Real-world effectiveness and causal mediation study of BNT162b2 on long COVID risks in children and adolescents

Qiong Wu,[a,b,c] Bingyu Zhang,[c,d] Jiayi Tong,[b,c,e] L. Charles Bailey,[f] H. Timothy Bunnell,[g] Jiajie Chen,[b,c] Elizabeth A. Chrischilles,[h] Dimitri A. Christakis,[i] Stephen M. Downs,[j] Kathryn Hirabayashi,[f] Aaron D. Mishkin,[k] Abu S. M. Mosa,[l] Nathan M. Pajor,[m] Suchitra Rao,[n] Hanieh Razzaghi,[f] Hayden T. Schwenk,[o] Marion R. Sills,[n,p] Huiyuan Wang,[b,c] Linbo Wang,[q] Yudong Wang,[b,c] Dazheng Zhang,[b,c] Ting Zhou,[b,c] Ravi Jhaveri,[r] Eric J. Tchetgen Tchetgen,[b] Jeffrey S. Morris,[b,***] Christopher B. Forrest,[f,**] and Yong Chen[b,c,d,s,t,u,*] on behalf of the RECOVER Consortium

[a]Department of Biostatistics and Health Data Science, University of Pittsburgh, Pittsburgh, PA, USA
[b]Department of Biostatistics, Epidemiology, and Informatics, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA, USA
[c]The Center for Health AI and Synthesis of Evidence (CHASE), University of Pennsylvania, Philadelphia, PA, USA
[d]The Graduate Group in Applied Mathematics and Computational Science, School of Arts and Sciences, University of Pennsylvania, Philadelphia, PA, USA
[e]Department of Biostatistics, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA
[f]Applied Clinical Research Center, The Children's Hospital of Philadelphia, Philadelphia, PA, USA
[g]Biomedical Research Informatics Center, Nemours Children's Health, Nemours Children's Hospital, Wilmington, DE, USA
[h]Department of Epidemiology, College of Public Health, University of Iowa, Iowa City, IA, USA
[i]Center for Child Health, Behavior and Development, Seattle Children's Research Institute, Seattle, WA, USA
[j]Department of Pediatrics, Atrium Health Wake Forest Baptist, Winston-Salem, NC, USA
[k]Section of Infectious Diseases, Lewis Katz School of Medicine at Temple University, Philadelphia, PA, USA
[l]Biomedical Informatics, Biostatistics and Medical Epidemiology, University of Missouri School of Medicine, Columbia, MO, USA
[m]Division of Pulmonary Medicine, Cincinnati Children's Hospital Medical Center and University of Cincinnati College of Medicine, Cincinnati, OH, USA
[n]Department of Pediatrics, University of Colorado School of Medicine and Children's Hospital Colorado, Aurora, CO, USA
[o]Department of Pediatrics, Stanford School of Medicine, Stanford, CA, USA
[p]Department of Research, OCHIN, Inc., Portland, OR, USA
[q]Department of Statistical Sciences, University of Toronto, Toronto, Ontario, Canada
[r]Division of Infectious Diseases, Ann & Robert H. Lurie Children's Hospital of Chicago, Chicago, IL, USA
[s]Leonard Davis Institute of Health Economics, Philadelphia, PA, USA
[t]Penn Medicine Center for Evidence-based Practice (CEP), Philadelphia, PA, USA
[u]Penn Institute for Biomedical Informatics (IBI), Philadelphia, PA, USA

Summary

Background The impact of pre-infection vaccination on the risk of long COVID remains unclear in the pediatric population. We aim to assess the effectiveness of BNT162b2 on long COVID risks with various strains of the SARS-CoV-2 virus in children and adolescents, using comparative effectiveness methods. We further explore if such pre-infection vaccination can mitigate the risk of long COVID beyond its established protective benefits against SARS-CoV-2 infection using causal mediation analysis.

Methods We conducted real-world vaccine effectiveness study and mediation analysis using data from twenty health systems in the RECOVER PCORnet electronic health record (EHR) Program. Three independent cohorts were constructed including adolescents (12–20 years) during the Delta phase (July 1–November 30, 2021), children (5–11 years) and adolescents (12–20 years) during the Omicron phase (January 1–November 30, 2022). The intervention is first dose of the BNT162b2 vaccine in comparison with no receipt of COVID-19 vaccine. The outcomes of interest include conclusive or probable diagnosis of long COVID following a documented SARS-CoV-2 infection, and body-system-specific condition clusters of post-acute sequelae of SARS-CoV-2 infection (PASC), such as cardiac, gastrointestinal, musculoskeletal, respiratory, and syndromic categories. The effectiveness was reported as (1-relative risk)*100 and mediating effects were reported as relative risks.

*Corresponding author. Department of Biostatistics, Epidemiology, and Informatics, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA, USA.
**Corresponding author.
***Corresponding author.
  E-mail addresses: ychen123@pennmedicine.upenn.edu (Y. Chen), forrestc@chop.edu (C.B. Forrest), jeffrey.morris@pennmedicine.upenn.edu (J.S. Morris).

Findings 112,590 adolescents (88,811 vaccinated) were included in the cohort for the analysis against Delta variant, and 188,894 children (101,277 vaccinated), and 84,735 adolescents (37,724 vaccinated) were included for the analysis against Omicron variant. During the Delta period, the estimated effectiveness of the BNT162b2 vaccine against long COVID among adolescents was 95.4% (95% CI: 90.9%–97.7%). During the Omicron phase, the estimated effectiveness against long COVID among children was 60.2% (95% CI: 40.3%–73.5%) and 75.1% (95% CI: 50.4%–87.5%) among adolescents. The direct effect of vaccination, defined as the effect beyond their impact on SARS-CoV-2 infections, was found to be statistically non-significant in all three study cohorts, with estimated relative risk of 1.08 (95% CI: 0.75–1.55) in the Delta study among adolescents, 1.24 (95% CI: 0.92–1.66) among children and 0.91 (95% CI: 0.69–1.19) among adolescents in the Omicron studies. Meanwhile, the estimated indirect effects, which are effects through protecting SARS-CoV-2 infections, were estimated as 0.04 (95% CI: 0.03–0.05) among adolescents during Delta phase, 0.31 (95% CI: 0.23–0.42) among children and 0.21 (95% CI: 0.16–0.27) among adolescents during the Omicron period.

Interpretation Our study suggests that BNT162b2 was effective in reducing risk of long COVID outcomes in children and adolescents during the Delta and Omicron periods. The mediation analysis indicates the vaccine's effectiveness is primarily derived from its role in reducing the risk of SARS-CoV-2 infection.

Funding National Institutes of Health.

Keywords: Causal mediation analysis; Long COVID; Pediatric research; Real-world effectiveness; Vaccination

---

Research in context

Evidence before this study
We conducted a search in PubMed from the database's inception to August 12, 2023, for studies published in English using the following terms: ("children" OR "adolescent" OR "pediatric" OR "paediatric") AND ("long COVID" OR "long-haul COVID" OR "post-acute sequelae of SARS-CoV-2" OR "PASC" OR "MIS") AND ("vaccine" OR "vaccination") AND ("effectiveness" OR "efficacy"). Among the 36 identified records, only one study quantitatively reported vaccine effectiveness against multisystem inflammatory syndrome in children (MIS-C). No studies focused on the vaccine effectiveness against long COVID or post-COVID-19 condition in children. Furthermore, although studies on vaccine effectiveness against long COVID in adults are available, their findings are inconsistent. Most existing studies in the adult population compare the long COVID following breakthrough infections in vaccinated individuals to infections in unvaccinated individuals, which only reveal the vaccine's effectiveness within the infected population. This approach does not accurately represent the true impact of vaccination on long COVID given the risk of infection is substantially reduced in the vaccinated group and introduces selection bias.

