Transcript

Dr. Beth Linas:

Good afternoon, and welcome to the RECOVER Research Review, our R3 seminar. My name is Beth Linas, and I'm an infectious disease epidemiologist with RECOVER Administrative Coordinating Center and the moderator of today's seminar. The goal of the seminar series is to catalyze a shared understanding of the research within the RECOVER Consortium. I want to start by thanking everyone who submitted questions in advance. Please submit any questions that arise during today's presentation using the Q&A feature in Zoom.

After the presentation, we will answer as many questions as many questions as possible. A Q&A document will be posted with the recording of the seminar on recovercovid.org. It will include the answers for submitted questions relevant to today's presentation. Questions about other scientific topics will be addressed in future seminars and answers to broader questions about RECOVER will be available in the FAQs at recovercovid.org, and as a reminder, we cannot answer individual questions about clinical care.

The topic of today's seminar is patterns and prevention of long COVID, findings from RECOVER EHR cohort studies. Our presenters today are Dr. Yongkang Zhang, Dr. Haneih Razzaghi, and our discussant is Dr. Ravi Jhaveri. Dr. Yongkang Zhang is an assistant professor in the Department of Population Health Sciences at Weill Cornell Medical College. His research uses large-scale healthcare data to understand patient health system and social characteristics associated with healthcare utilization, quality, and outcomes with a special focus on racial, ethnic, minority patients, and socially vulnerable patients.

Dr. Zhang is one of the core investigators in epidemiology/health services research component of the PCORnet EHR Hub under the RECOVER Initiative. Dr. Hanieh Razzaghi is the Director of Analytics for the Data Coordinating Center for PEDSnet, where she leads work on data integration, advanced data quality assessment, and clinical analytics. Her primary research interest is in secondary use of clinical data to better define health status in children and to improve the quality of healthcare. Her research focuses on data quality assessment that accounts for the analytic uses of data and on effective methods for automated phenotyping and analytics in large datasets, and Dr. Ravi Jhaveri: is Division Head for Pediatric Infectious Diseases at the Ann & Robert H. Lurie Children's Hospital in Chicago and Professor of Pediatrics at the Northwestern University Feinberg School of Medicine.

Dr. Jhaveri's research spans many aspects of hepatitis C virus, with particular focus on the burden, clinical outcomes, and treatment of HCV in infants, children, and pregnant women. He currently serves on the AASLD/IDSA HCV Guidelines Panel, as well as the AASLD Viral Hepatitis Elimination Task Force. Dr. Jhaveri is a fellow of the Infectious Diseases Society of America and currently serves as the Chair of the IDSA Standards and Practice Guidelines Committee. Today's speakers will share our current understanding, the gaps in our knowledge, and how recover will contribute to filling these knowledge gaps, and with that, I will turn it over to Dr. Jhaveri.

Dr. Ravi Jhaveri:

Thanks very much, Dr. Linas. I really appreciate being here and being with all of you. Thanks for being here. I wanted to just start with a very brief introduction. So, many of you who are familiar with the RECOVER Initiative know that there are many aspects to this project and that relates to direct clinical studies, the analysis of existing clinical data, as well as the collection of samples that are available on patients with COVID and post-COVID conditions. The theme of R2 topics and papers being discussed today really focus on that second bullet point, which is the EHR-related studies that are being scrutinized as part of RECOVER. So, really, we'll be focused on that EHR and real world data that's available to look at the impact of COVID and the development, risk factors, or potential impact of vaccine and treatment on post-COVID conditions. So, with that as a brief introduction, I'm gonna hand it over to Dr. Zhang, who's going to take us through our first presentation.

Dr. Yongkang Zhang:

Thanks so much, Dr. Jhaveri. Good afternoon, everyone. Thank you for the opportunity to present results from a newly published paper about to PASC among children and adults. So, my name is Yongkang Zhang, and I'm an assistant professor in the Weill Cornell Medical College or Weill Cornell Medicine. In this study, we compared select symptoms and the conditions 31 to 150 days after COVID-19 testing between those who are tested positive versus negative among both children and adults. Next please.

So, as Dr. Jhaveri has discussed, this study is part of the NIH RECOVER Initiative, and I'm from the one of the three EHR sites where we're using large-scale EHR data to generate high-quality real world evidence to understand what are the leading symptoms and the conditions among people who were tested positive of COVID-19 path conditions symptoms. Next please. So, our team leveraged very comprehensive and high-quality EHR data from Cornell to Patient-Centered Outcome Research Network, where we receive the data from more than 40 health systems in a standardized kind of format. I mean, this number of health systems vary. So, in the beginning, where we use the few, but as time goes by, we have a much more data from a lot of health systems to study PASC or long COVID.

The data from this health systems, we'll be going through some necessary steps, such as extraction, transformation, to make them readily available for researchers, and with this comprehensive data, we have been doing a lot of activities such as verbalization query, cohort identification to support a trial, for example, and more importantly, to generate real world evidence to improve our understanding of PASC or long COVID conditions among children and adults. Next please.

Our team has four key components. In the machine learning AI component, they help you using advanced machine learning methods to identify symptoms and the conditions that can be considered as PASC. So, as a member of epidemiology HSR team, we are now reaching the output from the AI and machine learning team to understand if there's any racial, ethnic, socio-economic, or geographic disparities in PASC, and we'll also understand if there is an exacerbation of preexisting conditions following a positive COVID-19 testing. We also have two other components led by physicians and the other colleagues. They're helping do important work in terms of defining PASC based on clinical expertise and conducting query to identify all kinds of potential PASC conditions and the symptoms. Next please.

So, the study I'm presenting today, it's a newly published study in BMC infectious disease. As discussed earlier, we compared the select symptoms and the conditions between 31 to 150 days after testing between those who are tested positive versus negative among children and adults. Next please. So, this study, we started this research project in 2021. So, a lot of things motivated the research of this project, probably has been changed, but when we plan the study, what we knew by that time was, like many study reported a significant proportion of people who infected with SARS-CoV-2 developed new and persistent condition symptoms, which we call PASC or long COVID.

The incidence varied among literature. So, because of variation in definition of PASC, the data source on how PASC was measured, but generally speaking, the literature by that time reported like 10 to 50% until COVID-19 patients developed PASC condition symptoms, and this PASC condition symptoms affected a wider range of organ systems. So, pretty much each part of your body will be infected by COVID-19 infection and we also observed the certain patient groups such as older adults and those who are hospitalized. Because of COVID, they have a higher incidence of PASC.

So, there were few large-scale population-based studies we identified back two or three years ago, but those large-scale population studies, we focused on various types of population, such as Medicare beneficiaries using Medicare claims data, and also lots of evidence from US veterans generated by a very productive team, but the US veterans as we know, they have very specific demographics and health conditions which may not be generalizable to the overall population. Next, please.

