Responses to Participants' Questions

This document provides responses to questions raised by webinar participants related to the following presentations:

- **Presentation 1:** *Findings on Clinical Phenotypes and Subphenotypes*
  PCORnet: Thomas Carton, PhD, MS and Rainu Kaushal, MD, MPH
  N3C: Melissa Haendel, PhD and Christopher Chute, MD, DrPH
  Clinical cohort discussant: Rachel Hess, MD, MS

- **Presentation 2:** *Epidemiology*
  PCORnet: Thomas Carton, PhD, MS and Rainu Kaushal, MD, MPH
  N3C: Melissa Haendel, PhD and Christopher Chute, MD, DrPH
  Discussant: Josh Fessel, MD, PhD

All Presenters: Questions and Responses

Q1. *ICD-10 code U09.9, Post COVID-19 condition, was implemented October 1, 2021, but folks like me had PASC for well over a year by that point and used Z86.16 as a proxy for documenting the fact that I had a history of COVID-19 infection. At a time when medical coding is not 100% reliable nor up to date, how can researchers infer the prevalence of PASC when not everyone has the appropriate diagnostic codes in their EHR [electronic health records]? Thinking more broadly, moving forward should we update our national/international pandemic response to create ICD codes faster to improve epidemiology tracking?*

Response:

Dr. Fessel: In my experience, it isn't just having the ICD-10 code, it's also the use adoption. We see a lot of variability across sites in the use or the U09.9, Z86.16, and B94.8 codes. This is one of the main rationales for an independent computable phenotype for PASC. Basically, we can use these codes as an input feature or a training component, but the computable phenotype to identify potential PASC patients should not rely solely on these codes at this time.

Q2. If I remember correctly, these data came from private insurance. Will there be a possibility for Medicare/Medicaid data to be utilized? Is there thought on connecting to a large database like CVS?

Response:

Dr. Haendel and Dr. Chute: N3C is not limited to private insurance and includes many safety-net hospitals and rural institutions. We're also getting insurance data.

Q3. Exertional intolerance and post-exertional worsening are often seen in Long COVID. Where would that feature show up in the clusters?

Response:

Dr. Haendel and Dr. Chute: Yes, we do see that in our clusters. However, we also recognize the need for including patient-reported outcomes to our clustering. An article by Deer at al., "Characterizing Long COVID: Deep Phenotype of a Complex Condition," published in *eBioMedicine* in 2021, provides an overview of these features and their complementarity and overlap: https://doi.org/10.1016/j.ebiom.2021.103722

Q4. How well do RECOVER or other EHR-based studies cover PASC outcomes in underserved and poorly surveilled communities like rural areas?

Responses:

Dr. Kaushal: I'll speak on behalf of PCORnet. PCORnet has made an intentional effort to include networks and consortia that treat underserved communities. Additionally, in PCORnet we are actively linking to public payer data, including Medicare data. I think that part of why we see, for example, a higher burden of PASC in New York City as compared with Florida, in addition to the severity of the initial infection, is an access issue. It's something we have to be cognizant of and dive into deeply.

Dr. Carton: I'll add—and we have some data that we'll present in the Epidemiology section—that while we have a lot of academic medical centers that are the backbone of PCORnet, there are also a good number of community-based hospital systems. Consequently, we do have those patient populations represented. With the size of the data sample, we can do subgroup analyses both regionally and by socioeconomic or racial disparities.

Dr. Hess: I'm in Utah, a rural and frontier state, and both the University of Utah and Intermountain Healthcare provide care to 80% to 90% of Utahans between the two of them. The University of Utah is in N3C and both organizations are in PCORnet. For other states served—such as Montana, Idaho, and parts of South Dakota—there are pockets that are not well covered. As the questioner was alluding to, we do have more complete coverage in cities and larger population centers.