Added value of this study
Using data from the largest pediatric electronic health record (EHR) clinical research network in the U.S., our study is among the first studies to assess the real-world effectiveness of the BNT162b2 vaccines at the population level while quantifying the vaccine's influence on long COVID outcomes, either beyond or through the prevention of SARS-CoV-2 infections, using causal mediation analysis. Our study found a high overall protective effect of BNT162b2 against long COVID during Delta period, and moderate protective effects during the Omicron period. The estimated direct and indirect effects indicated that the vaccine's primary advantage in protecting against long COVID outcomes stems from its ability to reduce the risk of SARS-CoV-2 infection in both Delta and Omicron period.

Implications of all the available evidence
Our findings demonstrate the effectiveness of BNT162b2 vaccines on long COVID in U.S. pediatric population. The causal mediation analysis suggested a comparable risk of long COVID for documented infections after vaccination to those occurring without prior vaccination. Given that the causal pathways for developing long COVID are still not fully understood, our findings underscore the importance of continually prioritizing the prevention of SARS-CoV-2 infections and using vaccination as a key focus of public health policy to mitigate the risk of long COVID.

Introduction

The scientific and clinical understanding of the long-term effects or sequelae following a COVID-19 infection caused by the SARS-CoV-2 virus continues to evolve. These post-acute sequelae of SARS-CoV-2 infection (PASC), also referred to as long COVID, are

persistent, exacerbated, or newly developed symptoms or other health effects that can affect multiple organ systems (e.g., cardiovascular, neurologic, mental, metabolic, and renal).[1-6] Significant work has been focused on characterizing the complex clinical representation of PASC with the development of standardized definitions. Research concerning PASC in pediatric populations revealed a difference in clinical features and incidence rates compared to adults.[7]

The efficacy and effectiveness of COVID-19 vaccines in preventing symptomatic and severe COVID-19 has been assessed through randomized controlled trials (RCTs[8,9]) and subsequent observational vaccine-effectiveness studies.[10-17] However, our understanding of how a COVID-19 vaccine administered prior to infection impacts long COVID outcomes is still unclear. Further, research conducted to date has been mainly centered on adults and has produced inconsistent findings. Some studies suggest a significant protective effect,[18-27] e.g., a reduced risk of the diagnosis of PASC or experiencing certain PASC symptoms. Meanwhile, other studies indicate mixed effects revealing considerable variations across different age groups, various dominant virus strains, and distinct PASC symptoms,[28-32] or even suggesting counter-protective effects.[33,34] Moreover, the majority of existing studies have reported effects by comparing breakthrough infection to infections in unvaccinated individuals, which is conditional on the infection status.[18,21,24,26,28,29,35] This approach only reveals the vaccine's effectiveness within the infected population, which does not accurately represent the true impact of vaccination on long COVID, as the risk of infection is substantially reduced in the vaccinated group.[10-17,36] Furthermore, as underscored by Hernan et al. (2023),[37] when certain factors affect the risk of infection and vaccination reduces the risk of infection, the infection status becomes a collider. If these same factors also affect the risk of long COVID, conditioning on infection status can introduce collider bias, leading to a biased estimate of the vaccine's effect.

To address these gaps in our knowledge of the effectiveness of COVID-19 vaccination on long COVID outcomes among the understudied pediatric population, we designed this study among children and adolescents during the Delta and Omicron variant-predominant periods using electronic health record (EHR) data from RECOVER PCORnet Program.[38,39] It is, to the best of our knowledge, among the first studies in the U.S. focusing on studying the effectiveness of COVID-19 vaccination on long COVID within the pediatric demographic. Further, to provide a more comprehensive understanding of the vaccine's effect and mitigate the collider bias in studies conditioning on infection status, we conducted a causal mediation analysis[40-42] that quantifies both the overall vaccine effectiveness and effects through specific mediating pathways. The overall vaccine effectiveness yields a quantification of the impact on long COVID at the population level involving both infected and uninfected individuals, not only infected patients, which delivers more generalizable findings. Meanwhile, the dissection of the overall vaccine effectiveness into direct and indirect components through causal mediation analysis allows for a nuanced assessment of the vaccine's influence on long COVID outcomes, *either beyond or through* the prevention of SARS-CoV-2 infections. To strengthen the reliability of our research findings, we also conducted the proximal analysis utilizing a set of negative control exposures and outcomes,[43,44] which helped to assess the potential residual bias due to unmeasured confounders in the EHR data. Last, given the challenges of diagnosing long COVID in pediatrics, a symptom-based computable phenotype definition of long COVID was implemented.[45] Five body-system-specific PASC clusters were evaluated to account for the heterogeneous representation of long COVID.

Methods
Data sources
This study is part of the NIH Researching COVID to Enhance Recovery (RECOVER) Initiative, which seeks to understand, treat, and prevent PASC. For more information on RECOVER, visit https://recovercovid.org/. Participating institutions in this study included: Cincinnati Children's Hospital Medical Center, Children's Hospital of Philadelphia, Children's Hospital of Colorado, Duke University, University of Iowa Healthcare, Ann & Robert H. Lurie Children's Hospital of Chicago, Medical College of Wisconsin, University of Michigan, University of Missouri, Medical University of South Carolina, Nationwide Children's Hospital, Nemours Children's Health System (in Delaware and Florida), OCHIN, Inc, Ohio State University, Seattle Children's Hospital, Stanford University, Temple University, University of California San Francisco, Vanderbilt University Medical Center, and Wake Forest Baptist Health. Data were transformed to either the PCORnet or Observational Medical Outcomes Partnership (OMOP) data models,[46,47] and forwarded to the RECOVER-PCORnet Coordinating Center. For this study, we used the s9 version of the data, collected until June 2023, which comprises 6,868,813 patients.

Construction of study cohorts
We identified three study cohorts to assess both the effectiveness and mediating effects of the BNT162b2 vaccine on long COVID risks associated with various strains of the SARS-CoV-2 virus in children and adolescents in the United States. Study 1 involved adolescents focused on the period when the Delta variant was prevalent, specifically, from July 1 to November 30, 2021. Study 2 involved children and Study 3 among adolescents covered a period when the Omicron variant

was prevalent from January 1 to November 30, 2022, ensuring a 179-day follow-up period for the observation of long COVID outcomes.

To be eligible for the study, children had to be aged between 5 and 11 years, while adolescents were between 12 and 20 years old at the start of each study. Participants could not have received a COVID-19 vaccination or had a documented SARS-CoV-2 infection at the start of the study period. Moreover, to confirm their active engagement with the healthcare system, participants must have had a prior interaction (either in-person, via phone, or through telehealth) within the 18 months leading up to their cohort entry. To ensure sufficient follow-up time for documenting infections among participants, individuals in the study cohorts were required to enroll at least three months before the end of the study for the Delta variant period, and at least four months before the end of the Omicron variant period.

The selection of participants in three study cohorts is summarized in an attrition table (i.e., Fig. 1). A detailed description of the cohort construction and observation windows is available in the Supplemental Appendix Section S1.

Study variables

The intervention of interest was vaccination, in comparison with no receipt of any type of COVID-19 vaccine. Considering that the BNT162b2 vaccine constituted the majority of recorded vaccinations for this age group in our database, our study mainly focused on evaluating the BNT162b2 vaccine. Vaccine doses were determined based on the presence of a vaccine code (CVX or RxNorm) designating an administered or patient-reported dose or a vaccine description containing the string "COVID" or "SARS" with a term "pfizer". Nevertheless, a comprehensive sensitivity analysis covering all types of COVID-19 vaccines available in the U.S. can be found in the Supplementary Appendix Section S11. To ensure comprehensive vaccine data, we included only sites with adequate vaccine data capture, by comparing the vaccination coverage from EHR data with CDC vaccination statistics for each study cohort.