In the meantime, there were fewer generalizable population studies like looking at your general adults. These studies have some significant limitations. For example, if you look at the studies published in 2021, '22, most of them focused on hospitalized COVID-19 patients, and many of them did not use a control group, which means that we don't know if those PASC conditions can be identified by this study or is it more prevalent among COVID-19 positive patients or does it have similar rates? We don't know, and many of them examine the PASC of a single organ system like respiratory or some other organ system and/or patients from specific region, and then more importantly, we found that to do studies probably failed to adjust for some potential confounders between a

COVID-19 infection and the PASC. Also, reviewing literature, we realized that overall, there was not a lot of evidence about PASC among children that are now hospitalized or adults. So, those are the motivations for us to start this project back to 2021, 2022. Next, please.

So, in this study, we introduced EHR data from PCORnet to examine if select symptoms and conditions were with SARS-CoV-2 infection among adults and children compared to a control group of people who never had a positive test. So, by selecting the terms conditions... I mean, this study is a follow-up study after an earlier study. So, the team had a genetic network open publication, which was a more descriptive analysis, and as you compare it to a more comprehensive of conditions and symptoms that are potential PASCs, and after the study, the team identified the select number of symptoms and conditions that were more prevalent among COVID-19 positive patients, but that study did not do a lot of adjustment. That's led to this study. So, we leveraged the output from a prior study to understand what would happen if we do a more rigorous adjustment using regression-based analysis? Next please.

So, this study used EHR data from 43 people on sites and those sites participated in a CDC-funded data surveillance program. Starting in April 2020, those health systems refreshed their data, and it is done monthly for patients who receive the care from their affiliated hospitals and clinics. So, those patients included anyone who had a documented SARS-CoV-2 laboratory tests, regardless of results. So, we included both positive and negative tests and also anyone had an ICD-10 diagnosis for respiratory disease and that may or may not indicate... that not limited to COVID-19 or medication use or some other treatment for COVID-19 infection. So, we included a lot of people from this health systems. Next please.

So, to be included in the study, we require that all patients should have a SARS-CoV-2 laboratory test between March 1st of 2020, and then May 31st of 2021, because we are looking at what happens after a person was tested positive. So, first, we have to know what happened to them before the COVID-19 testing. So, that's why we required first that everyone should have at least one encounter in about 31 to 18 months. So, 31 days to 18 months before their index date, which was defined as their first positive or negative testing data. So, this baseline period was used to screen what kind of conditions and symptoms each person had before their COVID-19 testing.

To understand what happened after a positive or negative testing, we additionally required each person to have an encounter in the period 31 to 150 days after their index date or after their first positive/negative date. So, we included two cohorts. For our children or adolescents cohort, each person was aged between 0 to 19 by the time orphans are testing, and the rest of people who aged over 20 years or older are included in adult cohort, and in both cohorts, we first categorized them based on if they were hospitalized following a COVID-19 testing. Next please.

For symptoms, so we examined a lot of symptom outcomes. I mean, a lot of symptoms could be considered as PASC, so we have a list of 14 or maybe a little bit more symptom conditions such as headache, shortness of breath, fatigue. So, based on the symptoms, we generate four outcomes. First is at least the one symptom, which require at least only one ICD-10 code for this symptom outcomes. We also look at the three or more symptoms, which includes at least three ICD-10 codes for the same or different symptoms. We also looked at two separate symptoms because they're helping each other reported as potential PASC, including fatigue and shortness of breath. So, we have four symptom outcomes. Next please. For condition outcomes, we looked at out of our comprehensive list of conditions such as including mental health conditions, include, for example, anxiety or depression, chronic kidney disorders, diabetes, either type I or type II, hematologic disorders, major cardiovascular events, neurological disorders, and respiratory disease. So, that each condition was examined separately as different outcome. Next please.

The key independent variable is the testing result of laboratory tests for COVID-19. So, we have the positive group based on the results including like presumably positive or detected, and the people who were defined as negative, so their results are negative or not detected. So, the positive group include anyone who had at least one positive test, and that group included people who never had a positive test or diagnosis for COVID-19. So, we defined the index date as their first positive or negative test date. Next please. So, we controlled for a lot of potential confounders or covariates in analysis. For both age cohorts, we controlled for age. Age squared to account for potential nonlinear relationship. We also controlled for their sex, race, ethnicity, and the width class. We have different cutoff for adults and the children, and also we controlled for number of encounters in the health system in the baseline before index date. Next please.

For adults, because adults have much more complicated health conditions than children, so we additionally controlled for a combined comorbidity score based on diagnosis and in the baseline, and then we control for their smoking status, including current smoker, never smoke, former or missing smoking status. For hospitalized people, because hospital people have very complicated, severe conditions, we also controlled for some variables that could represent the severity of their conditions during hospitalization, including length of hospital stay, medication use, and the mechanical ventilation during the hospitalization. Next please.

So, in terms of analysis, first of all, adults, we examined all the seven conditions, outcomes I just discussed, and each condition has their own models that we examined separately. So, we use the Cox proportional hazard model to do the time-to-event analysis accounted for the time from beginning of the post-acute period to the earliest documentation of the first diagnosis for each condition until the end of the outcome period. We controlled for all covariates, I just described in this models. So, for conditional outcomes, we look at the new conditions, which means that this person to be included in the mental health outcome. This person should not have any relevant diagnosis in the baseline. So, we only identify the people who newly diagnosed and newly developed each of the conditions following their COVID-19 tests, following after their index date. Next please.

For symptom outcomes, as you may know, those are like fatigue, headache, fever. Those are very common symptoms. It could happen to anyone. So, we did not exclude the people who had these symptoms before in the baseline because that means that we will exclude a lot of people because those symptoms are very common in clinical setting. So, instead, we controlled for like who had the symptom in the baseline. For example, if our outcome is fatigue, we identify the people who develop fatigue following their index date, and then we controlled for if this person also had a fatigue diagnosis in the baseline. So, this way, we'll keep a lot of people in our sample without losing too many patients to have enough power. For this one, we used logistic regression because we did not consider time during winter. We only looked at if the person developed a symptom following in that stage, and again, so this was conducted for both children and adults. Next please.

So, results, overall, we included about three million adults aged 20 years or older by the time of their index date, between March 1st, 2020 and the May 31st, 2021. About 10% of them, we identify them as positives based on the lab test results, and the rest of them were negative. We also identified about 675,000 unique children who aged below 20, and around 9% of them were tested positive, and the rest of them were negative. Next please. This table presents the descriptive results of each symptom or condition outcome by hospitalization status and if the COVID-19 testing is positive or negative.

