# Responses to Participants' Questions

The overarching goal of the R3 Seminars is to catalyze a shared understanding of the research being conducted by the scientific stakeholder community within the RECOVER Consortium. The R3 Seminars and the Q&As typically feature highly scientific material intended for researchers and clinicians. For other audiences interested in these topics, a link to the National Library of Medicine's MedlinePlus medical dictionary is provided at the end of the Q&As as a resource to help in understanding the scientific terminology.

This document provides responses* to questions raised by seminar participants related to the following presentations at the R3 Seminar Patterns of PASC with Initial COVID-19 Infections and Reinfections: EHR Insights held on October 29, 2024:

- ***Insights from an N3C RECOVER EHR-based cohort study characterizing SARS-CoV-2 reinfections and Long COVID***
  **Richard Moffitt, PhD**
  **Emily Hadley, MS**
- ***Long COVID incidence proportion in adults and children between 2020 and 2023: An EHR-based study from the RECOVER Initiative***
  **Richard Moffitt, PhD**
  **Hannah Mandel, MS**
- **Discussant: Ravi Jhaveri, MD, FPIDS, FAAP**

* Responses may have been edited for clarity.

## All Presenters: Questions and Responses

### Q. Are you able to offer any more granularity about age? So many of the patients that we see who have persistent symptoms after other viral infections like mononucleosis or flu or other things tend to be adolescents. Are we able to dig down to see if those are primarily adolescents, or are we still seeing that same proportion of 4% even into the younger ages?

**Response:**

**Dr. Moffitt:** I don't think we have that prepared to answer you today, but we can certainly revisit the data to answer that question relatively quickly at this point. I'm not aware that we've cut it up by age groups unless Ms. Mandel is remembering something that I'm forgetting. But yes, we have the ability because this is all retrospective analysis to do subset analysis in adolescents versus younger kids.

I see one other question in the chat, which is basically a follow-up on your age question about older adults. And one thing that's interesting in these electronic health records (EHR)-based studies—when you look at patients who are much older, like elderly patients that have complex pre-existing conditions, sometimes we can start to see a paradoxical effect where things that we would call symptoms of Long COVID stopped getting coded in the EHR. There are other reasons that those things might be present in a patient's record, so they're less likely to be recorded. For example, if there's an elderly patient with diabetes that's uncontrolled and has very high A1C levels, it's much less likely that a physician is going to code fatigue in their chart. So we see this sort of paradoxical effect where the sickest or the most vulnerable patients actually appear to have a little bit less Long COVID as a result. We try to avoid those whenever possible, but in general, whenever we slice up the patient population, unless otherwise noted, we see the general same effects in middle age versus older adults.

## Q. What can you say about the rates or likelihood of PASC among immunocompromised patients? Can we pull those patients out from the groups you mentioned?

**Response:**

**Dr. Moffitt:** Yes, there's a current subgroup that's looking at HIV positive patients, and there's other work throughout the consortium which I'm less familiar with that is looking at other immunocompromised groups.

## Q. How do you account for those infections that occurred at home where the individual did not seek medical attention, but subsequently patients may present for a visit for Long COVID-like symptoms? How have you managed to adjust your analysis to account for those patients?

**Responses:**

**Ms. Hadley:** This was definitely a question that we struggled with as part of our analysis. We didn't intentionally exclude people with home tests. We just don't get that information; people might test at home and not tell anyone what the result of their test is. So, one of the pieces that we aim to focus on—particularly with our understanding of the severity of first infection versus severity of reinfection—was to make sure we were especially focused on those groups who had a more severe first infection because we considered them more likely to have interactions with the healthcare system that might result in them documenting a future test, particularly because they might be engaging in a healthcare setting more often. And you'll notice that we put asterisks on a lot of our findings where we're really not talking about population level incidence here.

We are focused on people who have been interacting frequently with the healthcare system. We have also considered doing specific timing of our analyses. We are more confident that we are capturing more positive tests from earlier in the pandemic rather than later. We've done some investigation into simulating what might've happened if we had more tests later in the pandemic. I think that somebody else had mentioned the tests we're

likely missing are for people who have mild cases, which makes us more confident in some of our findings related to more severe outcomes and therefore less confident in some of our findings related to mild experiences.