We defined the index date for the intervention group as the date of receiving the first dose of the BNT162b2 vaccine. For the comparator group, an index date was assigned based on a randomly chosen medical visit
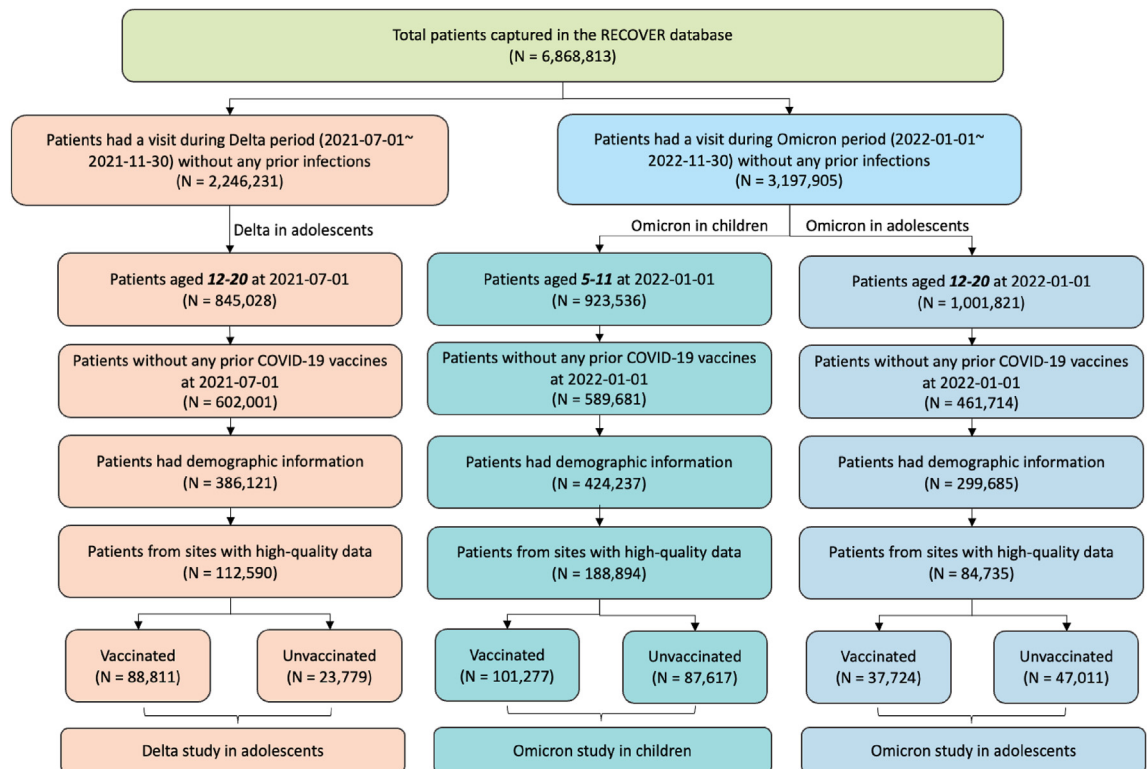

Fig. 1: Selection of participants for the three study cohorts evaluating the overall effectiveness, direct and indirect effect of the BNT162b2 vaccine on long COVID outcomes in (1) adolescents aged 12–20 years during the period when the Delta variant was prevalent, (2) children aged 5–11 years and (3) adolescents aged 12–20 years during the period when the Omicron variant was prevalent.

while ensuring the distribution of index dates for the comparator group matched the distribution in the vaccination group to control for temporal effects.[17]

The mediator in our study was identified as any documented SARS-CoV-2 infections defined by the occurrence of positive polymerase-chain-reaction (PCR), serology, or antigen tests or diagnoses of COVID-19, post-acute sequelae of SARS-CoV-2 (PASC), or multi-system inflammatory syndrome (MIS) regardless of the presence of symptoms. We defined the risk period of infections as 28 days after the index date such that participants with infections within 28 days were excluded.

The primary outcome of our study was the diagnosis of long COVID 28–179 days following a documented SARS-CoV-2 infection. Because long COVID is heterogeneous and likely underdiagnosed, we defined a conclusive or probable diagnosis of long COVID based on a computable phenotype as our primary outcome. A conclusive diagnosis of long COVID was determined by two or more medical visits indicating diagnoses of PASC or MIS. Recognizing the potential limitations and varying uptake of specific long-COVID diagnostic codes, we also identified a probable diagnosis of long COVID. This was characterized by a single visit indicating a diagnosis of PASC or MIS or the presence of a documented SARS-CoV-2 infection alongside a minimum of two long-COVID-compatible diagnoses, spaced at least 28 days apart. These long-COVID-compatible diagnoses were categorized based on previously identified clusters of codes associated with post-acute manifestations of COVID-19 in earlier research.[7,48] In addition, we identified body-system-specific PASC condition clusters, including cardiac, gastrointestinal, musculoskeletal, respiratory, and syndromic categories, as secondary outcomes given the heterogeneity of long COVID symptoms. The diagnostic criteria for these body-system-specific PASC clusters required a documented SARS-CoV-2 infection with at least two relevant diagnoses from the respective cluster, spaced a minimum of 28 days apart.

To account for potential confounding in the relation of intervention, mediator, and outcome, an extensive set of confounding variables was incorporated including demographic variables such as age, sex, and race/ethnicity; clinical factors like obesity status, a chronic condition indicator as defined by the Pediatric Medical Complexity Algorithm (PMCA), and a list of pre-existing chronic conditions; and healthcare utilization metrics including the number of inpatient and outpatient visits, emergency department (ED) visits, unique medications prescribed, and the count of negative COVID-19 tests administered prior to the cohort entry. The detailed definitions of study variables were included in the Supplementary Table S1 of the Supplemental Appendix.

Statistical analysis

We estimated the overall vaccine effectiveness and mediating effects of BNT162b2 vaccines on long COVID risks by conducting causal mediation analyses with documented infection as a mediating variable. A visual representation of the hypothesized effect pathway can be found in Fig. 2. A typical example of mediation analysis and its interpretation is available in Section S1.B of the Supplemental Appendix. The overall vaccine effectiveness was quantified by the total effect of the mediation analysis which represents a marginal effect of vaccination on the risk of long COVID. We further decomposed the overall vaccine effectiveness (total effect) into (natural) direct and indirect effects, to determine the roles of BNT162b2 vaccines in preventing long COVID outcomes. Direct effects measure the impact of vaccination on long COVID outcomes *beyond* its effect on COVID-19 infection, while indirect effects capture its impact *through* prevention of infection.

The overall vaccine effectiveness (total effect), and direct and indirect effects were estimated by implementing the weighting strategy to adjust for a large number of confounding variables. For each study cohort, we used propensity scores to mimic the treatment assignment in randomized experiments[49,50] and mediator probabilities to adjust for confounding in the mediator-outcome relationship.[51,52] For each study cohort, three logistic regression models were built to estimate (1) the propensity score as the probability of a participant belonging to the vaccination group, (2) mediator probability in the vaccination group as the probability of a participant in the vaccination group being infected during the follow-up period, and (3) mediator probability in the comparator group as the probability of a participant in the comparator group being infected during the follow-up period. We trimmed participants based on 5th and 95th percentile cutpoints of the propensity scores and mediator probabilities to stabilize weights and improve robustness.[53] We derived and implemented closed-form variance estimators for overall vaccine effectiveness (total effect), and direct and indirect effects using the empirical sandwich method.[54] Covariate balance is assessed using the standardized mean difference (SMD) before and after weighting, with an SMD value below 0.1 indicating an acceptable balance.[55]

The overall vaccine effectiveness was reported as (1- relative risk)*100 and relative risks. The direct and indirect effects of the three study cohorts are reported in relative risks (change in likelihood). The corresponding risk differences (change in incidence) are presented in Supplementary Fig. S3 of the Supplemental Appendix for the straightforward interpretations. All the data analysis was performed using R version 4.0.2.

