

TRANSCRIPT

HOST: NCHS has been a leader among federal agencies in developing a data linkage program that builds the data resources needed to better understand the health of the US population, and the effects of public health policies meant to protect or improve the health of all Americans.

Joining us today is Lisa Mirel, Director of the Data Linkage program at NCHS. So do you need approval then from the survey participants for their information be used in this kind of expanded fashion?

LISA MIREL: Yes definitely - I think that's a great question. We are actually only able to link data from survey participants that have consented to linkage and that have also provided sufficient personally identifiable information. So our group actually works very closely with some of the survey divisions to figure out really the best approach to get that kind of information from their survey participants. So we worked with them about coming up with questions about consent for linkage and maybe even changing how they are asking for some of the personally identifiable information but making sure that we are still getting enough that we can perform high quality linkages as well.

HOST: And how many of these survey participants are eligible to be part of the linkage activities?

LISA MIREL: So, right now it's probably between about 85 and 90% of the survey participants are eligible for linkage and this is actually a pretty big change from what was going on around like the mid-2000s. At that time we were only at about 50% linkage eligibility. So that really had implications for analysis and inference using the linked data when we only had about 50% of the people coming in that were linkage eligible, and then only some small percentage of them actually linking to the outside records. But you know as I mentioned, we worked really closely with the survey divisions when we started seeing these real declines in people wanting to provide their information for linkage. And we changed one of the questions in the questionnaire to ask a direct question about consent for linkage, basically like "Can we link your records?" and then we also started only asking for the last four digits of the Social Security number. So in earlier years we had asked for all 9 digits, and I think over time people became - they were feeling less likely that they wanted to share that type of information with us on the surveys so they would just provide, we started just asking for the last four digits of the Social Security number and that really increased the linkage eligibility rates. But even though we still have a pretty high rate right now our group spends a lot of time looking at what you might want to call "linkage eligibility bias" for those people who aren't even eligible for linkage, so they don't even come into the picture. So we actually have a report coming out - it should be coming out sometime in the next month; it just received approval in the clearance process - and it really illustrates what the linkage eligibility bias looks like in terms of different estimates coming out of the surveys. And then we also provide ways to help mitigate that bias and talk through some of the statistical methods that could be used to make up for those who are not eligible for linkage.

HOST: OK - could you share some examples of epidemiological and policy-related research that can come from data linkage?

LISA MIREL: I would be happy to. So one of the things that we do in the branches is we actually track the different citations of people who have used the linked data. So we do these bibliography updates about every six months and we post them on our website so that researchers can come in and see what types of research has been done already with the analyses and maybe get or generate some ideas of some analyses they might want to do. And I think our last count we had cited about over 950 publications that have used the linked data. So I guess some examples would be in using the data that have been linked to the National Death Index people have looked at deaths that are associated with being underweight, overweight or obese. People have also looked at air pollution exposure and heart

disease mortality and then also looked at educational differences and U.S. adult mortality. With some of the CMS data, people have looked at the characteristics of those who chose Medicare Advantage upon entering Medicare enrollment. And people have also looked at health services use among the previously uninsured who then once they enrolled into Medicare and then they had that health insurance. And we've done a lot using the Housing and Urban development data. People have looked at blood lead levels as a comparison for those who are living in HUD assisted housing compared to those who are not. People have looked at cigarette smoking and adverse health outcomes among adults who are living in federal assisted housing and I think that ties into some smoking legislation programs. And people also looked at housing assistance associated with insurance rates and unmet medical needs. So those are examples of some of the epidemiological work, but I did want to mention just a few of, like, the policy-related work that has been done using our data. Right around the time - I had mentioned previously about the Evidence Act - right around the time of the passage of that we were actually invited to be part of a book that was looking at a collection of case studies that highlighted many different approaches to using evidence-based data. And so in that book, we wrote a chapter about the HUD data and again that was looking at blood lead levels for children living in HUD assisted housing compared to comparable children who were not in HUD assisted housing. Just a couple other policy-related... the Congressional Budget Office has used our link data to look at Medicare spending and health-related aspects. Another example of some policy-related work working with our linked data is where people have looked at beneficiaries who receive disability insurance and what happens to them in the time period when they have to wait before receiving Medicare entitlement. And so there was a paper that compared their access to care during that waiting period. Another example is actually looking at some of the effects of folic acid fortification and looking at the health effects in older adults.

HOST: What about post-hospitalization mortality? What kind of information does that provide us?

LISA MIREL: Yeah that's a really good question. So that ties into the linkage with the National Hospital Care Survey data and that's that survey where we have hospitals that are sampled but then we get patient records for that year that that they may have been seen at the hospital. And what we've done is we have linked those patient records with the National Death Index. And so we get information about what happened. Let's say they were seen for some condition in the hospital and then by linking to the NDI we can see - did they die in 30, 60 or 90 days post hospitalization? So a really great example of how these data have been used was a recent National Health Statistics Report that came out that looked at opioid-involved emergency department visits and then they looked at the subsequent mortality. And the authors found that for approximately 22% of the patients who died of an opioid-involved drug overdose that their last emergency department only visit or hospitalization was opioid-involved. So I think it's just a really interesting way that you can kind of, you know what's going on at the hospital, you have that information, the hospitals have that information, but what they don't know is what happens when that person leaves the hospital. And so by linking these sources it's a really great way to understand what might be going on with that population and perhaps developing ways that they can help target or try to prevent some of those deaths.

HOST: Join us next time for part three of our conversation with Lisa Mirel on data linkage activities at NCHS.