So, a few highlights here. So, as you can see, generally speaking, those who are hospitalized, they had a higher prevalence or incidence of these outcomes, for example, where half adults were hospitalized with a positive test results. They developed at least one symptom versus about 40% show among those who tested negative. We also found that about 17% of hospitalized people with positive test, they developed shortness of breath. Among condition outcomes, we found there's a big difference between positive and negative in terms of respiratory disease, like about 14% of people who hospitalized, they're testing positive, they developed at least one respiratory disease versus only 7% who were negative. Among children, we found about 44% of children, who tested positive and hospitalized, developed at least one symptom conditions due to our descriptive result and adjustment. Next please.

So, those are adjusted analysis from regressions. We have two panels here. The upper panel is for adults and the lower panel is for children or younger adults aged below 20. Also, as you can see, it's very consistent with whether we observe from the descriptive table. So, among adults, those who were hospitalized, those people who were tested positive, like those orange dots, among hospitalized people, those who tested positive had a higher odds of developing any symptom, three or more symptom, fatigue or shortness of breath. As you can see, those who were not hospitalized, a positive testing is associated with higher odds of getting fatigue or shortness of breath.

Among children, as we can see, among those hospitalized, a positive COVID-19 testing was associated with higher odds of having at least one symptom and shortness of breath. Interesting that we find among those who were not hospitalized, a positive test that was associated as lower odds or decreased odds of having three or more symptoms. So, those are for the symptom outcomes. Next. For condition outcomes such as mental health, chronic kidney disease, what we found is among those who were hospitalized, a positive testing is associated with increased risk. So, those are from the Cox model. So, they adjust the hazard ratio instead of odds ratio. So, those

who tested positive had a higher risk of developing type I or type II diabetes, hematological disorders, or respiratory disease. So, the evidence is less significant among non-hospitalized people. We only find among non-hospitalized people, positive testing is associated with increased risk of developing hematological disorder, and we also found a positive testing is associated with reduced risk of some conditions. Next please.

So, this is what we found about the... To summary, we found adults with a positive test and that increased the odds of being diagnosed with certain symptoms and also were in a higher risk of being newly diagnosed with certain conditions as potential PASC between the 31 to 150 day time window following their tests, compared to those who never tested positive. We also found the hospitalized children with a positive test who also were in higher odds of being diagnosed with symptoms, including shortness of breath compared to those hospitalized children with negative testing results. Next please. Generally speaking, we found the difference in symptoms and condition following positive, negative tests were more evident among hospitalized patients than non-hospitalized patients. We also found the relatively small difference in symptoms and conditions between non-hospitalized patient who tested positive and those who tested negative. Next please.

So, that's some implications in terms of clinical care delivery and public health. So, first, given the results, it seems to tell us that the clinicians and the public health agencies should monitor for the development and the persistence of symptoms/conditions after coding and testing, especially amongst those who were hospitalized following the test, and also the higher burden of PASC symptoms/conditions of the COVID-19, especially among those with severe disease, also should encourage investment in clinical public health resources to deliver care to treat and prevent a PASC. Next, please.

So, it says there are some limitations because this is an EHR-based analysis. I mean, we only capture patient symptoms/conditions if they have an encounter with this health systems, but in reality, people could go anywhere to receive healthcare. So, there is a chance that we underestimate the real prevalence and incidence of these outcomes among these patients because... I mean, our data only captures those health systems and the hospitals and the clinics affiliated with those health systems. So, also like people who were identified as negative, they could have a test positive test in some point, like at home testing or some pharmacy that we're not capturing EHR data, but this will bias the result towards known, and also EHR data does not include the information about the [inaudible 00:27:37] person begins terminate the relationship with health system.

We only see this person had the encounter in this time period. If we don't find any information on this person, we don't know if this person just terminated or this person is healthy with all the receiving healthcare, or this person, something happened that we did not capture. For hospitalized people, I mean, we just don't know, like what's the reason for their hospitalization? So, could it be for those hospitalized that testing negative, it could include those hospitalized for some other conditions. Next please. So, this study was generously supported by CDC, and of course, the NIH funded the RECOVER Initiative. I think that's all I have for my part. So, thank you so much for your time, and I'll turn it over to the next speaker.

Dr. Hanieh Razzaghi:

Hi. Give me one second to get myself set up. All right. Hello. Thanks for joining. My name is Haneih Razzaghi, and I will be talking about the work that was recently published in studying the effectiveness of the COVID-19 vaccine against long COVID and children with some sensitivity analyses and for COVID infection using real world recovery data. Next slide. So, this work was made possible by the NIH RECOVER program, but this talk and its contents are my own and not representative of the official views of the NIH. I just want to take a moment upfront to thank all patients, caregivers, and community representatives for their continued engagement and the support and all the work that we do in RECOVER to try to improve the lives of those living with long COVID. Next slide.

So, the RECOVER program has four research aims in its study of the post-acute sequela of COVID-19, PASC, or long COVID. These include understanding how often long term symptoms occur, and who may have higher or lower risk, why these effects of COVID-19 happen, how they impact a person's long term health, and importantly, what we can do to prevent or treat long COVID. The work that we present here deals with the last of these aims, specifically with understanding the role of vaccines, that vaccines have played in preventing long

COVID, both from preventing infection in the first place and then in mitigating the symptoms associated with long COVID. Next slide.

It's been widely established in clinical trials that children who are vaccinated have fewer infections and reduce severity and infections, and that there are potential waning effects over time with new variants. Large-scale epidemiologic studies have largely replicated these findings. A recent study published in the Annals of Internal Medicine, for example, showed that vaccines reduced infection and severity of infection as measured in a four-tiered way in pediatric patients. However, very little is known about the effect of vaccines and protecting against long COVID, particularly in children. We do have some adult data demonstrating the protective effects, but for children, there's a dearth of observational studies, specifically on large sample sizes. Next slide.

So, to fill this gap, we wish to look more closely at the vaccine effectiveness in reducing long COVID in pediatric patients in large databases. This was driven primarily by the fact that long COVID as a primary endpoint for clinical trials is difficult, given that it's a rare and poorly defined entity. We therefore decided to use a large clinical database to estimate the vaccine effect, leveraging our access to the real world data on vaccines and healthcare utilization. The PCORnet RECOVER database comprises 40 clinical health systems in the US, and that served as the primary data to study this question. Slide.

So, one of the difficulties in studying the impact of vaccines on long COVID is the difficulty in defining long COVID. There are many manifestations of long COVID and we have demonstrated in several studies, that there are both syndromic features which are symptoms like fever or hair loss or chest pain, as well as systemic features which point to underlying diagnosed illness such as myocarditis, or kidney injury, and etc. The data here are from a study produced a couple of years ago, where we identify disease clusters that were more prominent in children with lab-confirmed COVID-19 versus those with tests negative for COVID-19. So, in 2022, that was possible given the general unavailability of at-home testing. Next slide.

