected by a disaster with post-disaster claims data, income tax data, and NDI mortality data to track the associations of pre-disaster morbidity with subsequent geographic mobility, health care utilization, and mortality. Targeted tracking surveys then could be used to investigate the determinants of substantially reduced health care utilization among people with evidence of high pre-disaster need for treatment. The main impediment to this kind of integrated analysis is lack of coordination among the agencies and organizations that maintain the many different administrative data systems that would be relevant to such undertakings. Legal constraints on sharing identifying information are important considerations here along with organizational inertia and structural disincentives to collaborate in inter-organizational initiatives. An inter-agency task force in the US federal government is currently grappling with these complex issues in an effort to develop a workable plan for the use of administrative databases in these ways in response to future disasters. In addition, legislation and regulations associated with the US federal government's Confidential Information Protection and Statistical Efficiency Act

(CIPSEA; [www.eia.doe.gov/oss/CIPSEA.pdf](http://www.eia.doe.gov/oss/CIPSEA.pdf)) call for increased data sharing among statistical units of federal agencies and for a correspondingly more extensive confidentiality umbrella over shared data.

## Design considerations

### *Panel versus trend study designs to monitor change*

Our Hurricane Katrina tracking surveys use a panel design (i.e. the same respondents interviewed repeatedly over time) rather than a trend design (i.e. a new sample of respondents selected in each interview) to monitor change. The panel design is preferable to the trend design when the main purpose of tracking is to use baseline information about risk to predict the subsequent onset of some adverse outcome that might be the subject of preventive intervention. There is considerable interest in the literature on PTSD, for example, in the extent to which baseline information obtained shortly after a disaster (the 'peritraumatic' time period) can help pinpoint which disaster victims will or will not subsequently develop PTSD (e.g. Shalev and Freedman, 2005; Simeon et al., 2005). Panel data are needed to investigate such individual differences. However, the panel design is inferior to the trend design when the purpose of the study is to monitor aggregate trends, as the problems of sample reactivity and attrition cumulate in a panel design but not in a trend design.

The decision to use a panel design in the Katrina surveys was based largely on the high costs and complexity of selecting the baseline sample. We did not have enough funds to select a new sample each time we carried out a subsequent wave of data collection. We attempted to deal with the attrition problem, which we expected to be higher than in most other panel surveys because of the instability of the housing situations of many baseline respondents, in a number of ways. First, we made it clear to respondents in the initial recruitment process that we planned to follow them over time to track the course of adjustment to the disaster and we asked for their commitment to stay with the project over a period of several years. Previous research has shown that commitment probes of this sort lead to significant improvements in respondent participation (Oksenberg et al., 1979).

In conjunction with the commitment probe, we characterized the sample to participants as a 'consumer advisory group' in an effort to build commitment to the ongoing enterprise and letting respondents know

that they were community advisors whose views were valued by the project team and the policy-makers who were the primary audience for study findings. Based on concerns about problems tracking the movements of respondents, we provided each respondent with a plastic identification card similar to a credit card that contained the project 800 number and web address. We asked respondents to use this card to contact us whenever they moved to give us their new contact information. We also gathered contact information for three people who were geographically stable that would know the whereabouts of each respondent if the respondent moved and we were unable to trace them. Finally, we sent respondents mailings of study results every six months in order to maintain rapport and to obtain mail address correction information when respondents moved and left a forwarding address. This set of approaches has been very successful in allowing us to track the baseline sample with over 90% success over subsequent waves.

While the strategies described in the last paragraph have the potential to maximize continued participation of baseline CAG members in subsequent interviews, they also have the potential to bias results by changing the cognitive schemas that respondents use in answering survey questions. One way to assess the magnitude of this problem is to carry out a trend survey in parallel with the panel survey to see the extent to which aggregate estimates differ in the two samples. We had originally intended to do this in the CAG, but financial constraints made it impossible to implement a parallel trend component of the design. More generally, though, some version of a mixed panel-trend design would generally be the preferred design in post-disaster needs assessment tracking when the complexities of sampling are not so great that this approach is prohibitively expensive. The mixed panel-trend is a preferred design because we will usually be interested both in aggregate trends and in individual-level change.

#### *Sampling time*

An important design consideration in longitudinal tracking studies is the time interval of assessment. Some tracking surveys are carried out every month and ask respondents to report their experiences over the past 30 days. Other tracking surveys are carried out every six months and ask respondents to report their experiences over a six-month recall period. Others still are carried out every six months and ask respondents

to report their experiences over the past 30 days. The first two of these designs are examples of the continuous time tracking design, one in which the researcher attempts to capture information across the entire time interval since the disaster. The third design (i.e. six-month intervals between data collection waves with 30-day recall questions), in comparison, is an example of a 'snapshot' design, one in which the researcher attempts to collect data only in a sample of time intervals rather than to capture information about experiences over the entire time interval since the disaster.