Dr. Haendel: We've worked very hard to include rural organizations, safety-net hospitals, and public institutions that serve underserved communities, and we've done a lot of analysis to look at the representation of the various demographics to ensure they are as representative as possible. One of the biggest challenges is the way we code that information within the EHR so we can do a better job of understanding the influences of environmental health, of social determinants of health, in the context of clinical outcomes. We've been working with multiple

sites to try to do that better and we've also been performing analytics to see how we can improve that coding. Things like Rural Urban Commuting Area (RUCA) codes and other measures of access have been integrated into N3C so that we can do that kind of multimodal analytics with these kinds of data assets at the same time in combination with the clinical data.

Q5. PASC shares symptomology with other post viral syndromes. Can or have EHR data differentiated any outcomes more specific to SARS-CoV-2?

Response:

Dr. Carton and Dr. Weiner*: We're applying a causal reasoning approach that explores the relative incidence of diagnoses that appear after COVID compared with the incidence of these diagnoses in similar patients without COVID. The same approaches can be used to compare outcomes in COVID patients with patients with non-COVID viral illnesses.

Q6. Of course, this all depends on what data are entered into the EHR. However, I'm noticing with many PASC patients who are feeling so weak that simply getting to a clinic feels overwhelming, and they are self-treating at home with OTCs, either on their own judgment or through advice from PASC patient support groups. Thoughts on … this data gap?

Responses:

Dr. Carton and Dr. Weiner*: This question reflects an important challenge in EHR data where the data available to analyze have already been filtered based on who presented for clinical care. This filter means we miss patients without access to care or who choose to not present for care. As the questioner recognizes, we can assess the scope of this gap by comparing our data with other sources of information on PASC, including the Clinical Science Cohort.

Dr. Fessel: Thank you so much for pointing out this critically important issue. Trying to understand what is happening for people who cannot (or who choose not to) participate in a study is always a challenge. Hopefully, RECOVER will be able to reduce the gaps through complementary approaches, including EHR, enrolling cohorts, mobile health, etc. None of these is sufficient on its own, but hopefully they're going to be quite powerful together.

Q7. Given that many Long COVID patients have outspokenly felt gaslit by their doctors over the past two plus years and that symptoms have been ignored or misdiagnosed as psychosomatic, how are researchers accounting for what is undoubtedly biased data found in EHRs?

Responses:

Dr. Carton and Dr. Weiner*: While some providers may not recognize the association of a patient's symptoms as being Long COVID, the provider should still document the array of a patient's symptoms. In our analysis, which includes billing diagnoses and symptoms that are recorded in the narrative of a clinical note, we can explore for symptoms that appear more frequently in patients with a history of COVID than in patients without COVID and recognize population-level associations that individual providers can miss.

Dr. Fessel: All data have biases, including EHR data. That's why any one single data source or data type is likely to be insufficient to address the complex issues in PASC/Long COVID, and why RECOVER is trying to bring together many different types and sources of data.

Q8. How reliable is the coding of Long COVID in electronic health records? What guidance has been issued to help providers determine the appropriate code for patients with Long COVID?

Response:

Dr. Carton and Dr. Weiner*: While our analyses respect the provider's assessment when deciding to apply the code for Long COVID, there remains a great deal of variability in its use. Consequently, our approach to designating patients as having Long COVID is based not only on the appearance of the Long COVID code, but also on finding patients with characteristics that are similar to characteristics in which the code is assigned, and in patients having diagnoses that appear more frequently in patients who have a history of COVID compared with patients without COVID.

*Dr. Mark Weiner, MD, is a Professor of Clinical Population Health Sciences at Weill Cornell Medical College. Dr. Weiner collaborated with Dr. Carton on the written responses, although he did not present at this webinar.

Webinar Slides

To request a copy of the R3 Webinar slides, please email RECOVER_ACC@rti.org.

To Learn More

- Information about RECOVER research and to volunteer for studies: https://recovercovid.org/research
- Frequently Asked Questions about RECOVER and PASC: https://recovercovid.org/faqs
- CDC information: Information for the general public and for healthcare providers about post-COVID conditions: https://www.cdc.gov/coronavirus/2019-ncov/long-term-effects/