So, since then, we've reproduced the similar findings, with updated data using a tree-based scan statistic as a mining tool for diagnosis codes that are more prominent in children either tested for or diagnosed with COVID-19, but then also children diagnosed with long COVID. So, we compared the second cohort, the long COVID analysis. We compared children with COVID-19 diagnosis, but not a long COVID diagnosis, compared to those with a long COVID diagnosis so that we can learn about the conditions and symptoms co-occurring with the long COVID diagnosis to come up with a symptom and condition feature set associated with long COVID. So, for in this example, for example dyspnea and its related symptoms, we found to be highly associated with long COVID diagnoses. Next slide.

So, based on our previous work, we developed two approaches to identifying long COVID in our database. First is a purely diagnosis based approach, where we require the presence of a long COVID diagnosis on two separate occasions in order to exclude patients who received the diagnosis as a rule out in the EHR system. However, we know that this approach will underestimate true long COVID in our population, and that the previous work has demonstrated that older patients and patients with respiratory symptoms are more likely to receive that diagnosis. So, we developed a probable long COVID outcome as well. So, this was defined as a recurring presence of diagnosis more common after COVID-19 or co-occurring frequently with a long COVID diagnosis code clustered by symptoms or organ systems.

We required the recurring presence of these diagnosis within a cluster in the post-acute period following the COVID-19 infection. So, this produced a frequency that was much closer to the prior estimates, in children, about 5%. So, importantly, these are patients who may not have recurrence of a long COVID diagnosis, but who are still suffering with new-to-them symptoms that are associated with long COVID. Next slide. So, the disease clusters are listed here and range from symptoms such as headache and hair loss to common manifestations, such as change in taste and smell, or respiratory signs and symptoms or fatigue, or as well as more acute conditions with higher short term risks, such as myocarditis or arrhythmias.

These define distinct subtypes of long COVID in children. A particular child can have one or more than one and different subtypes rise and fall over time based on the phase of the pandemic. So, in particular, the Alpha and Delta period, some more cardiac symptoms, where MIS-C was also very common then, but that's waned over time, and more respiratory symptoms have increased at the start of Delta, but really through and post-Omicron. Next slide.

So, the other challenging thing using EHR data was defining immunizations. So, we specifically used electronic health record datas as our data source. Vaccines are captured in three ways in this data source. They can be administered at the health system. They can be patient-reported, or they can be captured through a health information exchange with the institution like a vaccine registry. So, we wanted to ensure adequate vaccine capture for the geographic regions of the hospitals we included in our study. So, we compared our rates to the CDC county level vaccination rates. We selected health systems with rates greater than or equal to 60% of the CDC estimate, knowing that the CDC rates may be over estimates because of the methods of data capture.

So, for example, a person who received their two doses at two different places were counted twice in the CDC methodology. So, as a result, densely populated areas may overcount patients. We frequently observed in the CDC vaccination rates greater than or equal to 95% vaccination captures, which reflects an overestimate. Next slide. So, to compute an immunization rate for an institution, we first found all vaccinated and unvaccinated patients in the EHR who had a visit to the health system since 2020. We mapped patients to their home county based on the Census Block Group data that we have in our EHR RECOVER program and computed a weighted average based on the concentration of visits from each county to the health system. Then we compare to the CDC county rate.

The graph on the right shows the attrition of institutions after we applied our threshold for the vaccine completeness. The X-axis shows threshold cut off of the CDC rate and the Y-axis shows the number of institutions eligible for inclusion. The blue line is for analysis of the 5-to-11 age group and the red line for the 12-to-17 age group. So, we cut about half of the sites. As you can see on the graph, at the 60% mark, but given the different sizes of the institution, we actually were able to retain a little bit under 80% of the patients. Next slide. So, in our study, vaccinated patients enter the cohort after their vaccine and were matched to unvaccinated patients who had a visit at the same time.

We matched based on age group and time of visit to ensure that the secular trends of the pandemic were preserved across both cohorts. Age groups were stratified based on the immunization availability and then we conducted conditional logistic regression with long COVID as the primary outcome and a sensitivity analysis for any infection. We adjusted for sex, ethnicity, health system, comorbidity burden, and pre-exposure healthcare utilization. Vaccine effectiveness was measured as the percent reduction in the outcome. Next slide.

So, this shows the final attrition cohort showing that we start with a total of 4,418,148 patients. After ensuring suitable vaccine data and excluding institutions that did not meet our threshold and selecting patients with adequate follow-up data, we ended with a total of 1,037,936 patients who were eligible for matching. The eligibility period was defined as the period of time after which vaccines became widely available. We required at least one in person visit to the health system during the vaccine eligibility period. For the vaccinated group, this was the time of the vaccine. For the unvaccinated group, it was defined at any visit to the health system during that period. We required one contact during the baseline period, which was the 36 months prior to the vaccine or the visit to the health system for the non-vaccinated patients, and one, during the observation period, which our primary observation period is 12 months for our primary outcome and 6 and 18 months for the sensitivity analysis that I'll cover later.

So, there were a total of 480,298 children in the 5-to-11 group and 557,638 in the 12-to-17 age group. Matching was performed in these age groups to generate the final study at cohorts. One caveat to this study, we studied all children who had at least one recorded dose of vaccines, because it was more likely that they received a second dose elsewhere, or that there was a misrecording than that that they had stopped at one dose. However, we did conduct a sensitivity analysis for the children with both recorded vaccines, and it replicated our primary results. Next study. Next slide.

The cohort of eligible patients was drawn from across the US and was similar to the overall demographics of the US children. All ages from 5 to 17 years were represented, but the most frequent ages were between 9 and 15. So, this table shows the pre matched characteristics of each cohort. So, interestingly, children in the vaccinated group tended to be older and more likely to be Hispanic, Black children and white, non-Hispanic children were more likely to be in the unvaccinated group. Next slide.

Vaccinated and unvaccinated children did not have major differences in the number with chronic medical conditions using the Pediatric Medical Complexity Algorithm, the PMCA, as the metric to evaluate medical

complexity. Unvaccinated children had slightly more clinical visits, before the vaccine said the entry date, than vaccinated children. Before matching, vaccinated patients were more likely to be seen by clinicians earlier in the pandemic, which may be related to vaccine availability in the adolescent age group. Unvaccinated children were more likely to be seen for the first time after December 2021, which coincides with the rising acute illness of Omicron and other respiratory infections. So, this pronounced difference really drove our decision to match by time period in our study, to ensure that the secular trends didn't drive differences in outcomes between our two groups. Next slide.