**Dr. Moffitt:** Yes, I think Ms. Hadley said it well. And one other answer to this is that we have a limited amount of unstructured data from clinical notes. We don't have the free text notes. We have actually extracted pieces of information there, and we see that, for example, with notes, we can reverse that paradox I talked about earlier where certain symptoms of Long COVID are undercoded, but they're still there in the notes. It might be noted that the patient had a cough somewhere in the notes, but that's not what gets coded for and billed. But in general, with all these things that EHR data are missing, our strategy is to focus on the questions that we can answer well and try not to do bad science on questions that we don't have business answering, for example, around home testing with EHR data.

## Q. Could you discuss the impact of treatment for COVID symptoms?

**Response:**

**Dr. Moffitt:** Yes, I think the best way to answer this is to say that those kinds of studies are underway or have recently concluded, and might be the topics of upcoming R3 seminars. But yes, we have looked at different treatments and their effects on Long COVID. And as I answered in the chat here that EHR cohorts are a great way to sort of pre-test clinical trials. The question was about use of off-label drugs, and these data are a great way to see any evidence of efficacy before we might go launch a trial. Also, it can help inform how that trial could be designed.

## Q. In a previous R3 seminar, you covered some of the papers on the performance of vaccine in preventing symptoms of PASC subsequently. Are there any updates or new findings in this area?

**Response:**

**Ms. Mandel:** For the analysis I discussed, we definitely have caveats about the completeness of the vaccination data. I think around a quarter of adults had at least one COVID vaccination documented prior to their index event, and that number was lower for pediatric patients. So, if somebody does have a documented vaccination, I think we're confident they actually had a vaccination, whereas if a patient has an absence of documented vaccinations, we're really not sure what to conclude about that. We were careful with the wording there and we didn't look at a number of doses, type of vaccination, or anything like that. In our adjusted analysis, we did find previously vaccinated adults had a slight but significantly lower risk of Long COVID compared to unvaccinated patients. Given the limitations of our data, it's hard to make a conclusion.

## Q. As you think about the reinfection data and subsequent exposure to newer variants, do you feel like we have some data to perhaps model how these vaccines would perform if we used them more readily and more people got boosters?

**Responses:**

**Ms. Hadley:** That's a great question. It's interesting you bring that up because we were actually just this past month doing a booster analysis as part of some separate query work for NIH and N3C. This allowed us to see we have a fair amount of booster data in N3C. We haven't looked at the relationship of the boosters with reinfections and Long COVID yet. I think that would be a future analysis. So far in our reinfections paper, again, it's not causal, but we did find that the percentage of people with reinfections had a lower proportion of people who are vaccinated than the percentage of people without reinfections. So that would suggest that we might explore that further.

Some of our previous vaccination work is being incorporated into our next round of causal reinfection analysis, but we still have some of the EHR research limitations here though, right? Both related to who we get vaccination data for, who we get booster data for. We can't be confident that a lack of booster data means the person didn't get a booster. It might just be that they got a booster at a different facility that we didn't get that information for. So, I put some pretty big asterisks on it. But we do have a fair amount of booster data in N3C. Dr. Moffitt, I don't know if you have more to add there.

**Dr. Moffitt:** It's not about vaccination and boosters, but we do have evidence—if you remember back to the first figure, Ms. Hadley showed that the frequency of reinfection, there is a sort of one or two-wave lag protection effect. People that are getting reinfections are people whose previous infection was earlier, but that protection goes away relatively quickly. We don't have the analysis to back it up quite yet, but that may be one reason that there's a similar effect with vaccinations.

## Q. Is the incidence of Long COVID changing as we get further into the pandemic based off of your studies or just in observations even outside of your studies? If so, could this be related to vaccine uptake and/or repeated infections?

**Responses:**

**Ms. Mandel:** I think there are a lot of temporal factors that could be involved in what we're seeing. I mentioned some of them at the end of the presentation. We did not see a sizable decrease in Long COVID when we looked at those secular trends over time. So, although those temporal factors could be playing a role, maybe what we're seeing is not related to that. It could be, again, increasing recognition and documentation; there has been an increase in healthcare utilization, increasing awareness, multiple infections over time—which again, we did not look at for this, but there are a number of factors.

**Dr. Moffitt**: Another way to ask that same question is: Does the subsequent reinfection give you another or boosted chance at getting Long COVID? And that's the follow-up we're currently working on based off of Ms. Hadley's work that we've shown today. But the simple answer is that every reinfection is a new encounter which gives a non-zero chance to experience Long COVID.