Sensitivity analysis

We conducted a comprehensive set of sensitivity analyses to evaluate the robustness of the research findings. Sections S5–S12 of the Supplemental Appendix present sensitivity analyses for the impacts of cohort design.
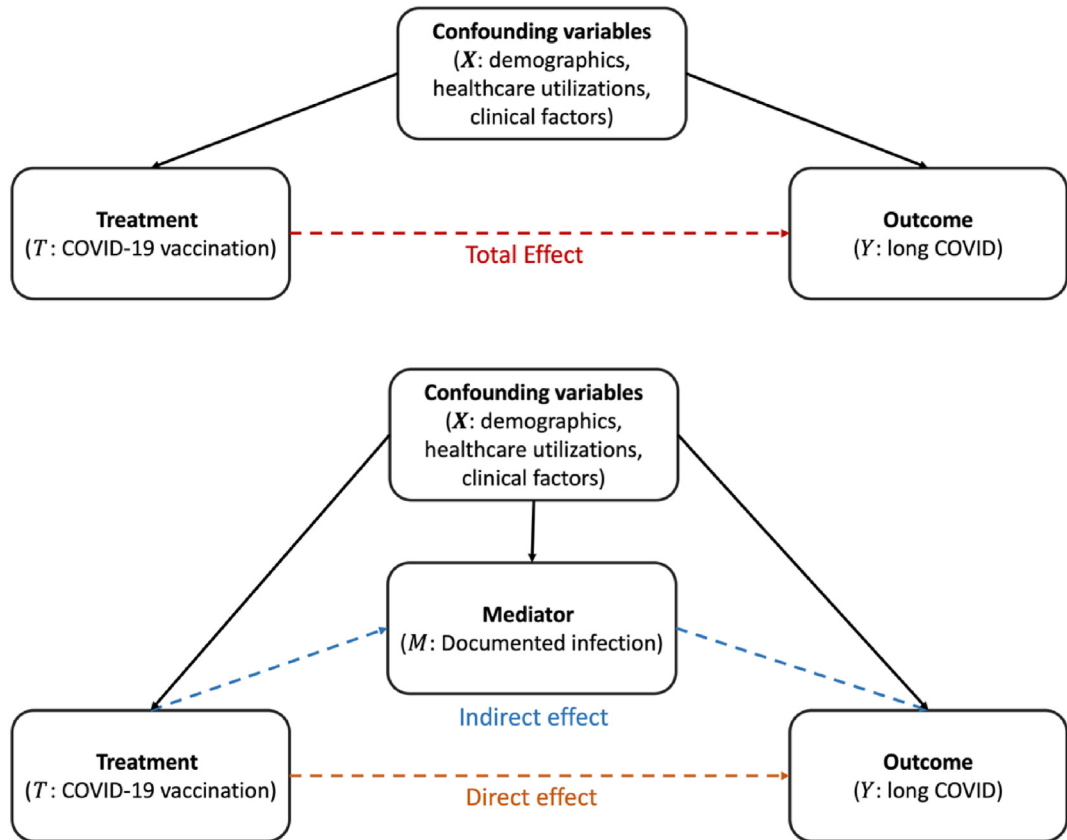
Fig. 2: Hypothetical graph of causal mediation pathways.

Since long COVID outside of MIS has been less studied in the pediatric population, we conducted a sensitivity analysis focusing specifically on non-MIS long COVID outcomes. We present the overall vaccine effectiveness and mediation effects for both Delta and Omicron studies focused on adolescents aged 12–17 years. In the Delta study among adolescents, the diagnosis of long COVID was defined based on 28–179 days following an infection in the Delta period, which could exceed the Delta period. We acknowledge the possibility of subsequent Omicron infections following an initial Delta infection prior to the diagnosis of long COVID. To ensure that our study specifically addresses long COVID as a consequence of Delta variant infections, we have narrowed the observation window for long COVID to the Delta-dominant period in a sensitivity analysis. Since the vaccination group had a notably lower percentage of patients entering through ED visits compared to the comparator group, we carried out a sensitivity analysis excluding those who joined the cohort due to an ED visit. To examine the dose-dependent effects of the BNT162b2 vaccine, we estimated the effects of a two-dose vaccination which was defined as the administration of a second dose at least 14 days before the

infection. Additionally, we provided estimates of the vaccine's effects including all available COVID-19 vaccine brands.

Sections S4, S13, and S14 of the Supplemental Appendix focused on sensitivity analyses regarding the statistical methodologies employed in the study. Beyond our primary causal mediation weighting, we present findings on the overall vaccine effectiveness and direct and indirect effects using a regression-based approach for causal mediation analysis. It's noted that the measured confounding variables may not be sufficiently rich to account for confounding. Hence, we conducted the proximal analysis using 4 pre-specified negative control variables from pediatric physicians which are believed to be not causally related to the intervention, mediator, and outcome.[43,44]

Ethics statement

This study constitutes human subject research. Institute Review Board (IRB) approval was obtained under Biomedical Research Alliance of New York (BRANY) protocol #21-08-508. As part of the Biomedical Research Alliance of New York (BRANY IRB) process, the protocol has been reviewed in accordance with the

institutional guidelines. The Biomedical Research Alliance of New York (BRANY) waived the need for consent and HIPAA authorization. Institutional Review Board oversight was provided by the Biomedical Research Alliance of New York, protocol # 21-08-508-380. This study used deidentified data with a waiver of consent.

Role of the funding source
The funders had no role in the design or conduct of the study; collection, management, analysis, or interpretation of the data; approval of the manuscript; or the decision to submit the manuscript for publication.

Results
Study population
Within the RECOVER network, we identified a total of 112,590 adolescents in which to investigate both the overall vaccine effectiveness and mediating effects of BNT162b2 on long COVID outcomes during a period dominated by the Delta variant (refer to Table 1 for baseline characteristics). For studying the overall vaccine effectiveness and mediating effects of BNT162b2 on the risk of long COVID of Omicron infections, the cohort comprised 188,894 children and 84,735 adolescents (baseline characteristics provided in Table 2). The highest incidence rate of both the long COVID outcome and documented infection was observed in the unvaccinated group of the Delta study in adolescents (i.e., 3.54 and 53.10 per 10,000 person-week). In contrast, the vaccinated group from the same study exhibited the lowest incidence rates, being 0.11 and 1.97 per 10,000 person-weeks, respectively (Table 3). In our database, the vaccination records for children and adolescents showed 89.3% for BNT162b2, 2.1% for mRNA-1273, and 8.6% for unspecified COVID-19 vaccines. Across all three cohorts, there was a minor imbalance in testing rates prior to cohort entry between the vaccinated and unvaccinated groups. However, after applying propensity score adjustments, all covariates achieved the balance between the vaccinated and unvaccinated groups, evidenced by an SMD of less than 0.1 in all three cohorts (Supplementary Figs. S2–S4 of the Supplemental Appendix).