The decision as to whether the continuous time design or the snapshot design is preferable depends on a number of substantive and logistical considerations that can vary from one study to the next. The most commonly used design in post-disaster needs assessment surveys is a mixed design in which the time interval between waves of data collection is fairly long (6–12 months), some information is collected in a continuous-time framework (e.g. retrospective questions about the persistence of PTSD over the entire time interval since the last survey), while other information is collected in a snapshot framework (e.g. questions about current needs for services). However, this is unlikely to be the optimal design for addressing the research questions that these studies are typically designed to address. The mixed design is the right one, as needs assessment surveys always have multiple goals and it is important to build in the flexibility to include questions that focus on diverse time intervals. However, the long-time intervals that typically exist between waves are sub-optimal, as they make it likely that recall bias will be magnified and that potentially important short-term trends will be missed.

Based on these considerations, a strong argument could be made for a continuous tracking design using the mixed panel-trend approach described in the last sub-section. A variety of mixed panel-trend designs exist (Kish, 1987). One of the most appealing is the rolling panel design, in which new trend survey respondents are recruited on a regular basis (e.g. in monthly samples) and followed over a specified series of panel waves that overlap in time with new trend surveys. This is the design used, for example, in the Bureau of Justice Statistics ongoing National Crime Victimization Survey (NCVS; [www.icpsr.umich.edu/NACJD/NCVS](http://www.icpsr.umich.edu/NACJD/NCVS)), where monthly surveys include samples of people who are interviewed for the first through sixth times with six-month follow-ups between waves of interviewing.



Random effects regression analysis can be used to estimate the impact of non-response bias in the panel component of the data on estimates of trends by taking into consideration systematic variation in trend estimates across the sub-samples (Verbeke and Molenberghs, 2001).

Given that the tracking period for post-disaster needs assessment surveys is typically rather short (no more than several years), a useful variant on the rolling panel design would be to begin with a rather large baseline sample interviewed as soon after the disaster as possible in order to assess peritraumatic stress reactions and to obtain rapid response information about need that can be provided quickly to service planners. In addition, smaller trend samples could be selected on a weekly or monthly basis for a period of six months or so in order to provide fine-grained tracking information on aggregate patterns of persistence or remission of symptoms. Fine-grained tracking could be especially useful when carried out in conjunction with monitoring of mass media messages and treatment recruitment efforts in order to provide information about the effects of social marketing interventions on KAB. Respondents in the baseline interviews could then be re-interviewed after the initial six-month trend period in a panel design that might have a six-month time interval between waves.

The panel component could be carried out with the full baseline sample in a rolling panel framework (e.g. respondents initially interviewed in Month 1 re-interviewed in Month 7, those initially interviewed in Month 2 re-interviewed in Month 8, those initially interviewed in Month 6 re-interviewed in Month 12) to collect information continuously each month, possibly including a small trend component (i.e. a small representative sample of new respondents interviewed each month in Months 7+). Or the panel interviews could be carried out only in a probability sub-sample of baseline respondents that over-samples those with baseline indicators of long-term risk (e.g. retrospectively reported pre-disaster history of psychopathology, extreme peritraumatic stress reactions, high exposure to disaster-related stressors).

This sort of mixed design would maximize flexibility in addressing a wide range of substantive issues and would allow for the rapid assessment of population response to mini-interventions (e.g. an announcement that special funds have been allocated by the federal government for disaster relief; an announcement that

Environmental Protection Agency (EPA) tests documented that fears of toxic chemical exposure were unfounded) both through the investigation of time series in point prevalence of mental disorders and through the inclusion of new public opinion questions on weekly or monthly waves of the survey that ask explicitly about awareness of and reactions to the mini-interventions.

It is important to recognize that the notion of 'continuous' time sampling is a misnomer, as retrospection is always needed in longitudinal data collection even when the time interval between waves is very short. Recall bias can easily creep into retrospective reports, especially with regard to reports of emotional experiences. Indeed, methodological research has shown that bias can be found in emotion reports even over a recall period as short as 24 hours (Diener and Seligman, 2004). Researchers interested in reducing this bias have developed the method of Ecological Momentary Assessment (EMA) (Stone et al., 1999). EMA uses beepers programmed to go off at random times in the day and diaries to have respondents record moment-in-time feelings across a sample of moments and days. An EMA trend study, for example, might recruit a separate random sample of disaster victims each week for one year and ask them to complete moment-in-time assessments at five randomly selected moments on each of the seven days of the week. EMA assessment can sometimes be a very useful adjunct to more conventional panel data collection (e.g. deVries, 1987; Wang et al., 2004). When EMA is considered too molecular, a daily diary can be used instead, with respondents are asked to record the experiences of their day before they go to bed each evening over the course of a one-week or two-week diary period (e.g. Chepenik et al., 2006; Henker et al., 2002).

