

Q&A

Pitfalls in using autism claims data: Q&A with Lindsay Shea

BY ANNA GOSHUA

14 DECEMBER 2021

Insurance claims databases are a rich resource for autism researchers: These vast fonts of procedure billing codes make it possible to track services delivered to hundreds of thousands of people on the spectrum.

Studies vary considerably, though, in how they identify autistic people using these data, which can make it difficult to compare results, according to **Lindsay Shea**, director of the Policy and Analytics Center at the A.J. Drexel Autism Institute in Philadelphia, Pennsylvania. Key differences include how many claims researchers use to verify a person's diagnosis and the time periods over which claims are assessed, according to an analysis Shea and her colleagues published on 1 September in the *Journal of Autism and Developmental Disorders*.

Shea spoke with *Spectrum* about the various ways **researchers tap insurance claims data** and how best to standardize those practices to improve health services for autistic people.

***Spectrum:* What led you to do this work?**

Lindsay Shea: Scott Grosse [a study investigator] and I have been talking for a long time about how variable the approaches are across autism research studies that use insurance claims data. The real underlying problem is that when researchers use such different criteria, it's incredibly difficult to compare findings across studies. We felt the need to put on the table all of the different approaches autism researchers take with claims data to understand if there's any room for comparability. It turned out that these approaches are so different that we needed a call to action around standardizing how we think about using claims data for these purposes.

S: What are some of the problems you found with how these data are used?

LS: For one, there are studies that use only one autism claim versus two or more autism claims on

separate dates to verify a diagnosis. The number of claims researchers use to confirm autism depends on various factors, such as the care setting. For example, inpatient claims may be less likely to have coding errors than outpatient claims because inpatient claims are typically coded by specialized staff at hospitals, which have entire units focused on billing.

The **Chronic Conditions Data Warehouse**, which is a set of algorithms commissioned by the U.S. Centers for Medicare and Medicaid Services, gives you the criteria to identify more than 60 conditions, autism among them. This database uses two outpatient claims or one inpatient claim within two years to identify an individual as having an autism diagnosis.

Reference periods, or the time period over which claims are assessed, can also vary dramatically across studies. When you're thinking about a reference period, such as the recommended two-year reference period, what happens if you have only one year of data? Does that mean that your study is impossible? The availability of data or the type of data that someone is looking at could potentially warrant a different reference period. For example, claims among children may be different than claims among adults, and there are likely a variety of reasons why researchers select different reference periods. But that heterogeneity introduces differences in who we are examining and what questions we are asking of the data.

S: How can researchers strive toward health equity using claims data?

LS: Administrative claims data are an exceptionally powerful source to examine issues of health equity, when we look at the entirety of service delivery to a population. At the population level, we're able to think about ways we can change systems to help improve access and outcomes across different populations.

There's tremendous opportunity to think about administrative claims to address the call to action around health, inequities and autism. That is further underscored by the nature of the insurance systems. For example, when we look at people who receive services or have access to health insurance through Medicaid, we know that's a group that is at higher risk for health inequities than people who have employer-based insurance, for a variety of reasons. We know the Medicaid population is more diverse and more likely to live in poverty. But it's problematic when we examine those health inequities with different criteria because it doesn't give us an accurate picture of these inequities across different populations and lifespans.

There was one large study that I was fortunate to be a part of in the United States in which we looked at claims and then examined records of individuals to ensure that the claim did indeed represent an autism diagnosis. We need more of that research. We've not examined whether things such as one claim or two claims, different reference periods and inpatient versus outpatient claims result in differences in claims data for children compared with adults, for example. We need to understand those differences so that we can apply these criteria in a meaningful way.

This work is resource intensive — in addition to the claims data, you also need clinical experts; you need access to medical records or to people to do clinical assessments. But that work is going to be a critical next step for us to be able to deploy these data as a strategy for looking at health inequities across the lifespan.

S: What are the next steps to improve and expand this type of research?

LS: First, we need studies in which we link claims data to individuals and the clinical sources of data to have the answer to the question: Who are the people in the claims data? What do they represent? If we identify an 80-year-old autistic adult in the claims data, in what way is that person different from a 20-year-old or a 10-year-old?

Another priority for this work is that we are still at the very beginning of understanding health inequities across the lifespan. Claims data are among the only sources available in which we have a large enough sample to take a close look at these issues. We can use our results to compel change by partnering with policymakers to work on systemic reform. Both of those areas are places where claims data are highly advantageous. They are cost efficient, and they present a great starting point for where we can go next with this research.

Cite this article: <https://doi.org/10.53053/YUXJ9225>