Overall vaccine effectiveness on long COVID
Table 4 summarizes the estimated overall vaccine effectiveness and direct and indirect effects of BNT162b2 vaccine on long COVID. During the Delta variant phase, the BNT162b2 vaccine exhibited an estimated overall vaccine effectiveness against long COVID of 95.4% (95% CI: 90.9%–97.7%) among adolescents. During the Omicron phase, the estimated overall vaccine effectiveness against long COVID was 60.2% (95% CI: 40.3%–73.5%) for children. For adolescents, it was 75.1% (95% CI: 50.4%–87.5%). The BNT162b2 vaccines have shown higher effectiveness during the Delta

variant period and are of greater magnitude in adolescents compared to children.

Direct and indirect effects on long COVID
The direct effect quantifies the vaccine's impact on long COVID outcomes, separate from its protection against SARS-CoV-2 infection. Conversely, the indirect effect measures the vaccine's benefit derived from its protective role against the SARS-CoV-2 infection. We refer to Fig. 2 for a visual representation of the hypothesized effect pathway and Section S1.B of the Supplemental Appendix for a typical example of mediation analysis.

During the phase dominated by the Delta variant, the BNT162b2 vaccine showed an estimated direct effect against long COVID with a relative risk of 1.08 (95% CI: 0.75–1.55), which suggests that, outside of its preventive function against SARS-CoV-2 infection, vaccination prior to infection does not significantly modify the risk of long COVID. The corresponding indirect effect was estimated as 0.04 (95% CI: 0.03 to 0.05), underscoring that the benefit of the vaccine in preventing long COVID outcomes is largely attributable to its capacity to mitigate the risk of SARS-CoV-2 infection.

During the Omicron variant phase, for children, the direct effect against long COVID was estimated as 1.24 (95% CI: 0.92–1.66), while the indirect effect was estimated as 0.31 (95% CI: 0.23–0.42). For adolescents in the same phase, the estimated direct effect against long COVID was 0.91 (95% CI: 0.69–1.19), while the indirect effect was 0.21 (95% CI: 0.16–0.27). These findings during the Omicron phase align with those from the Delta phase, suggesting that the primary effectiveness of the vaccine in mitigating long COVID is through its protection against the initial risk of SARS-CoV-2 infection.

Body-system-focused PASC condition clusters
Table 5 presents the estimated overall vaccine effectiveness, along with direct and indirect effects on PASC condition clusters focused on specific body systems, namely cardiac, gastrointestinal, musculoskeletal, respiratory, and syndromic clusters. The findings generally align with the vaccine's overall effectiveness and mediating effects on the conclusive or probable diagnosis of long COVID.

During the phase dominated by the Delta variant, the BNT162b2 vaccine demonstrated consistent effectiveness across different body-system-focused PASC condition clusters, ranging from 90.2% to 97.7%. In the Omicron phase, among children, the BNT162b2 vaccine demonstrated the highest effectiveness against the musculoskeletal cluster at 71.5% (95% CI: 56.5%–81.1%), which was slightly lower against the respiratory cluster 44.3% (95% CI: 16.5%–62.9%) and gastrointestinal cluster 46.7% (95% CI: 18.5%–65.2%). Meanwhile, in the Omicron phase among adolescents, the BNT162b2 vaccine displayed the highest effectiveness

Delta study in adolescents			
	Vaccinated (N = 88,811)	Unvaccinated (N = 23,779)	Overall (N = 112,590)
Age			
Median [Q1, Q3]	14 [13, 16]	15 [13, 17]	15 [13, 16]
Distribution			
12	16,867 (19.0%)	3331 (14.0%)	20,198 (17.9%)
13	15,697 (17.7%)	3107 (13.1%)	18,804 (16.7%)
14	15,127 (17.0%)	3214 (13.5%)	18,341 (16.3%)
15	14,204 (16.0%)	3138 (13.2%)	17,342 (15.4%)
16	10,091 (11.4%)	2985 (12.6%)	13,076 (11.6%)
17	7580 (8.5%)	2771 (11.7%)	10,351 (9.2%)
18	4098 (4.6%)	2043 (8.6%)	6141 (5.5%)
19	2936 (3.3%)	1778 (7.5%)	4714 (4.2%)
20	2211 (2.5%)	1412 (5.9%)	3623 (3.2%)
Gender			
Female	46,650 (52.5%)	12,830 (54.0%)	59,480 (52.8%)
Male	42,161 (47.5%)	10,949 (46.0%)	53,110 (47.2%)
Ethnicity			
White	26,971 (30.4%)	12,373 (52.0%)	39,344 (34.9%)
Black/AA	16,044 (18.1%)	4833 (20.3%)	20,877 (18.5%)
Hispanic	35,147 (39.6%)	4112 (17.3%)	39,259 (34.9%)
Asian/PI	3521 (4.0%)	426 (1.8%)	3947 (3.5%)
Multiple	1521 (1.7%)	520 (2.2%)	2041 (1.8%)
Other/Unknown	5607 (6.3%)	1515 (6.4%)	7122 (6.3%)
Hospital			
A	4989 (5.6%)	3160 (13.3%)	8149 (7.2%)
B	10,894 (12.3%)	2395 (10.1%)	13,289 (11.8%)
C	7408 (8.3%)	1865 (7.8%)	9273 (8.2%)
D	3115 (3.5%)	1431 (6.0%)	4546 (4.0%)
E	998 (1.1%)	740 (3.1%)	1738 (1.5%)
F	1386 (1.6%)	643 (2.7%)	2029 (1.8%)
G	1655 (1.9%)	1167 (4.9%)	2822 (2.5%)
H	1000 (1.1%)	1345 (5.7%)	2345 (2.1%)
I	3189 (3.6%)	2959 (12.4%)	6148 (5.5%)
J	13,315 (15.0%)	1855 (7.8%)	15,170 (13.5%)
K	37,399 (42.1%)	2855 (12.0%)	40,254 (35.8%)
L	2234 (2.5%)	1376 (5.8%)	3610 (3.2%)
M	1229 (1.4%)	1988 (8.4%)	3217 (2.9%)
Entry time			
07/2021	40,073 (45.1%)	9351 (39.3%)	49,424 (43.9%)
08/2021	48,738 (54.9%)	14,428 (60.7%)	63,166 (56.1%)
Obesity			
Nonobese	27,459 (30.9%)	10,475 (44.1%)	37,934 (33.7%)
Obese	50,789 (57.2%)	10,385 (43.7%)	61,174 (54.3%)
Unknown	10,563 (11.9%)	2919 (12.3%)	13,482 (12.0%)
PMCA			
0	64,877 (73.1%)	14,509 (61.0%)	79,386 (70.5%)
1	14,988 (16.9%)	4653 (19.6%)	19,641 (17.4%)
2	8946 (10.1%)	4617 (19.4%)	13,563 (12.0%)
Negative tests prior entry			
0	73,356 (82.6%)	13,309 (56.0%)	86,665 (77.0%)
1	11,450 (12.9%)	6220 (26.2%)	17,670 (15.7%)
2	2611 (2.9%)	2213 (9.3%)	4824 (4.3%)
≥3	1394 (1.6%)	2037 (8.6%)	3431 (3.0%)

Table 1: Baseline characteristics of adolescents 12–20 years of age in the study of overall effectiveness, direct and indirect effect of the BNT162b2 vaccine on long COVID during the period when the Delta variant was prevalent.