#### *Before–after designs*

An important limitation of virtually all disaster needs assessment surveys is that respondents are only interviewed after the disaster, making it impossible to make direct before–after comparisons that could estimate the impact of the disaster on the prevalence of mental disorders in the population. There are some exceptions to this general problem. For example, the Epidemiological Catchment Area (ECA) Study in St Louis (Regier et al., 1984) was carried out shortly before the 1985 flood, dioxin exposure scare, and subsequent mass evacuation of Times Beach, a small town on the outskirts

of the St Louis Metropolitan Area that was in the ECA sample. This created an opportunity to carry out a before–after comparison of mental health associated with the Times Beach disaster. But situations of this sort are rare. The much more typical situation is for studies of the mental health impact of disasters to be carried out only after the fact. Information about pre-disaster psychopathology is collected retrospectively.

Two practical approaches exist to introduce before–after information on a more routine basis into post-disaster needs assessment surveys. The first is to use tracking information from ongoing government health surveys to construct an appropriate *post hoc* pre-disaster comparison group. Three major ongoing national surveys exist that could be used in this way: the US National Health Interview Survey (NHIS; [www.cdc.gov/nchs/about/major/nhis/his.sample.htm](http://www.cdc.gov/nchs/about/major/nhis/his.sample.htm)), which carries out face-to-face interviews weekly with a nationally representative sample that includes approximately 43,000 households each year; the CDC Behavioral Risk Factor Surveillance Survey (BRFSS; [www.cdc.gov/brfss/about.htm](http://www.cdc.gov/brfss/about.htm)), which carries out weekly telephone interviews with a sample in each of the 50 United States, the District of Columbia, Puerto Rico, the US Virgin Islands, and Guam that includes more than 350,000 interviews each year; and the SAMHSA National Survey on Drug Use and Health (NSDUH; [www.oas.samhsa.gov/redesigningNHSDA.pdf](http://www.oas.samhsa.gov/redesigningNHSDA.pdf)), which carries out annual face-to-face interviews with a nationally representative sample of approximately 70,000 respondents with an over-sample of the most populous states and of youth. Importantly, all three of these surveys include a version of the K-6 scale of psychological distress (Kessler et al., 2002; Kessler et al., 2003), the core global screening measure of the Diagnostic and Statistical Manual of Mental Health, fourth edition (DSM-IV) anxiety-mood disorders that we use in our model post-disaster mental health needs assessment tracking survey.

This truly massive resource of baseline information, with roughly one out of every 600 adults in the entire US being interviewed in one of these surveys each year, could be used to provide baseline information to assess the effects of disasters on the mental health of local populations by selecting sub-samples appropriate for comparison with targeted disaster populations. To illustrate the potential of this approach, consider the case of Oklahoma City (3 450 000 residents in the 2000 Census), the site of a 1995 terrorist attack on a US

government office complex that killed 168 people. Given the size of the three surveys described earlier and the size of Oklahoma City, a sample of roughly 5000 adult residents of Oklahoma City would have been interviewed in one of these surveys in the 12 months before the terrorist attack if all three surveys had been in place in the mid-1990s. A sample as large as this would create a very stable baseline for assessing the mental health effects of the terrorist attack. In the case of smaller disaster areas, such as Graniteville, South Carolina (population 7112), we could combine information from three surveys in similar communities collected over the prior 12 months to construct an approximate pre-disaster comparison group. Or we could combine data from interviews with residents of areas in the vicinity of the disaster site collected over a decade or more before the disaster with post-disaster interviews in the affected area and use interrupted time series analysis (McDowell et al., 1980) to estimate the effect of the disaster on the mental health of residents.

There are bureaucratic impediments to carrying out this type of analysis in that the government agencies that administer the three ongoing surveys have restrictions on making information available to researchers about small area geographic characteristics of individual respondents. Even more important, the agencies are slow in releasing the survey data for public use, making it impossible to obtain pre-disaster data in a time frame that would be useful for disaster response planning purposes. These impediments made it impossible for us to use data from any of these surveys in pre-post analyses of the mental health effects of Hurricane Katrina even though we estimate that more than 6000 residents of the areas affected by Katrina were respondents in one of these three surveys in the 12 months before the hurricane. Efforts have been made recently to decrease the time delays in producing usable data files from these surveys. We hope that the creation of the NIMH center for post-disaster mental health needs assessment surveys will help cartelize these efforts and make it possible to use these surveys to create pre-disaster comparison groups that can be used in needs assessment studies of future disasters.