So, results now. This graph shows vaccine effectiveness. First, I'm showing that acute infection outcome. Infection was defined as positive antigen or PCR test or specific COVID-19 diagnosis. Followup was 12 months. Since the 12-to-17-year-olds had a vaccine available to them prior to Omicron, we computed vaccine effectiveness pre and post-Omicron. So, here, we see that the vaccine was most effective in a 12-month followup among adolescents and teens in the pre-Omicron era. The overall vaccine effectiveness across age groups and time periods was greater than 50%. Next.

So, here, we look specifically at effectiveness of vaccines in preventing long COVID. Overall, the vaccines demonstrate more than 35% effectiveness in preventing long COVID, defined by our probable definitions, which was our condition cluster definition across all age groups and time periods. The effect was the highest for adolescents and teenagers, particularly during and after the Omicron period. The effect was even stronger when we looked only at the patients who received a diagnosis of long COVID. It was 41.7% overall with nearly 60% effectiveness for adolescents. We performed two major sensitivity analyses on this data. First, we wanted to observe the effect of vaccines against long COVID when the vaccine was received following a prior episode of COVID-19. The sample size was too small for the diagnosed PASC, but we observed a protective effect greater than 50% in the adolescent group and an overall effect of about 48% in the full cohort.

Second, we conducted a mediation analysis to examine how infection mediates the relationship between vaccine and long COVID. The data show that the principal mechanism for reducing long COVID is by reducing the acute risk of COVID-19 and serious infection. Next slide. This graph shows that like vaccine effectiveness on acute COVID-19, the protection wanes over time for preventing long COVID. The blue lines represent the six-month followup from vaccines. The red, 12 months, which was our primary outcome, and the green, 18 months. So, the vaccine was most effective at six months and least effective at 18 months.

The overall protective effects across all age groups was greater than 50%, six months after immunization, whereas it was 34% at 12 months, which was our primary outcome, and not statistically significant in 18 months. So, some part of this observation could be due to the impact of the changing variants and how vaccines target those or it could be due to the waning effects of vaccines over time. Next slide.

So, we have demonstrated the vaccine studies should be done using real world data that can or should be done to learn about the impact on preventing COVID-19, as well as long COVID. Our vaccine rates and long COVID rates have replicated other studies providing face validity for our methods, and most importantly, we've replicated findings from clinical trials showing protective effects of the vaccine against long COVID and real world data. The use of this data is accessible and more time-efficient than trying to conduct clinical trials for everything from scratch. So, producing these results offers some reassurance about the validity of EHR research in studying vaccine effectiveness particularly for the pandemic.

We found that our vaccines were most protective in our adolescent age group, and that the protectiveness wanes over time. At 12 months, we're still seeing protective effects, but statistical significance are less pronounced than 18 months. So, more research is needed to understand whether this is driven primarily by changing variants or the waning protection of the mRNA vaccines. Next slide. A few limitations and caveats. First, EHR-based studies are utilization based and therefore may be biased. So, we attempted to account for this limitation in several ways, including matching our cohort, adjusting for covariates, careful selection criteria and rigorous exposure and outcome definitions. Second, there are a lot of temporal changes, including differences in diagnosis and presentation of long COVID over the past few years. So, we attempted to address this by using our previous research to define long COVID dynamically and heterogeneously.

Third, as discussed on previous slides, vaccine effectiveness wanes over time, and finally, more phenotyping work is required to understand the presentation of long COVID in younger children, as they're the group with the fewest incidences of long COVID diagnosis. So, there are several next steps. First, we need to

implement some more methods to account for the potential of additional types of biases. We should also look at mediation analysis to understand the steps towards long COVID vaccines or rather the protective effects of vaccines on long COVID.

Third, the data we show are primarily for children who received the Pfizer-BioNTech mRNA vaccine, which was approved earlier for children. So, comparative effectiveness by vaccine type may be of interest to community members and those impacted by long COVID. We've observed in the study that vaccine following infections still offers protection against long COVID and this will be important for future prevention efforts. Fourth, especially as the RECOVER pediatric cohort gathers data, we'll learn how to refine our methods for identifying long COVID in EHRs and reproducing these results will be required, and finally, re-analyzing the results is important to track new viral variants and clinical practice. We should try to understand better what is driving the waning protection that we observed in the study. Thank you, and that concludes my presentation.

Dr. Ravi Jhaveri:

All right. Thanks to both our presenters. I think they're both really great examples of how we can leverage the current EHR data to try to look for patterns and obviously also detect the potential effectiveness of vaccines and potentially other therapeutics down the line. I wanted to perhaps start with one of the questions that comes up, I think, pretty often, which is as we... The pandemic obviously has taken us through many phases and both the studies highlight the challenges with doing these studies over time and comparing COVID in 2020, with 2021, and later in 2022, as we go through original Wuhan strains to Alpha and Delta strains through to Omicron and beyond strains.

I wonder if both of you could just comment a little more on how you accounted for that in the current work? I know Dr. Zhang, your analysis stops before many of the variants emerged, but how you are accounting for variant emergence and the changes in symptoms as you carry these studies forward?

Dr. Yongkang Zhang:

Yeah, that's a good question, very good question. I think EHR data is very hard to select what variant this person had because we only know if the test result is positive or negative, but there's no additional information, it's like Omicron or Alpha or Delta. So, what we did, I think probably I did not present that in master's, is to control for when this person was tested positive by using like a time. Like March 2020, April 2020, like January 2021, this could be a proxy for the variant, but not perfect because EHR data does not have such information available for researchers.

Dr. Hanieh Razzaghi:

Yeah, so we attempted to look at this by understanding or trying to encompass a wide range of symptoms because of the changing variance. I mean, the vaccines didn't come until when it did in the pandemic. So, especially in the 5-to-11 age group, there were a lot of the MIS-C, for example, would not have been that much of an interest. So, because of that, that was excluded from our definition. So, we did try to align the time periods with the disease clusters that we came up with to account for the differences that we see of presentation and then the long COVID diagnosis code obviously is underdiagnosed, like it's not used as frequently as it should, but it does offer the added benefit of clinician diagnosing and saying, "This patient has long COVID."

Dr. Ravi Jhaveri:

Yeah. One of the other things that I think comes up as we try to do these analyses, and both of you alluded to it, is the immense changes in the way people were tested over time. So, obviously, right. In the beginning, there were no testing available, and then it became that you could really only be tested at a healthcare facility or a specialized testing location, and then it became that everyone was testing at home, and now no one

tests ever, and it's only if they happen to show up and they're like, "Oh, I have the symptoms," and then somebody says, "Oh, do a COVID test," and they realize they actually have COVID. So, can you just talk a little bit about how you tried to account for the patient reporting or when a patient said that they had a positive test at home, and obviously, some of the challenges with trying to do that?