	Omicron study in children			Omicron study in adolescents		
	Vaccinated (N = 101,277)	Unvaccinated (N = 87,617)	Overall (N = 188,894)	Vaccinated (N = 37,724)	Unvaccinated (N = 47,011)	Overall (N = 84,735)
Age						
Median [Q1, Q3]	8 [6, 10]	7 [6, 9]	8 [6, 10]	15 [13, 17]	15 [13, 17]	15 [13, 17]
Distribution						
5	14,831 (14.6%)	18,260 (20.8%)	33,091 (17.5%)			
6	14,424 (14.2%)	15,771 (18.0%)	30,195 (16.0%)			
7	14,142 (14.0%)	13,143 (15.0%)	27,285 (14.4%)			
8	14,027 (13.9%)	11,545 (13.2%)	25,572 (13.5%)			
9	14,245 (14.1%)	10,445 (11.9%)	24,690 (13.1%)			
10	14,484 (14.3%)	9587 (10.9%)	24,071 (12.7%)			
11	15,124 (14.9%)	8866 (10.1%)	23,990 (12.7%)			
12				8208 (21.8%)	6972 (14.8%)	15,180 (17.9%)
13				5260 (13.9%)	6538 (13.9%)	11,798 (13.9%)
14				4948 (13.1%)	6288 (13.4%)	11,236 (13.3%)
15				4650 (12.3%)	6331 (13.5%)	10,981 (13.0%)
16				4350 (11.5%)	5953 (12.7%)	10,303 (12.2%)
17				4224 (11.2%)	4963 (10.6%)	9187 (10.8%)
18				2631 (7.0%)	3941 (8.4%)	6572 (7.8%)
19				1894 (5.0%)	3510 (7.5%)	5404 (6.4%)
20				1559 (4.1%)	2515 (5.3%)	4074 (4.8%)
Gender						
Female	49,080 (48.5%)	41,313 (47.2%)	90,393 (47.9%)	20,617 (54.7%)	25,109 (53.4%)	45,726 (54.0%)
Male	52,197 (51.5%)	46,304 (52.8%)	98,501 (52.1%)	17,107 (45.3%)	21,902 (46.6%)	39,009 (46.0%)
Ethnicity						
White	22,763 (22.5%)	35,040 (40.0%)	57,803 (30.6%)	9532 (25.3%)	22,538 (47.9%)	32,070 (37.8%)
Black/AA	19,547 (19.3%)	17,902 (20.4%)	37,449 (19.8%)	9750 (25.8%)	9499 (20.2%)	19,249 (22.7%)
Hispanic	40,242 (39.7%)	22,020 (25.1%)	62,262 (33.0%)	13,506 (35.8%)	9389 (20.0%)	22,895 (27.0%)
Asian/PI	8136 (8.0%)	2904 (3.3%)	11,040 (5.8%)	1734 (4.6%)	1153 (2.5%)	2887 (3.4%)
Multiple	2317 (2.3%)	2419 (2.8%)	4736 (2.5%)	662 (1.8%)	920 (2.0%)	1582 (1.9%)
Other/Unknown	8272 (8.2%)	7332 (8.4%)	15,604 (8.3%)	2540 (6.7%)	3512 (7.5%)	6052 (7.1%)
Hospital						
A	5167 (5.1%)	8083 (9.2%)	13,250 (7.0%)	2183 (5.8%)	4623 (9.8%)	6806 (8.0%)
B	14,578 (14.4%)	10,424 (11.9%)	25,002 (13.2%)	5905 (15.7%)	5526 (11.8%)	11,431 (13.5%)
C	6081 (6.0%)	5947 (6.8%)	12,028 (6.4%)	1870 (5.0%)	3438 (7.3%)	5308 (6.3%)
D	3265 (3.2%)	4092 (4.7%)	7357 (3.9%)	1534 (4.1%)	3821 (8.1%)	5355 (6.3%)
E	582 (0.6%)	1502 (1.7%)	2084 (1.1%)	255 (0.7%)	1605 (3.4%)	1860 (2.2%)
F	4853 (4.8%)	2621 (3.0%)	7474 (4.0%)	652 (1.7%)	1462 (3.1%)	2114 (2.5%)
G	1026 (1.0%)	2246 (2.6%)	3272 (1.7%)	662 (1.8%)	2409 (5.1%)	3071 (3.6%)
H	1200 (1.2%)	1727 (2.0%)	2927 (1.5%)	757 (2.0%)	1618 (3.4%)	2375 (2.8%)
I	5893 (5.8%)	8899 (10.2%)	14,792 (7.8%)	2116 (5.6%)	3619 (7.7%)	5735 (6.8%)
J	10,309 (10.2%)	9539 (10.9%)	19,848 (10.5%)	3813 (10.1%)	4307 (9.2%)	8120 (9.6%)
K	39,011 (38.5%)	12,749 (14.6%)	51,760 (27.4%)	16,144 (42.8%)	6422 (13.7%)	22,566 (26.6%)
L	1895 (1.9%)	3340 (3.8%)	5235 (2.8%)	1063 (2.8%)	3103 (6.6%)	4166 (4.9%)
M	1356 (1.3%)	5821 (6.6%)	7177 (3.8%)	770 (2.0%)	5058 (10.8%)	5828 (6.9%)
N	433 (0.4%)	1142 (1.3%)	1575 (0.8%)			
O	1626 (1.6%)	2555 (2.9%)	4181 (2.2%)			
P	2700 (2.7%)	2936 (3.4%)	5636 (3.0%)			
Q	1302 (1.3%)	3994 (4.6%)	5296 (2.8%)			
Entry time						
01/2022	54,823 (54.1%)	31,111 (35.5%)	85,934 (45.5%)	17,583 (46.6%)	16,438 (35.0%)	34,021 (40.1%)
02/2022	19,133 (18.9%)	15,962 (18.2%)	35,095 (18.6%)	7425 (19.7%)	8231 (17.5%)	15,656 (18.5%)
03/2022	9420 (9.3%)	11,456 (13.1%)	20,876 (11.1%)	4136 (11.0%)	6190 (13.2%)	10,326 (12.2%)
04/2022	4891 (4.8%)	7551 (8.6%)	12,442 (6.6%)	2422 (6.4%)	4660 (9.9%)	7082 (8.4%)

(Table 2 continues on next page)

	Omicron study in children			Omicron study in adolescents		
	Vaccinated (N = 101,277)	Unvaccinated (N = 87,617)	Overall (N = 188,894)	Vaccinated (N = 37,724)	Unvaccinated (N = 47,011)	Overall (N = 84,735)
(Continued from previous page)						
05/2022	4160 (4.1%)	7982 (9.1%)	12,142 (6.4%)	2148 (5.7%)	4384 (9.3%)	6532 (7.7%)
06/2022	4565 (4.5%)	7291 (8.3%)	11,856 (6.3%)	2162 (5.7%)	3752 (8.0%)	5914 (7.0%)
07/2022	4285 (4.2%)	6264 (7.1%)	10,549 (5.6%)	1848 (4.9%)	3356 (7.1%)	5204 (6.1%)
Obesity						
Nonobese	34,372 (33.9%)	38,168 (43.6%)	72,540 (38.4%)	11,249 (29.8%)	19,806 (42.1%)	31,055 (36.6%)
Obese	55,691 (55.0%)	40,622 (46.4%)	96,313 (51.0%)	21,757 (57.7%)	21,651 (46.1%)	43,408 (51.2%)
Unknown	11,214 (11.1%)	8827 (10.1%)	20,041 (10.6%)	4718 (12.5%)	5554 (11.8%)	10,272 (12.1%)
PMCA						
0	78,261 (77.3%)	59,388 (67.8%)	137,649 (72.9%)	28,138 (74.6%)	30,080 (64.0%)	58,218 (68.7%)
1	14,888 (14.7%)	15,513 (17.7%)	30,401 (16.1%)	6016 (15.9%)	8789 (18.7%)	14,805 (17.5%)
2	8128 (8.0%)	12,716 (14.5%)	20,844 (11.0%)	3570 (9.5%)	8142 (17.3%)	11,712 (13.8%)
Negative tests prior entry						
0	72,618 (71.7%)	45,099 (51.5%)	117,717 (62.3%)	29,406 (78.0%)	25,020 (53.2%)	54,426 (64.2%)
1	19,079 (18.8%)	24,973 (28.5%)	44,052 (23.3%)	5833 (15.5%)	13,486 (28.7%)	19,319 (22.8%)
2	5893 (5.8%)	9905 (11.3%)	15,798 (8.4%)	1561 (4.1%)	4750 (10.1%)	6311 (7.4%)
≥3	3687 (3.6%)	7640 (8.7%)	11,327 (6.0%)	924 (2.4%)	3755 (8.0%)	4679 (5.5%)