The availability of before–after data can be very useful in addressing an important question about need that we noted in the introduction: that the socio-demographic correlates of need for treatment found in post-disaster surveys might have existed before the disaster, in which case they could be unrelated to the

disaster. An illustration of such an analysis is our use of data collected in the 2001–2003 National Comorbidity Survey Replication (NCS-R) (Kessler and Merikangas, 2004) among respondents in the two Census Divisions subsequently affected by Hurricane Katrina to approximate a before–after comparison of the prevalence of serious mental illness (Kessler et al., 2006). The K-6 was used to screen for 30-day DSM-IV anxiety and mood disorders in both the NCS-R and the baseline CAG survey. Based on previous K-6 validation (Kessler et al., 2003), scores on the 0–24 scale in the range 13–24 were classified probable serious mental illness (SMI).

A variety of socio-demographic correlates of SMI were assessed in a comparable way in the two surveys. The estimated prevalence of SMI was found to be dramatically higher in the CAG than the NCS-R. Socio-demographic variation in this between-survey difference was assessed by pooling the data in the two surveys into a single data analysis file and estimating logistic regression equations to predict SMI from a 0–1 dummy variable (0 = the NCS-R, 1 = CAG), the socio-demographic variables, and interactions between the survey dummy and the socio-demographic variables. A great many significant socio-demographic correlates of SMI were found in the CAG, such as female gender, low education, and pre-hurricane unemployment. However, all of these associations were also found in the NCS-R and none of the associations was significantly stronger in the CAG than the NCS-R. This is consistent with the view that the adverse mental health effects of Katrina were equally distributed across broad segments of the population (in the sense that rates of SMI increased proportionately in each group) despite the fact that SMI was significantly more common in some socio-demographic segments of the CAG sample.

We noted earlier that there are two practical approaches to introduce before–after information on a routine basis into post-disaster needs assessment surveys. The first one, which we just reviewed, is to use information from ongoing government health surveys to construct an approximate pre-disaster comparison group for pre-post trend analysis. The second is to use the same sort of data for panel analysis. The latter is often referred to as a ‘follow-back’ design (Castle et al., 2004; Seeman et al., 1989). In this approach, respondents who participated in a government survey some time prior to the disaster could be traced and re-interviewed after the disaster to provide individual-level pre-post information. As noted earlier, we estimate that more than

6000 residents of the areas affected by Hurricane Katrina were respondents in one of the three major government surveys that collect K-6 information in the 12 months before the hurricane. It might have been difficult to trace all these people by trying to contact them at their pre-hurricane addresses and searching for them on safe lists and ARC-FEMA lists, but the degree of tracing success would in itself have been useful to know along with the substantively useful information that would have been obtained from the individual-level pre-post comparisons of K-6 scores and pre-disaster predictors of individual-level changes in these scores. Although we are unaware of any previous use of this design to evaluate the effects of disasters, we plan to use this design as part of our collaboration with the BRFSS in future post-disaster mental health needs assessment studies.

#### *Surveying help-seekers*

The ARC and FEMA lists of people who apply for assistance are made up entirely of people who sought help. Help-seekers presumably differ from other residents of disaster populations in a number of ways, including both in the extent of their need for help (i.e. the extent to which they experienced property loss in the disaster) and in the extent to which they are motivated and capable of making an application. Although we might expect to find a meaningful number of victims with high need who did not seek help due to extreme physical restrictions (e.g. housebound in a wheelchair), possibly in conjunction with extreme social isolation and communications problems (e.g. no access to a telephone, unable to speak English, blind or deaf), relief agencies make efforts to find such people through a variety of community outreach and household screening programs. Based on this fact, it is not unreasonable to think that fairly representative data on demand for services could be obtained by sampling people who applied for relief even though the sample might not represent all people with need for services.

In the case of need for treatment of mental disorders, a very important set of post-disaster help-seekers consists of those who call the various mental health crisis hotlines that are typically established by local and national mental health associations. The largest and most important of these is the National Suicide Prevention Lifeline, the only national suicide prevention and intervention telephone line sponsored by the federal government ([www.suicidepreventionlifeline.org](http://www.suicidepreventionlifeline.org)). The

Lifeline was launched in December 2004 to link callers to staff in more than 120 mental health crisis centers around the country. SAMHSA used the lifeline crisis phone number as the hub for mental health referrals during the aftermath of Hurricane Katrina and it is likely to do so again in future mass disasters. Follow-up needs assessment surveys with callers of Lifeline and other crisis hotlines could be useful components of larger post-disaster mental health needs assessment efforts in at least three important ways.