Dr. Yongkang Zhang:

Oh, yeah. I could go first though. So, this is such an important question, which I think all the EHR-based studies have this question like, "How do you know if this negative person, which you identify in the EHR truly negative?" And also you consider like capacity constraints we had in the beginning pandemic, anywhere in the country. I mean, there's a lot of problems with this issue. So, we did a few things to account for this potential bias. I mean, first, we did a sensitivity analysis by excluding patients who were tested in an early phase like first three, like March, April, May in 2020 because we know, during this time period, it was a mess because testing capacity was very limited and there's no standard of how to code a positive test like back in those days. If you'd consider, like the physicians are using a variety of ICD-10 codes to code this person has COVID.

So, we saw probably the testing results and diagnosis in the early phase pandemic are... It's less reliable compared to later phase. For many of our studies, we excluded like first three months pandemic in 2020, making sure our results are consistent, and also, we did that to make sure, like positive or true positive, negative or true negative is like... Among those who test the negative, we also make sure that those people did not have a diagnosis suspicious as COVID. We also make sure they don't receive a therapeutic for COVID, like remdesivir or some other treatment the physicians used in the early phase pandemic to treat COVID.

So, we did a lot of the sensitivity analysis to see if the results aren't consistent, but again, this issue is very hard to address, but we were able to do all kinds of analysis to make sure we did our best to address this issue.

Dr. Hanieh Razzaghi:

Yeah, and for ours, we did require for the symptom-based diagnosis for long COVID to have any kind of evidence of COVID-19. So, that could be lab or it could be just a diagnosis code that the clinician enters. So, if a patient was seen for COVID-19, regardless of testing or where or how, it would have been recorded. So, that was our kind of attempt to do that. Of course, we're not accounting for the patients who had COVID-19, but didn't report it or weren't sick enough to see a clinician.

However, we did exclude the patients who had a history. So, if they were being seen repeatedly for other things and they had a history of COVID-19, we did exclude those patients. We did not require original infection for the diagnosed long COVID group.

Dr. Ravi Jhaveri:

Thanks. I guess I would add maybe one editorial comment for the audience, which is when we think about vaccination studies of any type, and impact of vaccine, I think it's really important that we think about how we try to measure that and there are some absolute numbers that Dr. Razzaghi talked about as far as protection against infection, but I think there's also the added benefit of modifying potential illness. So, Dr. Zhang's study actually, I think, is a really nice example of how the potential impact of vaccine may be most important in causing a very severe illness to be more mild and that would still count as an infection and blunt the absolute response but still be a really meaningful impact of vaccine.

So, I just want the audience to just keep in mind. I know Dr. Razzaghi did try to catch that, but it's one of the things that as a vaccine researcher, I really try to emphasize, is that sometimes the numbers themselves don't truly capture the impact of vaccine and going from severe to even mild or moderate, could have a huge impact on a patient's risk of subsequently developing a long COVID, and while we try to account for those factors, we can't perfectly do it in these studies, but it's really important to keep that in mind.

Dr. Hanieh Razzaghi:

Yeah, and I would add that even just reporting COVID-19 to your clinician. So, when we say we see increased all infection, that I think incorporates also like a proxy severity because the fact that the patients with immunization are less likely to even pick up their phone and say, "I have these symptoms and I have COVID-19," does demonstrate the both infection overall, but most importantly, to make it severe enough to just report it in the first place.

Dr. Ravi Jhaveri:

Yeah, and obviously, I think the point to acknowledge is that in the current era right now, our vaccination rates are very, very low across all age groups, older adults of any kind with boosters, and certainly young kids, because the vaccine became available to them very late. Very few of the youngest kids have been vaccinated at all with any type of vaccine. So, I think the importance and impact of this research is, again, trying to figure out what the protective effects are and to continue to define better, particularly in children, what the impact of COVID is and what the long COVID conditions really are, and ah, let's see. Did either of you had any comments about your counterpart's presentation that you wanted to make or any questions? Not to put you on the spot, but I guess I am putting you on the spot.

Dr. Yongkang Zhang:

Yeah, I understand the study about vaccination because... I mean, as many people know, it's extremely hard to capture vaccination status in the EHR data because most people get a vaccination in pharmacy, not hospital. That information is not shared with hospital in EHR data. That's why you can see most studies about PASC did not control for vaccination status. On labs, they have very robust EHR like VA. They probably have more comprehensive information. So, I can imagine that's a very tremendous work to identify the vaccination status.

I mean, we are working on combining EHR with the state registry, which they have our comprehensive vaccination data. So, I just want knowledge the effort, like this is a very important factor to consider and it's very hard as to getting this study completed. It's great to inform policy and public health initiatives.

Dr. Hanieh Razzaghi:

One of the biggest issues we had was making sure that our vaccine data was somewhat reliable and so we actually surveyed the sites and asked, "Who is getting the health information exchange?" We did this whole data quality... We're actually going to publish that paper as well to make it more widely available and known, but we tested for association between those institutions who say, "Yes, we participate in a health information exchange with vaccine registry and the more accurate measures." So, that was a thing that we had to spend a lot of time on. I do have a question also for Dr. Zhang. I'm curious how you disentangle long COVID from prior conditions and patients with medical complexity in your study.

Dr. Yongkang Zhang:

Yeah, that's very good question. So, for condition outcomes like mental health, diabetes, chronic kidney disease, we're looking at newly diagnosed conditions. It's like we made sure there was no diagnosis for this condition prior to COVID testing. Like in this baseline, 1-month-to-18-month period, we made sure this person did not have any diagnosis for diabetes, and if we observe a newly diagnosed condition, and if newly diagnosed diabetes is much higher among positive versus negative, we can consider it as a PASC because we never know

which one's PASC. It's like that's much more likely to be a PASC if we see a newly diagnosed diabetes, much higher incident among positive compared to negative.

I think it's hard to disentangle for symptom conditions because fatigue, shortness of breaths are very common symptoms. It could happen to anyone for any condition, and also, as you may know, EHR data probably does a poor job of capturing symptom compared to capturing conditions because many physicians probably don't bother to code the data in EHR. The headache they keep, that's minor. But now, we're recording in the EHR data. So, this is probably harder to disentangle, ascertain compared to conditions. When we did [inaudible 01:03:16], we just look at who had the headache after testing, and okay, if this person also had a headache before, which would be very reasonable for some other conditions. Whereas controller for the prior headache, I think that's what we can do in the most suitable conditions. So, yeah.