Table 2: Baseline characteristics of children 5–11 and adolescents 12–20 years of age in the study of overall effectiveness, direct and indirect effect of the BNT162b2 vaccine on long COVID during the period when the Omicron variant was prevalent.

against the respiratory cluster at 87.4% (95% CI: 74.9%–93.7%), and the lowest effectiveness against the cardiac cluster, with an estimate of 27.2% (95% CI: 0.0%–63.5%).

In the analysis of body-system-focused PASC condition clusters across three study cohorts, the majority of the estimated direct effects did not attain statistical significance, except the gastrointestinal cluster in the Delta phase among adolescents, respiratory cluster in the Omicron phase among children, and cardiac cluster in the Omicron phase among adolescents. The estimated indirect effects indicate that the majority of the vaccine's protective effect was achieved through protecting against SARS-CoV-2 infection.

Sensitivity analysis

Sections S5–S12 presents the results of sensitivity analyses that were conducted to evaluate the robustness of research findings regarding the study design. Section S5 presented the results from excluding MIS cases when defining the conclusive or probable PASC, which yields consistent results with the primary analysis. Section S8 of the Supplemental Appendix outlines the estimated effects observed during the Delta and Omicron phases when focusing on the adolescent age group of 12–17 years. In the sensitivity analysis, which utilized a narrowed observation period within the Delta phase to ensure that the observed long COVID outcomes resulted from infections occurring in that phase, the results presented in Section S9 are consistent with the primary research findings. Moreover, Supplementary Tables S34, 36, and S38 showcase the percentage of participants who completed a primary series of BNT162b2 vaccines across

three different studies. Section S12 summarizes the overall vaccine effectiveness and mediating effects of two-dose vaccination, which yields consistent results with the primary analysis.

Sections S4, S13, and S14 the Supplemental Appendix present the results of sensitivity analyses evaluating the impact of statistical methods. Section S13 presents findings on the overall vaccine effectiveness and mediation effects, leveraging a regression-based approach for causal mediation analysis. The results yield reasonably consistent findings as in the primary analysis. In addition, Section S14 details the proximal analysis from three studies, incorporating two negative control exposures and two negative outcomes. After adjusting for the potential of unmeasured confounders through proximal analysis, the results yield statistically insignificant direct effect estimates confirming the findings from the primary analysis that the predominant benefit of the vaccine in protecting against long COVID outcomes arises from its protective effect against SARS-CoV-2 infections.

Discussion

We identified three study cohorts to assess the population-level effectiveness of the BNT162b2 vaccine on long COVID outcomes using data from a national network of pediatric health systems in the U.S. Utilizing causal mediation analysis, we estimated the vaccine's overall effectiveness and direct and indirect impacts via specific mediating pathways. Our findings indicated a high overall protective effect of BNT162b2 against long COVID during the period dominated by the Delta

variant and moderate protective effects during the Omicron period. The estimated direct effects suggested that beyond the protective role of vaccination against SARS-CoV-2 infection, pre-infection vaccination does not significantly modify the likelihood of long COVID (i.e., conclusive or probable long COVID as well as body-system-focused PASC condition clusters). The estimated direct and indirect effects indicated that the vaccine's primary advantage in protecting against long COVID outcomes stems from its ability to reduce the risk of SARS-CoV-2 infection. The higher effectiveness of the BNT162b2 vaccine during the Delta phase, relative to the Omicron period, can be attributed both to its protection against Delta infections and to the fact that the primary benefit of the vaccine on long COVID stems from its capacity to prevent infections. With the genetic evolution of Omicron strains that generally are more highly transmissible with much lower severity of symptoms, infection rates are higher and vaccine effectiveness has been lower.[56,57]

Our study employed the causal mediation analysis to comprehensively investigate the effectiveness of the BNT162b2 vaccine on long COVID outcomes, which possesses several attractive features compared to existing studies. First, most of the existing studies focused on the vaccine's effectiveness against long COVID risks within the infected population. By employing the causal mediation analysis, the overall vaccine effectiveness reported in our study reflected the protective benefits of vaccination to the pediatric population that were both infected and uninfected, which offers substantial relevance to public health initiatives. We note that there are a few parallel studies assessing the effectiveness of vaccination against long COVID in the pediatric population.[45,58–60] Notably, a study conducted in Italy found that receiving at least one dose of the vaccine was associated with a reduced, although not statistically significant, risk of developing long COVID.[59] Second, in the existing studies, the approach by conditioning on post-treatment variables (i.e., infection) may suffer from the collider bias highlighted in Hernan et al. (2023),[37] which could produce spurious negative associations. In contrast, the direct effect in causal mediation analysis does not condition on infection but rather on a scenario where the vaccine has no impact on the risk of infection, thereby mitigating the collider bias. Third, by employing causal mediation analysis to disentangle the vaccine's impact on long COVID outcomes, the study provides a more nuanced understanding of the role of vaccination and informs decision-making. Our findings on direct effects suggest that documented infections after vaccination could pose a comparable risk of long COVID compared to documented infections occurring without prior vaccination. Particularly because the causal pathways for developing long COVID are still unknown, our findings highlight the necessity of persistently focusing on preventing SARS-CoV-2 infections using vaccination

as a focus of public health policy to mitigate the risk of long COVID.

Furthermore, our study has several additional strengths. First, to the best of our knowledge, this is the first study of the impact of COVID-19 vaccines on long COVID outcomes that offers insights into both the overall vaccine effectiveness and effects through distinct mediating pathways. Second, considering the complexity of long COVID, the effectiveness against different definitions of long COVID was evaluated: conclusive or probable diagnosis of long COVID through computable phenotype algorithms and PASC condition clusters relevant to body systems. Third, the study cohorts were sourced from a national network of academic medical centers covering a diverse population

	Vaccinated	Unvaccinated	Overall
Delta study in adolescents			
Follow-up			
Total follow-up — no. of person-wk	1,521,152	397,762	1,918,914
Median [Q1, Q3]	17 [15, 19]	17 [15, 19]	17 [15, 19]
Incidence rate per 10,000 person-wk			
Documented infection	1.97	53.10	12.56
Conclusive or probable PASC	0.11	3.54	0.82
Respiratory cluster	0.01	0.63	0.14
Musculoskeletal cluster	0.03	0.68	0.17
Cardiac cluster	0.01	0.70	0.16
Syndromic cluster	0.03	1.26	0.29
Gastrointestinal cluster	0.02	0.38	0.09
Omicron study in children			
Follow-up			
Total follow-up — no. of person-wk	4,120,783	3,265,395	7,386,178
Median [Q1, Q3]	44 [39, 46]	40 [31, 45]	42 [35, 46]
Incidence rate per 10,000 person-wk			
Documented infection	4.56	16.43	9.81
Conclusive or probable PASC	0.33	1.07	0.66
Respiratory cluster	0.15	0.41	0.27
Musculoskeletal cluster	0.04	0.15	0.09
Cardiac cluster	0.05	0.14	0.09
Syndromic cluster	0.08	0.23	0.15
Gastrointestinal cluster	0.06	0.15	0.10
Omicron study in adolescents			
Follow-up			
Total follow-up — no. of person-wk	1,491,914	1,747,295	3,239,209
Median [Q1, Q3]	43 [36, 46]	40 [31, 45]	41 [33, 45]
Incidence rate per 10,000 person-wk			
Documented infection	3.91	21.19	13.23
Conclusive or probable PASC	0.24	1.43	0.88
Respiratory cluster	0.03	0.29	0.17
Musculoskeletal cluster	0.05	0.27	0.17
Cardiac cluster	0.07	0.23	0.16
Syndromic cluster	0.08	0.43	0.27
Gastrointestinal cluster	0.01	0.16	0.09

Table 3: Patient follow-up and clinical measures in studying BNT162b2 vaccine on long COVID risks in children and adolescents. PASC: post-acute sequelae of SARS-CoV-2 infection.