First, an important under-studied issue concerns patterns and determinants of unmet need for treatment of mental health problems after disasters (Boscarino et al., 2005; Stuber and Galea, 2005). A useful way to study this issue would be to carry out follow-up interviews with callers of mental health hotlines that were given a referral for treatment. The information obtained in these interviews about modifiable barriers to treatment could be organized using existing conceptual frameworks (Rogler and Cortes, 1993) that might provide insights into potential values modifications in the referral process. We know of no previous research of this sort carried out with callers to post-disaster mental health referral lines. However, we have established collaborations with the ARC and with Mental Health America (MHA; formerly known as the National Mental Health Association) as well as with a number of MHA affiliates, including the National Suicide Prevention Lifeline, to implement this type of study as part of a larger plan for the proposed NIMH center for post-disaster needs assessment tracking.

Second, relatively little is known about the quality of care provided to patients referred by crisis hotlines after disasters to local mental health treatment centers. This quality control problem could be addressed, at least in part, by carrying out systematic follow-up interviews that assess patient satisfaction. Surveys of this sort are now a routine part of many treatment quality assurance programs, the most notable example being the Consumer Assessment of Healthcare Providers and Systems (CAHPS) program ([www.chaps.ahrq.gov/default.asp](http://www.chaps.ahrq.gov/default.asp)), which now includes a behavioral health care component. Publicizing the 'report cards' generated by the results of these surveys in conjunction and other quality indicators has been shown to influence consumer choice of health plans (Jin and Sorensen, 2006; Oetjen et al., 2006) which, in turn, is hoped to influence health plan performance. As part of the proposed collaboration with MHA noted in the last

paragraph, we plan to develop a similar system that will carry out CAHPS-like follow-up surveys with patients who are referred to post-disaster mental health services. It is important for these surveys to be very inexpensive because the goal would be to give all patients a chance to respond so as to obtain countable information for as many service providers as possible. As a result, patients who have an email address will be surveyed using inexpensive web survey technology (Schonlau et al., 2002), while other patients will be interviewed using inexpensive IVR technology.

Third, there is considerable uncertainty about the most appropriate interventions to use in treating the emotional problems of disaster victims (Watson and Shalev, 2005). This uncertainty is due in no small part of the difficulties involved in carrying out controlled treatment studies in disaster situations. A potentially useful way to address this problem would be to build in randomization of referrals of help-seekers to different treatment settings and types in conjunction with the follow-up interviews described in the last two paragraphs. This approach could be used to evaluate a highly specified treatment approach that is experimentally provided to a small probability sub-sample of help-seekers in comparison to the usual care provided to all other disaster victims. Alternatively, all help-seekers could be randomized across the range of seemingly appropriate treatment settings available in a given disaster situation and follow-up questionnaires of the sort described in the last paragraph could be used to determine whether effectiveness varies significantly across these settings both in the aggregate and for patients with particular characteristics. With regard to the latter, the large numbers of patients included in the randomization in a major disaster would make it possible to determine whether overall treatment effectiveness could be improved by some type of patient-program matching.

### Overview

As noted in the introduction, a wide range of measurement, design, and analysis issues present themselves in planning a consistent approach to the implementation of post-disaster mental health needs assessment surveys. We focused here only on design issues due to the fact that these have been much less widely discussed in the literature than measurement or analysis issues. It needs to be recognized, though, that consistency of measurement has to be the first step in the process of

coordinating needs assessment. The recent history of mental health assessment among survivors of Hurricane Katrina illustrates the problem. The Louisiana Department of Public Health documented substantial psychopathology among the 50,000 Katrina survivors cared for in ECs shortly after the hurricane based on a measure of unknown validity (Centers for Disease Control and Prevention, 2006b), while the Centers for Disease Control and Prevention carried out a household needs assessment survey that found half of adults still living in New Orleans to have clinically significant psychological distress using a completely different unvalidated measure (Centers for Disease Control and Prevention, 2006a). Two public opinion polls, one carried out jointly by Gallup, CNN, and *USA Today* in a sample of people who sought ARC assistance (Page, 2005), and the other carried out by the *New York Times* in a sample from the ARC safe list (Dewan et al., 2006), also asked a small number of questions about mental health, but without attempting to assess clinical significance. A probability survey of families with children still residing in FEMA-sponsored trailers or hotel rooms in Louisiana as of mid-February 2006 found 44% of adult caregivers to have clinically significant psychological distress, but this conclusion was based on yet another unvalidated measure (Abramson and Garfield, 2006). The CAG subsequently carried out a general population survey of Katrina survivors using the validated K-6 scale to assess mental illness, but the CAG results cannot be compared with the results of the earlier more focused studies because of non-comparability of measures. The joint results of these studies would have been much more useful if they had all used comparable measures to assess mental illness. It should be an easy matter to coordinate these measures, so this should be the first priority in studies of future disasters.