## Q. One concern is that something being labeled as first infection can actually just be the first documented infection, and so this could undermine the finding of lower or Long COVID prevalence in the reinfected group. Have you considered this limitation at all in your paper?

**Response:**

**Ms. Hadley:** This is a great question. It was less of a concern with our initial paper just because early in the pandemic, there weren't that many home tests available. We have more confidence in many of the tests that we have during the ancestral period and the Alpha period, and some of the early Delta period as well. It's definitely become more of an issue as we've gotten later in the pandemic. We don't know how many people still have yet to have experienced a COVID infection. When we have an initial infection that pops up much further out than the initial pandemic period, we are curious if we might've missed some initial infections. It's a hard limitation to overcome.

I think that Dr. Moffitt pointed out earlier, it's possible we might be able to glean some information from the notes where a clinician wrote, this is actually the third time this person's had this infection, even if it's the first time it's documented in their record. But even that might be somewhat unlikely. It's a limitation of our research, especially as we move further out from the initial date of the pandemic. Again, it goes back to that piece of we're very cautious about how we generalize our findings and are really focused on people who are likely to interact with the healthcare setting and have a rather robust record.

## Q. Are there any potential explanations for why we're seeing the average age of reinfection is younger than the initial infection?

**Responses:**

**Ms. Hadley:** We had a couple of different suggestions for this in our paper. One of them was related to behavior. We thought that younger folks may be more likely to engage in behaviors like going to work or engaging in daily life settings that perhaps other individuals might be deferring to other people. There was some evidence that suggested early on in the pandemic that younger folks may be less likely to wear masks or take part in some of the other recommendations to prevent the spread of COVID or might just be less likely to be concerned about getting COVID again. Therefore, they might not be taking the same precautions that folks who were older in age were

taking. So those were our thoughts around likely increased risk of exposure to second infections and then considerations too related to vaccination that we covered earlier. Dr. Moffitt, other pieces?

**Dr. Moffitt**: I think the only thing that maybe you didn't touch on was a lot of folks in the working population age . were surveilled as part of their jobs. There might just be better selection for folks that have more testing available even for asymptomatic testing.

## Q. How might the increased recognition of Long COVID over time impact the reliability of longitudinal trends observed in the data?

**Response:**

**Dr. Moffitt:** I would say that as recognition increases over time, we basically have a monotonic increase in information and all of the computable phenotypes and machine learning approaches that Ms. Mandel mentioned are at our core based on this Long COVID diagnosis. The more that that happens, the better for us data analyst types. One thing that's going to be perhaps different over time is that the type of people who are going to get those diagnoses might change. We have other work from ourselves and colleagues that has taken a deeper dive into the difference between the population that actually gets a diagnosis versus the general population of COVID patients, also versus the population that the models suggest could have a Long COVID diagnosis but don't have it documented.

## Q. Are you aware of other evidence that confirms your findings related to lower albumin levels surrounding COVID infections? And do you know if this is of interest to trialists as their potential therapeutic target?

**Response:**

**Dr. Moffitt:** There's a difference between a significant difference and a clinically significant difference, and the differences that we see are statistically different, but it's unclear that the magnitude of the change would be clinically relevant,—and in fact, it might take a whole handful of these several biomarkers. If you combine them all together, that might be enough information to start to look at a risk factor biomarker panel or even a diagnostic factor biomarker panel. But albumin alone is never going to be important enough to really make that decision. It's going to have to be one of many pieces of evidence.

## Q. Do you adjust for the treatment type received in your adjusted analysis for Long COVID?

**Response:**

**Ms. Hadley:** I should clarify that this is actually not an adjusted analysis. This is just an observation of how frequently we are seeing Long COVID diagnosed following an initial infection versus reinfection. Our current analysis that we are working on publishing will include a series of adjustments and we'll have more of a causal interpretation, but this was really to understand if we see Long COVID diagnosed among people with reinfections in our data. And it actually doesn't even include the computable phenotype that Ms. Mandel discussed. We'll be including that in the next round of analysis as well. But I would say this was very much meant as a preliminary "can we do research related to reinfections and Long COVID with this work?" question.

## Q. How does this research incorporate findings that polymerase chain reaction (PCR) testing can be inaccurate in identifying COVID cases in children, and how might that impact the study's results?

