

## HRSM Session 1    Advances in Collecting Health Data Survey Data

John Boyle (ICF): Propensity to Participate in Surveys with Physical Measures and Biomarker Collection: Potential Opportunities and Limitations

Kristen Becker (RAND): Examining the Effectiveness of Push-to-Web Mixed Mode Approaches on Response Rates on an Emergency Department Setting – a Randomized Study

Martha Kovac (Mathematica): Conducting a Mixed-Mode Web and Telephone Survey with an Adult Population with Disabilities: Will a Web Survey Mode Improve Response and Data Quality?

David Dutwin (NORC): From RDD to ABS: The Migration of Health Interview Surveys  
Chair: Jen Dykema (UWSC)

Discussant: John Stevenson (UWSC)

Rapporteurs: Eran Ben-Porath (SSRS); Jessica Graber (CDC); Jacquie Hogan (Westat)

### **Rapporteur Overview**

This session addressed the changing landscape of health survey data collection in light of declining response rates, new technologies, and available secondary data sources. Each paper addresses one or more of the following issues: propensity to respond to surveys; multi-mode design, optimizing contact strategies; and sampling strategies.

#### **(1) Can we estimate propensity to respond?**

ICF conducted a formative research study using a non-probability online panel. Participants were asked a series of questions about their willingness to participate in a health study where certain physical measures were collected. Based on these results, they found that those with poorer health status and greater numbers of doctors and hospital visits are more likely to report intent to participate as are those with serious or chronic health conditions and with unmet health needs. Those findings are consistent with what is reported in NHIS and actual participation in NHANES health exams.

But how predictive are attitudes of actual behavior?

Not trying to get at predictive validity. He doesn't know how well we can move the needle and increase response rates. He is focused on salience and the topics of health and thinks that we will engage Rs who are engaged on this topic. But we need to find ways to frame the study in different ways to engage people to make the study salient for them.

The idea is that by measuring intent to participate we are identifying issues of salience to respondents. Knowing these can help us tailor our modes, incentives, outreach, and refusal conversion activities to specific groups rather than using a one-size-fits all approach.

The study is looking at reported motivation and not actual behavior. No consideration is being given to any structural barriers or ability to participate in surveys. While this study looked at topical salience, attitude, motivation and desire, actual health surveys need to facilitate the R's ability to take action.

Additionally, knowing in advance the groups that are more likely to participate and cases that can be completed more easily and cheaply will allow for planning and time to work the more difficult cases.

Since NHANES data demonstrates that most participants who complete the household interview also complete the physical exam, these propensity models are applicable to more than just biomeasure response rates.

Other general population studies have shown the opposite. Jennifer Dkyema suggests additional testing of intent using vignettes or multifactorial approaches may be more predictive of Rs to participate.

Concern expressed that studies typically show that the less-healthy have lower propensity to participate in biomeasures. Boyle suggests that this could be related to barriers: while the less healthy may be inclined to participate more, it may be more difficulty for them to follow-through on this propensity because their health forces limitation on scheduling and mobility.

## **(2) Can we optimize contact strategies to improve response rates?**

Becker presented a study conducted at 16 hospitals who surveyed patients discharged from Emergency Departments. Cases were randomized into 8 experimental treatment arms with unique "push-to-web" contact strategy approaches. The 9<sup>th</sup> arm was a control group who received a mailing and a follow up telephone call, which is the current protocol.

The approach that produced the highest response rate was the addition of a single email prior to the mailing and the telephone call. While an email address was not available for all sampled respondents, for those for whom it was, the response rate increased approximately 7 percent.

Implementing the texting protocol was also difficult given the TCPA requirements for consent. Documented consent was required, and each hospital collected this differently with results ranging in email coverage as low as single percentage points to over fifty percent.

Kovac stresses that the differential-incentive in her mix mode study, in which web respondents received a higher incentive, led to higher web-response than anticipated.