Design and analysis issues are much more difficult. Design issues are especially complex due to the fact that the circumstances of disasters and the resources available to construct sampling frames after disasters differ greatly across disaster situations. We reviewed in this paper the main design challenges as we see them with a focus on implementation of post-disaster surveys in the US. A more complex set of challenges exists in other countries, especially less developed countries, where the survey methodology infrastructure is less well developed than in the US. Because of this highly developed infrastructure, we focused more on opportunities than challenges, as a number of very important

and heretofore neglected opportunities exist substantially to improve the quality and scope of US post-disaster needs assessment surveys by exploiting existing survey infrastructure and technology.

The three large ongoing federal government surveys that all include the K-6 are especially noteworthy opportunities both as a basis for before–after trend comparisons and for follow-back panels. There is no reason other than bureaucratic roadblocks that post-disaster needs assessment surveys should not be coordinated with these three surveys. We are especially pleased, in this regard, that we have been able to develop an agreement with the BRFSS for collaboration in future telephone needs assessment surveys. We hope this agreement can serve as a model for subsequent collaborations with the NHIS and the NSDUH.

Our burgeoning collaborations with the ARC, MHA, and the National Suicide Prevention Lifeline are also noteworthy because follow-up needs assessment surveys with disaster survivors selected from their lists will give us ready access both to samples of high-risk survivors for purposes of needs assessment and to representative samples of help-seekers for purposes of studying barriers to treatment and perceived quality of services provided as well as to carrying out experimental evaluations of intervention effects. The last of these could be especially important in light of the continued paucity of information on the effectiveness of the treatments typically provided to disaster victims. This lack of information is certainly understandable in light of the fact that treatment effectiveness studies require careful prior planning and implementation that are often impossible to realize in disaster situations. However, by preparing an infrastructure for post-disaster needs assessment in collaboration with the large national organizations involved in ongoing disaster relief, we hope to render these problems tractable and to help build a platform for creating evidence-based standards for effective disaster mental health interventions that will help address the enormous unmet need for services that so often exists among survivors of disasters.

#### Acknowledgements

Preparation of this paper was supported by NIH Research Grants R01 MH070884-01A2 and R01 MH081832 from the US Department of Health and Human Services, National Institutes of Health (NIH), the Office of the Assistant Secretary of Planning and Evaluation, the Federal Emergency Management Agency, and the Administration for Children and Families.

## Declaration of Interests

Dr. Kessler has been a consultant for GlaxoSmithKline Inc., Kaiser Permanente, Pfizer Inc., Sanofi-Aventis, Shire Pharmaceuticals, and Wyeth-Ayerst; has served on advisory boards for Eli Lilly & Company and Wyeth-Ayerst; and has had research support for his epidemiological studies from Bristol-Myers Squibb, Eli Lilly & Company, GlaxoSmithKline, Johnson & Johnson Pharmaceuticals, Ortho-McNeil Pharmaceuticals Inc., Pfizer Inc., and Sanofi-Aventis. The remaining authors have nothing to declare.