**Responses:**

**Dr. Moffitt:** It's always a good question to consider how our inclusion criteria are working, conspiring to help us or hinder us in certain conclusions. I think the best answer is that we really try and focus on the pieces of the analysis that we feel are good and sound. And if we're underestimating diagnoses in a certain population based on the testing, unless we have solid evidence that this is a non-random phenomenon, we hope that generally it washes out in the large numbers. Presumably there would be some relationship to the negative tests also being related to less severe disease perhaps related to differences in viral load. In this way, it would just lead us back to the same sort of bias about not being able to capture home tests or mild symptoms.

**Ms. Mandel:** Just to add onto that, we also included a number of different inclusion criteria. So not just testing. We looked at diagnoses including acute COVID and Long COVID, as well as MIS-C for children. We also looked at IgG test results. If patients had indication of one of those, we thought they must have been COVID positive at some point. If they had no other indication of an acute infection, we tried to impute the date of their acute COVID infection for the purpose of this analysis. That's in more detail in the pre-print, but we tried to cast a broader net.

**Dr. Moffitt:** Whereas in contrast in the reinfection work, we were much more strict in terms of defining when that reinfection occurred because we wanted to be very aware of the fact that sometimes diagnoses are carried forward in EHR. So just seeing the COVID diagnosis, it's not exactly clear when doing the analysis if that was a new COVID diagnosis or it is the physician mentioning that COVID happened sometime before. We really leaned heavily on the actual test results for determining reinfection.

**Dr. Jhaveri**: I think there are so many variables too that go into the test that have nothing to do with the performance of the test. You obviously have to have a high-quality sample, and for young kids who are screaming

and crying, an adequate sample sometimes is a real challenge. That in and of itself can affect test performance. Similarly, we have many challenges with patients who persistently test PCR positive, and the result, it doesn't really impact or isn't relevant to the current clinical status that's going on. So I think we are left with making our best guess on both sides of the equation. Those caveats come into play whenever we're talking about a test, in particular, COVID PCR testing.

## Q. Do you think that the results from your studies could be used to improve documentation for medical professionals with patients who have Long COVID?

**Responses:**

**Dr. Moffitt:** Yeah, I hope so. Maybe not these two studies that we talked about, but other work from our groups and our colleagues. If you look just back at the beginning, the computable phenotype definitions where we're really just trying to find evidence of Long COVID in EHR, all of those come with lists of diagnoses which are associated with Long COVID. And I think the new definition from the National Academies of Sciences, Engineering, and Medicine (NASEM) is extremely inclusive and echoes a lot of the findings that our groups have found empirically from the EHR.

**Dr. Jhaveri:** I would chime in with the same hope that yes, that's part of the reason why we're in this doing this work—to try to help inform not only for patients but also for providers about the spectrum of PASC and to improve it. I think the insights that we've also had from some of our natural language processing data show that when you actually look at some of the notes, there may be many mentions of Long COVID, but in fact there's no code that's used when a patient is seen. So there are a lot of discrepancies that we think by shedding light on and doing the research to show these patterns that will help improve provider recognition and subsequent documentation that enables patients to get better care and the support services that they need. It's definitely the goal for all of us.

## Q. Are you aware of any evidence that repeat infections may prolong or increase the effects of Long COVID?

**Response:**

**Dr. Moffitt:** The duration and severity of Long COVID is unfortunately another thing that's really tricky to study in EHR. First of all, there's only that one diagnostic code. Either you have it or you don't. There's no severity encoded within that. But more than that, the first crutch that people generally go to with when trying to decide if Long COVID is resolved or not is looking for an absence of new complications that define Long COVID. We want to be really careful to avoid conflating stopping seeking care with that person's Long COVID having resolved. And from an EHR perspective, we have no way to disambiguate between someone who's fed up and stopping seeking care versus someone who's continuing to seek care and generating more data in the EHR because the care is working.

So, this one falls under the category of things we are going to avoid studying directly for now in favor of things that the data are better suited to answer for us.

## Q. What recommendations might you make to healthcare providers for monitoring and diagnosing Long COVID moving forward?

**Responses:**

**Dr. Moffitt:** I would say one of the key takeaways from the work Ms. Mandel presented is that Long COVID is still happening at a rate that is not drastically reduced from where it was at the beginning of the pandemic, with respect to the fraction of patients who are getting it versus the number getting infected. And at the same time, from the work Ms. Hadley presented, every time you get a reinfection, you have yet another chance to get Long COVID. Both of these details conspire to suggest that avoiding infection and subsequent reinfection is your best way to prevent getting Long COVID.