Dr. Ravi Jhaveri:

I guess one of the points that I would add before we transition to the QA, and this is, I guess, perhaps a question to pose to both of you is I think, obviously, the EHR strength is identifying those new conditions, but how do you try to account for patients who, let's say, had existing anxiety or depression that was managed, perhaps with counseling or without medication, but post-COVID becomes severe enough for us or exacerbated where they need medication or further therapy or become recalcitrant? You could say the same thing about patient with diabetes or kidney disease or the like. Can you just talk about some of the challenges and trying to account for that, some of the ways you try to address that, and perhaps some of the ways you're looking at that going forward?

Dr. Yongkang Zhang:

Yeah, that's a good question. So, I mean, it has to be defined as newly diagnosed condition or exacerbation of preexisting conditions. So, we started to look into this question. So, we started with diabetes, for example. So, we looked at people who had type II diabetes before pandemic and defined a few indicators that may potentially indicate exacerbation of type II diabetes such as elevated HBA1C testing results from a well-controlled, or too poorly controlled, based on the testing results. Increased use of anti-diabetic medication, like only one medication class become two or more and also... So, we started looking into questions starting with diabetes because diabetes outcomes are easier to define based on EHR data because you have labs. You have medication prescription. For depression, anxiety, I think that will narrow some questionnaire skills. I mean, people will answer questions.

So, it's probably a bit harder to do with EHR data along with how to combine EHR with some other data. Of course, we can look at the prescription. For people with mental health conditions, we can look at if there are any change in terms of medication use before and after COVID testing data. So, that could be something we wish to explore in the future, but that is very important question. Not only newly diagnosed condition, but also the exacerbation of prior conditions after COVID.

Dr. Hanieh Razzaghi:

Yeah, so this is a very difficult question and disease exacerbation. So, there's a couple of ways to think about it. So, we have been doing, and this was a different study, but we've been doing disease focused analysis. We did one on type I diabetes, for example, showing how the A1C does actually get a little bit worse, but then normalizes over time for most patients after COVID-19, and then we are currently doing one on sickle cell. So, there's certain kinds of conditions that we are doing deep dives on to look at this for these large-scale analyses. What we do is we have a washout period, so we don't say completely new to the patient, but for example, if they weren't seen for six months for their depression because they were handling it okay, or their whatever condition it was, but then we see a new utilization, we will say that that is still counting. It's likely potentially a exacerbation. So, that's the way that we attempted to balance the question about the exacerbation.

Dr. Ravi Jhaveri:

All right. Thanks to both of you. All right. I'm gonna hand it back to Dr. Linas to take us through the audience submitted questions.

Dr. Beth Linas:

Thanks, Jhaveri. Hey, everyone. I'm going to run through some of the questions, and again, those that we don't get to, we'll try to answer, and they'll be on the website. So, I think the first question I'll start with came as a pre-question for this talk today, and it had to ask us, please discuss successes and barriers to obtaining info from EHRs for research and comment on strategies for inter-institutional data sharing, and that's to both speakers.

Dr. Yongkang Zhang:

Oh, I'm having to start first. So, yeah, this is definitely a challenge. I mean, if you consider like New York City, we have such a fragmented market. We have so many big hospitals in New York City like Cornell, Mount Sinai, NYU, Columbia, Montefiore are five major ones in Manhattan, I guess, only, and not to mention like Long Island. So, people go anywhere. For researchers, we would like to capture the comprehensive information as much as possible to track patients across different institutions to have a whole-person information, full picture. That means that only using one hospital's data is not completed. You miss lots of information about the positive testing, medication use, diagnosis.

So, we leverage this PCORnet framework. So, if hospitals are members of this network, they are able to track the same patient across different hospitals. So, which we have much better, much comprehensive information for each person in our data. Of course, I mean, because many people could also use healthcare in some other hospital, not members of this network. In the other case, we will not be able to capture information happened outside this network, and something else we have been considering, for example, using claims data, which will be much comprehensive whether you don't have lab or some other data from claims.

So, I don't think there's answer for this question. I think each project, we just identify the key data elements we need for each research question and do our best to combine data from different sources to find the most reliable, most robust, possible data for the question. So, I'll pass it to Dr. R, for your discussion.

Dr. Hanieh Razzaghi:

Yeah. So, EHR data is very challenging. So, the part of the way... I mean, we can't get around it, right? The EHR data creates what we call collider bias in terms of who gets healthcare, and therefore, we're already starting with bias. So, given that, one, we try to really account for that in the analyses that we do. So, making sure that the two cohorts where we are looking for differences in effect are similar enough or are comparable so that we can draw on observations from the data that we have. So, there's the utilization bias. There's what gets recorded and all of these things.

So, in addition to accounting for these statistically, we try to do what we call study-specific data quality analysis to understand the extent of the bias that we're dealing with or how diagnosis codes are represented or not represented across different institutions. Are they using things a little bit differently to represent the same idea? So, we really tried to mitigate the these kinds of biases by spending a lot of time upfront evaluating the different kinds of biases and what we can do to address them.

Dr. Beth Linas:

Great. Thank you so much. I think the next question I'm going to send to Dr. Razzaghi. The question is pretty long, but it says with the burgeoning number of research studies being done on long COVID, I wonder if it would be helpful for the research community to communicate to clinicians to code for long COVID in EHR. I have had classic long COVID For over a year and have encouraged my PCP to code for it in part to assist any future research done using EHRs. She has not given me the diagnosis. However, as she states there's no diagnosis despite their existing quite a few codes for it. I wonder how many clinicians feel this way and the effect it could have on many things, including on research?

Dr. Hanieh Razzaghi:

Yes, and that's a really great observation, because you can only study something if it's recorded. So, that's why we tried to come up with these data mining techniques to understand what is occurring more frequently in these groups and then generalize it to a larger population so that we can overcome these barriers. I think that clinician outreach has to be an important part of what we do and within our health systems, feeding back this learning health system model where we provide information back to clinicians.

I bet if you look at your notes, the clinician notes, they probably mentioned long COVID. So, as part of the RECOVER program, what we're doing now is going through and extracting from clinician notes observations about patients that we can't get just by the structure data. So, we're trying to get inside the mind of the clinician whether or not they diagnose officially in the EHR.

Dr. Beth Linas:

Great. Thank you. This question will be for both of you. There was actually several questions around the same topic. So, I'm just going to read one, but it speaks to several comments. Our 14-year-old daughter was a long hauler for over a year. Her main and most debilitating symptom is anxiety and suicidal depression. It's such a common symptom among all long haulers. Why was this symptom not included in the probable diagnosis disease cluster slide? With all due respect, if you do not look very closely at mental health problems caused by even mild acute COVID infection, you're missing the mark.