## References

- Abramson D, Garfield R (2006). On the Edge: Children and Families Displaced by Hurricanes Katrina and Rita Face a Looming Medical and Mental Health Crisis. New York: Mailman School of Public Health, Columbia University.
- Ben-Ezra M, Essar N, Saar R (2006). Gender differences and acute stress reactions among rescue personnel 36 to 48 hours after exposure to traumatic event. *Traumatology* 12: 139–42.
- Boscarino JA, Adams RE, Stuber J, Galea S (2005). Disparities in mental health treatment following the World Trade Center Disaster: implications for mental health care and health services research. *J Trauma Stress* 18: 287–97.
- Brewin CR, Andrews B, Valentine JD (2000). Meta-analysis of risk factors for posttraumatic stress disorder in trauma-exposed adults. *J Consult Clin Psychology* 68: 748–66.
- Brick MJ, Dipko S, Presser S, Tucker C, Yuan Y (2006). Non-response bias in a dual frame sample of cell and landline numbers. *Public Opinion Quarterly* 70: 780–93.
- Brodie M, Weltzien E, Altman D, Blendon RJ, Benson JM (2006). Experiences of hurricane Katrina evacuees in Houston shelters: implications for future planning. *Am J Public Health* 96: 1402–8.
- Castle K, Duberstein PR, Meldrum S, Conner KR, Conwell Y (2004). Risk factors for suicide in blacks and whites: an analysis of data from the 1993 National Mortality Followback Survey. *Am J Psychiatry* 161: 452–8.
- Centers for Disease Control and Prevention (2005). Emergency Preparedness and Response: Strategic National Stockpile. Available: <http://www.bt.cdc.gov/stockpile/index.asp> [11 April 2006].
- Centers for Disease Control and Prevention (2006a). Assessment of health-related needs after Hurricanes Katrina and Rita – Orleans and Jefferson Parishes, New Orleans area, Louisiana, October 17–22, 2005. *MMWR Morb Mortal Wkly Rep* 55: 38–41.
- Centers for Disease Control and Prevention (2006b). Surveillance in hurricane evacuation centers – Louisiana, September–October 2005. *MMWR Morb Mortal Wkly Rep* 55: 32–5.
- Chepenik LG, Have TT, Oslin D, Datto C, Zubritsky C, Katz IR (2006). A daily diary study of late-life depression. *Am J Geriatr Psychiatry* 14: 270–9.
- deVries MW (1987). Investigating mental disorders in their natural settings. *J Nerv Ment Dis* 175: 509–13.
- Dewan S, Connelly M, Lehren A (2006) Evacuees' lives still upended seven months after hurricane. *New York Times* 22 March, A1.
- Diener E, Seligman MEP (2004). Beyond money: toward an economy of well-being. *Psychological Science in the Public Interest* 5: 1–31.
- Fisher N, Turner SW, Pugh R, Taylor C (1994). Estimating numbers of homeless and homeless mentally ill people in north east Westminster by using capture-recapture analysis. *Br Med J* 308: 27–30.
- Flay BR, Kessler RC, Utts JM (1989). Evaluating media campaigns. In Coyle SL, Boruch RF, Turner CF (eds) *Evaluating AIDS Prevention Programs*. Washington, DC: National Academy Press, pp 3(1)–3(47).
- Fullerton CS, Ursano RJ, Wang L (2004). Acute stress disorder, posttraumatic stress disorder, and depression in disaster or rescue workers. *Am J Psychiatry* 161: 1370–6.
- Galea S, Nandi A, Vlahov D (2005). The epidemiology of post-traumatic stress disorder after disasters. *Epidemiol Rev* 27: 78–91.
- Gidron Y (2002). Posttraumatic stress disorder after terrorist attacks: a review. *J Nerv Ment Dis* 190: 118–21.
- Henker B, Whalen CK, Jamner LD, Delfino RJ (2002). Anxiety, affect, and activity in teenagers: monitoring daily life with electronic diaries. *J Am Acad Child Adolesc Psychiatry* 41: 660–70.
- Ironson G, Wynings C, Schneiderman N, Baum A, Rodriguez M, Greenwood D, Benight C, Antoni M, LaPerriere A, Huang HS, Klimas N, Fletcher MA (1997). Posttraumatic stress symptoms, intrusive thoughts, loss, and immune function after Hurricane Andrew. *Psychosom Med* 59: 128–41.
- Jin GZ, Sorensen AT (2006). Information and consumer choice: the value of publicized health plan ratings. *J Health Econ* 25: 248–75.
- Kalton G, Anderson D (1986). Sampling rare populations. *J Royal Stat Soc Series A* 149: 65–82.
- Kessler RC, Merikangas KR (2004). The National Comorbidity Survey Replication (NCS-R): background and aims. *Int J Methods Psychiatr Res* 13: 60–8.
- Kessler RC, Galea S, Jones RT, Parker HA (2006). Mental illness and suicidality after Hurricane Katrina. *Bull World Health Organ* 84: 930–9.
- Kessler RC, Andrews G, Colpe LJ, Hiripi E, Mroczek DK, Normand SL, Walters EE, Zaslavsky AM (2002). Short screening scales to monitor population prevalences and trends in non-specific psychological distress. *Psychol Med* 32: 959–76.
- Kessler RC, Barker PR, Colpe LJ, Epstein JF, Gfroerer JC, Hiripi E, Howes MJ, Normand SL, Manderscheid RW,