**Ms. Hadley:** I would just add, I think a lot of our work, when we were first putting this out, there were still a lot of questions around do people get reinfections? How often do they get reinfections? What is the reinfection experience like? I think our paper is really helpful evidence that people do get reinfections and they get reinfections more than once and that the experience might change with different variant processes and that you can still, to Dr. Moffitt's point, be diagnosed with Long COVID afterwards, but you can also still be hospitalized. You can also still have a really negative experience with this illness. I think that people continue to move past this, thinking “I've got COVID once, I'll never get it again.” That's not what we're seeing and so I think it is really worth paying attention to reinfections and educating clinicians and others that COVID is still very much out there and people still get it.

**Dr. Jhaveri:** Yeah, I would echo all of those comments and say this—Long COVID is actually a big part of the discussion I have with patients and families to encourage vaccinations or boosters. The data that's out there suggests that there is a positive effect. It's not 100%, we know that, but still as both Dr. Moffitt and Ms. Hadley and others have said, whatever impact we can have to reduce your risk overall, it's certainly worth considering.

## Q. As a patient, I am interested in "bottom up" data driven results. As such, I am interested in CURE ID, which uses FDA reports from patients using off-label medications. How can we combine efforts to help people?

**Response:**

**Dr. Moffitt:** EHR studies offer a great way to look for evidence of efficacy for off-label drugs and inform subsequent trials. We didn't speak about this type of work today, but we are doing it.

## Q. In relation to the home testing and under-diagnosed population, are there initiatives to further educate providers and third parties, such as insurance payers and short-term disability providers, on diagnosis criteria?

**Response:**

**Dr. Jhaveri:** Thank you for this question. I'm not aware of specific initiatives to address the very real issues you raise in your question. Now that the guidance about work and school exclusion for COVID-19 infections are more in-line with those in place for flu and other viral infections, providers should be less concerned about the implications of diagnosis. I think COVID fatigue is a real thing amongst providers as well as the general public.

## Q. How do the EHR cohort findings compare to findings in the RECOVER observational cohort?

**Response:**

**Dr. Moffitt:** The two cohorts are complimentary. The EHR cohorts are very large, but don't have the ability to follow-up to, for example, ask questions about potential symptoms and timing of Long COVID. As a result, documented Long COVID diagnoses in the EHR are relatively infrequent. As Ms. Mandel is speaking to now, EHR-based definitions of Long COVID are a much more sensitive way of capturing Long COVID, compared to the diagnosis, but surveys enable more precise identification of Long COVID.

## Q. How does this research incorporate the finding that PCR testing can miss more than 50% of COVID cases in children, and how might that impact the study's results?

**Response:**

**Ms. Mandel:** Just noting that our inclusion criteria for children included evidence of MIS-C and positive nucleocapsid IgG test results to capture more patients. If they had only this type of evidence or a diagnosis code indicating Long COVID, but no documented acute infection, we included them and imputed the date of acute COVID infection for this analysis.

## Q. Why did you exclude patients positive by home tests? Are you not excluding an unknown number of patients in your analysis?

**Response:**

**Dr. Moffitt:** We know that positive home tests are not frequently captured in the EHR, which is what this analysis is based on. We do not deliberately exclude these data, but we acknowledge the limitations associated with missing these data.

### Q. Not being able to include home testing positives undercounts the total COVID positives, and these missing people might be less likely to get Long COVID because their cases were less severe, which is associated with lower Long COVID risk. Thus, the excess number might be overestimated.

**Response:**

**Dr. Moffitt:** Agreed. Unfortunately, we don't have the data in the EHR to answer this conclusively, but the RECOVER observational study should be able to shed more light on this question.

### Q. Is there an ICD code for Long COVID?

**Response:**

**Ms. Mandel:** Yes, U09.9, but this was not available at the beginning of the pandemic.

### Q. So, U09.9 is the only current ICD 10 code?

**Response:**

**Dr. Moffitt:** There's one legacy code which we saw being used (B94.8), especially before U09.9 was available.

# Webinar Slides

To request a copy of the R3 Seminar 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/covid/long-term-effects/
- For medical/scientific terminology: https://medlineplus.gov/healthtopics.html