## **(3) What is the effect of mode?**

Boyle noted that regardless of stated intent to participate, the survey design, including mode, will affect actual response. The impact of the interaction with an interviewer – in a respondent's home, on the telephone, or at some other location - allows for the opportunity to build rapport and for persuasion, regardless of the salience of the survey topic. Web and mail surveys do not allow for such interactions.

Becker's presentation demonstrated that respondents will take advantage of a web response option but not at the rates desired. The need for CAPI and PAPI response options is still clear as respondents selected those modes as their preferred way to respond. Further analyses are also needed to understand the benefits and in terms of cost and time savings by implementing a web response option and the effects of text messages on response rates.

Kristen Olson (UNL) asked whether administrative records were available from hospitals to provide additional context to the response mode selected by the respondent. Per Becker, a measure of severity was available for each case and web respondents were found to be more likely to have a less urgent problem.

Kovac described the implementation of a demonstration project for the Social Security Administration where a mixed mode survey was administered to adult respondents with disabilities.

Respondents received a letter inviting them to participate and providing them with a web address where they could complete the survey online or by telephone. An incentive was offered for their participation, with the web incentive being higher than the telephone incentive.

The web survey was designed to be 508 compliant, with simple formatting. The URL was short and easy to type. The questionnaire was programmed so respondents could skip any question and there were no hard edit checks to cause usability issues.

Overall, response to the survey online was higher than over the telephone but it varied by demographic group. Web respondents were more likely to be women, white, non-Hispanic, younger, and have higher educational levels. While CATI respondents were more likely to report their health as poor, web respondents were more likely to report a physical health condition or emotional health limitation. Kovac suggested that the phone respondents may also be different as they were the late respondents, completing their survey weeks into the field.

Response rates were high overall but sampled respondents had previously opted to participate in a demonstration project and therefore were likely more willing.

Shane Desselle (Research and Social Administrative Pharmacy) commented that many funders of research want a web-based survey because they think they are cheaper and asked if there is evidence of the cost effectiveness of mixed modes or that they reach under covered populations? Dutwin cautioned that in survey design, cost should not be the only consideration. For example, opt-in research may have double the bias and 3-4 times the variability. So, cost is irrelevant if in the end the survey produced poor quality and inaccurate results.

#### **(4) What are the tradeoffs when using Address Based Sampling?**

Dutwin presented on the transition from RDD to ABS and the implications of this change.

One aspect is the use of secondary data (Census Planning Database (CPDB), USPS Delivery Sequence File, consumer data) to conduct nonresponse analyses, model sampling strata, etc. These secondary data sources can be used to enhance study operations – planning for materials translated in multiple languages, creating strategies for areas with low response scores (LRS) in the CPDB, etc. In general population surveys, the stratification is generally now done by modeling rather than by geography.

Stephen Blumberg (NCHS) asked what we know about motivating respondents in low response score strata. This led to discussion about the use of increased incentives, additional mailings and alternative modes of response. It was noted that CATI response options are important for low income and low literacy respondents, characteristics highly correlated with LRS strata. Brad Edwards (Westat) mentioned that adding face-to-face interviewing and ASACI are effective ways of reaching LRS/hard-to-reach populations, in addition to other modes offered.

Boyle suggested reframing and tailoring the survey's communications for nonrespondents. Rather than assume that the original message that was effective for the early responders is appropriate, understand what is salient to the nonrespondents and address those concerns in ensuing waves.

Another discussion focused on the question of how much improvement in data-quality comes with the possibilities of weighting ABS data to more variables such as those in CPDB or appended consumer data. In response to the concern of added variability due to this weighting, Dutwin suggests weighting only to parameters that could correlate with variables in interest, a growing tolerance in the field for higher design effect, and the need for analysis of the tradeoffs between bias reduction and variance inflation. He also adds that even though investing resources into efforts of reaching LRS respondents makes data collection more expensive, and reduces response rate, it can improve data quality.

Lastly, as far as switching from RDD to multi-mode data-collection, there was concern expressed with SAQs (mail) in complex health insurance studies, where respondents fail to follow skip instructions.