- Walters EE, Zaslavsky AM (2003). Screening for serious mental illness in the general population. *Arch Gen Psychiatry* 60: 184–9.
- Kish L (1987). *Statistical design for research*. New York: John Wiley & Sons.
- Kohn R, Levav I, de Almeida JM, Vicente B, Andrade L, Caraveo-Anduaga JJ, Saxena S, Saraceno B (2005). [Mental disorders in Latin America and the Caribbean: a public health priority]. *Rev Panam Salud Publica* 18: 229–40.
- McDowell D, McCleary R, Meidinger EE, Hays RA (1980). *Interrupted Time Series Analysis*. Beverly Hills, CA: Sage Publications.
- Norris F (2005). Psychosocial Consequences of Major Hurricanes and Floods: Range, Duration, and Magnitude of Effects and Risk Factors for Adverse Outcomes: A National Center for PTSD fact sheet. Available: [http://www.ncptsd.va.gov/facts/disasters/fs\\_range\\_hurricane.html](http://www.ncptsd.va.gov/facts/disasters/fs_range_hurricane.html) [14 April 2005].
- North CS, Pfefferbaum B, Tivis L, Kawasaki A, Reddy C, Spitznagel EL (2004). The course of posttraumatic stress disorder in a follow-up study of survivors of the Oklahoma City bombing. *Ann Clin Psychiatry* 16: 209–15.
- Oetjen D, Fottler MD, Unruh LY, Rehman Z (2006). Consumer determinants of the use of health plan information in plan selection. *Health Serv Manage Res* 19: 232–50.
- Oksenberg L, Vinokur A, Cannell CF (1979). The effects of instructions, commitment and feedback on reporting in personal interviews. In Cannell CF, Oksenberg L, Converse JM (eds) *Experiments in Interviewing Techniques*, DHEW Publication No (HRA) 78–3204. Washington, DC: Department of Health, Education, and Welfare, pp 74–108.
- Page S (2005). Many evacuees to stay away. *USA Today* 14 October, A1.
- Regier DA, Myers JK, Kramer M, Robins LN, Blazer DG, Hough RL, Eaton WW, Locke BZ (1984). The NIMH Epidemiologic Catchment Area program. Historical context, major objectives, and study population characteristics. *Arch Gen Psychiatry* 41: 934–41.
- Rogler LH, Cortes DE (1993). Help-seeking pathways: a unifying concept in mental health care. *Am J Psychiatry* 150: 554–61.
- Rosenbaum S (2006). US health policy in the aftermath of Hurricane Katrina. *JAMA* 295: 437–40.
- Schonlau M, Fricker R, Elliott M (2002). *Conducting Research Surveys via Email and the Web*. Santa Monica, CA: RAND.
- Seeman I, Poe GS, McLaughlin JK (1989). Design of the 1986 National Mortality Followback Survey: considerations on collecting data on decedents. *Public Health Rep* 104: 183–8.
- Shalev AY, Freedman S (2005). PTSD following terrorist attacks: a prospective evaluation. *Am J Psychiatry* 162: 1188–91.
- Simeon D, Greenberg J, Nelson D, Schmeidler J, Hollander E (2005). Dissociation and posttraumatic stress 1 year after the World Trade Center disaster: follow-up of a longitudinal survey. *J Clin Psychiatry* 66: 231–7.
- Skinner C, Rao JNK (1996). Estimation in dual frame surveys with complex designs. *J Am Stat Assoc* 91: 349–56.
- Stone AA, Shiffman SS, DeVries MW (1999). Ecological momentary assessment. In Diener E, Schwarz N, Kahneman D (eds) *Well-being: The Foundation of Hedonic Psychology*. New York: Russell Sage Foundation.
- Stuber JP, Galea S (2005). Barriers to mental health treatment after disasters. *Psychiatr Serv* 56: 1157–8.
- Subar AF, Heimendinger J, Patterson BH, Krebs-Smith SM, Pivonka E, Kessler RC (1995). Fruit and vegetable intake in the United States: the baseline survey of the Five A Day for Better Health Program. *Am J Health Promot* 9: 352–60.
- US Environmental Protection Agency (2005). *Norfolk Southern Graniteville Derailment: Final Information Update as of January 21, 2005*. Atlanta, GA: Department of Health and Human Services, Public Health Service, Agency for Toxic Substances and Disease Registry, Division of Health Assessment and Consultation.
- Verbeke G, Molenberghs G (2001). *Linear Mixed Models for Longitudinal Data*. New York: Springer-Verlag.
- Wang PS, Beck AL, Berglund P, McKenas DK, Pronk NP, Simon GE, Kessler RC (2004). Effects of major depression on moment-in-time work performance. *Am J Psychiatry* 161: 1885–91.
- Watson PJ, Shalev AY (2005). Assessment and treatment of adult acute responses to traumatic stress following mass traumatic events. *CNS Spectr* 10: 123–31.

*Correspondence:* R. C. Kessler, Department of Health Care Policy, Harvard Medical School, 180 Longwood Avenue, Boston, MA 02115, USA.  
 Telephone (+1) 617-432-3587  
 Fax (+1) 617-432-3588  
 Email: [kessler@hcp.med.harvard.edu](mailto:kessler@hcp.med.harvard.edu)