Dr. Hanieh Razzaghi:

Yeah. I, first, just want to acknowledge that that sounds like a very painful experience and I'm sorry for that. So, I felt it was important to address this, or I do feel it's very important to address this question. So, the mental health impact. There was a cluster for affective cognitive functioning. So, as mental health impacts functioning, we're hoping that those were captured, because they were occurring more frequently, and then two other things. So, when we do these analyses, what we've also noticed outside of kids who were diagnosed with COVID-19 was the large rise of mental health issues, just from the social impact of the pandemic in general. So, this was very difficult to disentangle from the direct impacts of COVID in general.

So, a lot of these signals actually were not prominent when we were comparing across the groups. However, that doesn't mean that it's not real. So, as I mentioned in my previous question about the notes, and the specific disease or condition-focused exacerbation is that we are trying to make mental health a priority in terms of looking at... given its impact and the way that the social impacts of the pandemic have also made it difficult to disentangle. A deeper dive is warranted into this. So, we're doing that and also with the clinician notes, using these mental health features, as well to extract from clinical notes, what's going on a little bit better. So, that is a limitation of the way that we were forced to demarcate these and label these probable conditions. So, thank you for asking that and describing your experience.

Dr. Ravi Jhaveri:

Can I just jump in for just one sec?

Dr. Beth Linas:

Oh, sure.

Dr. Ravi Jhaveri:

We have spent this session talking about electronic health record studies and some of the advantages, but also some of the disadvantages. It is important to remember that RECOVER is many things, as I talked about at the beginning, and one of the specific advantages of some of the prospective studies is to be able to ask some of these questions in a more detailed way or in a different way to try to account for some of the shortcomings of relying exclusively on EHR-based research or on lab-based research or the likes. So, thinking about the comprehensive nature of what the RECOVER Initiative is trying to do is really important and the whole initiative is trying to come at it from different directions.

Dr. Beth Linas:

Great, thank you for the reminder. Dr. Zhang, do you have any comment or I can go on to the next question?

Dr. Yongkang Zhang:

We can go next one.

Dr. Beth Linas:

Okay, great. So, I'm gonna send it to you actually. The question is, when controlling for mechanical ventilation, did you also test for the effect of the number of MV days on COVID diagnosis? I'm thinking this should have been significant, and there's a second question from someone else very similar that says, "Rather than just control for MV days, I hope you can test for the effect of MV days on long COVID diagnosis as well as on variables such as RDW.

Dr. Yongkang Zhang:

Sorry. My connection was not stable. So, I missed the first half of what you just said.

Dr. Beth Linas:

So, the first question was when controlling for mechanical ventilation, did you also test for the effect of a number of MV days on COVID diagnosis?

Dr. Yongkang Zhang:

Yes, that's a very good question. I think the tricky part is that in the EHR data, we only observe like who had the mechanical ventilation based on the procedure code, or like, "Okay, this is all the EHR data we have for this hospital encounter," and we'll observe a procedure code, a PPX code, the ICD-9 procedure code, ICD-10

procedure code for mechanical ventilation. We know this person. How did this become a ventilation? It's very hard for us to count how many days this person used mechanical ventilation because... I mean, I don't think EHR data has like a per-day record that we only know like what happened during the encounter or the medications or diagnosis or the procedures without knowing like, "Okay, probably this procedure repeated five days." So, it's hard to capture the EHR data about the agreements. It's a really important thing to consider.

Dr. Beth Linas:

All right. Thank you. So, another question for both of you. Have symptoms of long COVID been observed to worsen or reemerge following a non-COVID infection or due to other triggering factors? If so, what are these triggers?

Dr. Yongkang Zhang:

We have to look at into this, but I would think if you think about it like other hardship, everyone experienced the pandemic of social isolation, for example, which of us experienced regardless of our COVID status, especially during like first-year pandemic, that could exacerbate everyone's mental health status regardless of if you were positive or negative, and if you were thinking about it like unemployment people experienced in 2020, food access, and you don't have access to healthcare for elective purpose. So, all this hardship we experienced could exacerbate a lot of conditions, even for people who weren't COVID-negative at that time. So, that's what I think.

Dr. Hanieh Razzaghi:

So, do you mind repeating the question? I want to make sure I address the question asked.

Dr. Beth Linas:

Sure. Have symptoms of long COVID been observed to worsen or reemerge following a non-COVID infection or to other triggering factors? If so, what are these triggers?

Dr. Hanieh Razzaghi:

Yeah, so we haven't done a lot of research into this. I will say right now, one of the things that we are looking at is the effect of subsequent infections on after initial COVID-19 infection. So, for example, if you look at children with COVID-19 and another respiratory illness, is both the frequency and the severity and the types of illness, for example, RSV, is it more severe after initial infection? So, we're actually currently doing this analysis and I don't have any preliminary results yet to show, but this is something that as the pandemic goes on, we are trying to understand how COVID-19 is or is not different than other respiratory infections.

So, this, I think, is going to be an important question about what triggers long COVID, what triggers vulnerability to subsequent infections, and things like that. So, thanks for asking this question. It's an important one.

Dr. Beth Linas:

Great, thanks, and another one for both of you, or if one wants to take it. Do we know the effects acute and long term of administrating mRNA and non-mRNA vaccine on patients who are presently with long COVID? Is there an exacerbation of symptoms?

Dr. Hanieh Razzaghi:

So, what we have observed is that if you get a vaccine after initial infection, you are protected from long COVID, but I think the question is asking if you currently have long COVID, what will happen if you get the vaccine, and that is actually that we have not yet studied that. So, I don't think we have any data to back that up. I will say that anecdotally, there have been reports that there's been improvement of symptoms after the administration of the vaccine, but we have not specifically looked at this question.

Dr. Ravi Jhaveri:

I would also just add, I think that there are so few individuals who have gotten the protein vaccine. The submitter of the question is asking that it's really hard to denote any specific effects of that. So, obviously, we have millions upon millions of people have gotten mRNA vaccine, and many, many, many fewer have gotten the Novavax protein vaccine. So, it's really hard to know. Other than the general protection level that it offers, we don't really have the data yet for the downstream benefits.

Dr. Beth Linas:

Well, thank you so much for responding to so many questions. We really appreciate it. We're gonna wrap it up here. Thank you so much to our presenters and thank you to our audience for attending the seminar and engaging with the Q&A. As a reminder, a recording of today's seminar will be available on recovercovid.org within a few weeks. We will also be posting a Q&A document that has responses to the questions we received today, including some that we did not have time to address. Information about future R3 seminars will be posted on the RECOVER website. We have some exciting topics coming up and hope to see you at future sessions. Additionally, you'll see a short survey come up on your screen, which asks for your feedback on the seminar. We would appreciate if you could take a minute to fill out this brief survey. Thank you so much and have a great day.