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. Has the data linkage program made any changes to further address the implementation of the Evidence Act of 2018?

LISA MIREL: Yeah - we actually have. In light of the Evidence Act, we've been looking for new and untapped sources of data that we can link to to help build evidence for evaluation of public health programs and policies. And one example of that is our upcoming linkage that we're going to be conducting with the survey data in the Veterans Affairs data. We are also doing a lot more outreach than we ever had before to really try to make people aware that NCHS has been linking in building data to support evidence for like the last two decades and I think it's really important that the linkage work it's not new for NCHS but what is new and exciting is this requirement within the Evidence Act to really expand the accessibility of the data. So our program is now actively working on new projects that will help expand the accessibility of the linked data, including looking into creating fully synthetic linked data files and possibly even an interactive data analysis tool that will be hosted on the website.

HOST: So what are some of the data linkage resources that are available for outside researchers?

LISA MIREL: So we have we have many different types of resources. Most of the linked data are only available through the NCHS research data center or with our partners within the federal statistical research data centers and most of them are restricted because as soon as we start combining the sources - you know kind of goes back to what I was talking about earlier in terms of protecting privacy - so we have to really ensure that once all these sources are put together that we're still protecting the confidentiality of our survey participants. But we really do want to make data more available for researchers because we recognize how important it is to be able to access the data. So we actually do release a couple of public use files. We have our public use linked mortality files that are available on our website for researchers and they can link the linked mortality file with the public use survey data and really get… great information in terms of looking at … having information from the survey and then subsequent outcome of mortality. The file has a limited number of causes of death – it is not as broad as we have in our restricted data - but it gives researchers definitely some understanding of what might be happening with their population and then if they are interested in kind of digging deeper, looking to more causes of death or looking at smaller subgroups, they can do that within the research data center. And the last public use file that we put out is what we call our feasibility files and we make these available for researchers for when they may be drafting up their research data center proposal. And again, they can link the feasibility files with the public use survey data. And the feasibility files have an indicator about whether or not that survey participant matched to the administrative record. So they definitely get a sense of sample size - they can look at particular subgroups to see is this, are they going to have the power to do the analysis that they want to do and they can do a lot of that leg work prior to coming into the research data center and even putting together their proposal. So we really also try to make a tremendous amount of information about the quality of the linkages available to researchers so that they have as much information as possible so to support accurate interpretation of their study results.

HOST: What are some future data linkage activities that will be happening down the road?

LISA MIREL: Yeah, so we have some really exciting linkages that are in the pipeline - actually starting within this next year. I mentioned several times about the linkage with the Veterans Affairs data. We're also going to be linking to the Medicaid TMSIS data. So in the past, we've linked with the Medicaid analytic extract files but CMS has actually worked with the States and they've kind of got this transformed Medicaid data set that we're going to be looking to, and we're going to be one of the first federal agencies to actually link survey data to these newly transformed Medicaid data files and we're really excited about that - it's going to be a unique resource to be able to answer a lot of different types of research questions… you know, looking at the health care policy and how it might affect the health status for Medicaid recipients. And we also have, in this next year we're going to be updating a linkage that - the last time we had done it was in 2008 - but we're going to be linking to a database that has information on end stage renal disease and there are not many people in the population that have end stage renal disease. It's about 1% of Medicare beneficiaries, but their health care costs are equal to about 7% of the total Medicare spending, at over $100,000 per year per beneficiary. So this is just a really important resource to help understand some of that spending and by having some of that self-reported information from the survey, you can really start to understand that population a little bit better. And another one that I just thought I would mention - we are in negotiations right now with Social Security Administration to link to their data on disability. You know, we have done a linkage with SSA in the past and we are looking to update that. So we are hoping, and you can put that on our list of what's coming up in the future because we do hope that that will come through. I think there's some really important research questions that could be answered using those data, including what are the population health factors that influence disability and unmet need for SSA beneficiaries?

HOST: Anything else you want to add or is there any sort of take-home message you'd like to share?

LISA MIREL: I think the biggest take home message is just really seeing how important it is to have these types of resources of combining data that each source can answer a question on its own and their wonderful in and of themselves, right? The survey data can be used to answer a plethora of issues related to health and health behaviors and you know like with NHANES with nutrition and the examination. And then, you know, having on the other side the administrative data - you can really get an understanding of what's happening in terms of who's… how many people are getting coverage. But by linking them together it just creates this tremendous resource that can answer so many questions and really feed into a lot of the evidence-based policymaking talk that's been so prominent in the past couple of years. So I just - I think what we're doing in the program it really exciting. I think there's a lot of really exciting opportunities to come… and I think for us to be able to get information from stakeholders, from researchers, to understand what would be most beneficial for them if we were to put out more types of synthetic data that would be just a wonderful thing to get feedback on. And what really is an obstacle in terms of coming into the research data center and what could we do to help researchers by having a little bit more public-facing data available.

HOST: Well thank you for joining us Lisa.