	Total effect		Direct effect	Indirect effect
	Vaccine effectiveness (in %) and 95 CI	Relative risk and 95% CI	Relative risk and 95% CI	Relative risk and 95% CI
Delta study in adolescents				
Conclusive or probable PASC	95.4% (90.9, 97.7)	0.05 (0.02, 0.09)	1.08 (0.75, 1.55)	0.04 (0.03, 0.05)
Omicron study in children				
Conclusive or probable PASC	60.2% (40.3, 73.5)	0.40 (0.27, 0.60)	1.24 (0.92, 1.66)	0.31 (0.23, 0.42)
Omicron study in adolescents				
Conclusive or probable PASC	75.1% (50.4, 87.5)	0.25 (0.12, 0.50)	0.91 (0.69, 1.19)	0.21 (0.16, 0.27)

Total effect: The estimated marginal effect of BNT162b2 vaccines on the diagnosis of long COVID in children or adolescents, including both direct and indirect effects. Direct effect: The estimated effect of BNT162b2 vaccines on the diagnosis of long COVID beyond its effect on the mediator (i.e., documented infection of SARS-CoV-2). Indirect effect: The estimated effect of BNT162b2 vaccines on the diagnosis of long COVID through the mediator (i.e., documented infection of SARS-CoV-2).

Table 4: **Estimated overall effectiveness, direct and indirect effects of BNT162b2 vaccines on conclusive or probable diagnosis of long COVID in children and adolescents.**

with broader pediatric demographic, which not only provided a robust sample size but also empowered the detection of rare long COVID outcomes. Fourth, the study included a diverse representation of U.S. pediatric populations from primary care, specialty care, emergency departments, testing centers, and inpatient settings. Fifth, the comprehensive nature of the EHR data enabled us to investigate the effectiveness against diagnosis of long COVID as well as body-system-focused condition clusters while adjusting for a wide array of confounders. Finally, by conducting proximal analysis using negative control variables, our findings offer insights into the impacts of unmeasured confounding variables.

Our study also has several potential limitations. First, effectiveness was investigated in a cohort without previous infection, but potential bias resulting from undocumented infections cannot be ruled out, especially if these occurred differentially in the vaccinated and unvaccinated cohorts. Our inclusion of previous negative COVID-19 tests as a confounder can adjust for the propensity to get tested which could partially adjust for this factor. Second, in the Omicron study involving adolescents, the cohort included adolescents who had their

	Total effect		Direct effect	Indirect effect
	Vaccine effectiveness (in %) and 95 CI	Relative risk and 95% CI	Relative risk and 95% CI	Relative risk and 95% CI
Delta study in adolescents				
Respiratory cluster	97.7% (95.4, 98.9)	0.02 (0.01, 0.05)	0.78 (0.44, 1.40)	0.03 (0.01, 0.05)
Musculoskeletal cluster	92.9% (86.0, 96.4)	0.07 (0.04, 0.14)	1.63 (0.88, 3.04)	0.04 (0.02, 0.07)
Cardiac cluster	96.6% (93.3, 98.3)	0.03 (0.02, 0.07)	0.77 (0.21, 2.74)	0.04 (0.01, 0.14)
Syndromic cluster	96.2% (92.5, 98.1)	0.04 (0.02, 0.08)	1.16 (0.58, 2.33)	0.03 (0.01, 0.06)
Gastrointestinal cluster	90.2% (80.5, 95.1)	0.10 (0.05, 0.20)	2.41 (1.47, 3.94)	0.04 (0.02, 0.06)
Omicron study in children				
Respiratory cluster	44.3% (16.5, 62.9)	0.56 (0.37, 0.84)	1.87 (1.43, 2.45)	0.29 (0.22, 0.38)
Musculoskeletal cluster	71.3% (56.5, 81.1)	0.29 (0.19, 0.44)	0.49 (0.23, 1.05)	0.57 (0.27, 1.21)
Cardiac cluster	58.3% (36.6, 72.5)	0.42 (0.28, 0.63)	1.17 (0.41, 3.39)	0.34 (0.12, 0.97)
Syndromic cluster	61.9% (42.2, 74.8)	0.38 (0.25, 0.58)	1.13 (0.54, 2.37)	0.33 (0.16, 0.68)
Gastrointestinal cluster	46.7% (18.5, 65.2)	0.53 (0.35, 0.82)	1.63 (0.86, 3.09)	0.32 (0.17, 0.59)
Omicron study in adolescents				
Respiratory cluster	87.4% (74.9, 93.7)	0.13 (0.06, 0.25)	0.93 (0.61, 1.41)	0.10 (0.07, 0.16)
Musculoskeletal cluster	76.9% (53.8, 88.4)	0.23 (0.12, 0.46)	0.58 (0.35, 0.95)	0.30 (0.19, 0.49)
Cardiac cluster	27.2% (0.0, 63.5)	0.73 (0.37, 1.00)	2.72 (1.89, 3.91)	0.20 (0.14, 0.29)
Syndromic cluster	73.3% (46.7, 86.7)	0.27 (0.13, 0.53)	0.69 (0.42, 1.13)	0.29 (0.18, 0.48)
Gastrointestinal cluster[a]	–	–	–	–

[a]Unreliable estimate due to small number of events.

Table 5: **Estimated overall effectiveness, direct and indirect effects of the BNT162b2 vaccine on body-system-focused PASC condition clusters in children and adolescents.**

first vaccine after January 1, 2022. Since the use of BNT162b2 vaccines was authorized in adolescents aged 12–15 years on May 10, 2021, this cohort represents a population with late vaccines that may reduce the generalizability of the findings. Third, to ensure sufficient follow-up time for documenting infections, we required patient's enrollment at least three months before the end of the study for the Delta variant period and at least four months before the end of the Omicron variant period, which may limit the generalizability of the findings.

Fourth, vaccine records may be incomplete for individuals who received vaccine doses outside the network's care delivery sites. To address this, for each study cohort, we only incorporated sites with adequate vaccine data capture, by cross-referencing the EHR data from participating institutions with vaccination statistics from the Centers for Disease Control and Prevention (CDC), which enhances the reliability of the vaccination data in our study. We defined a reference vaccination rate by applying weights to the CDC's county-level statistics based on the residential addresses of patients. A site was considered to have adequate vaccine data capture if its reported vaccination rate was at least 60% of the CDC's reference rate.

Fifth, the identification of long COVID in children using EHR data is challenging which may introduce potential bias from inaccurate capturing of outcomes. In this study, besides the ICD10-CM U09.9 diagnosis code of long COVID, we defined probable long COVID using a computable phenotype algorithm, which includes features known to be associated with long COVID in previous statistical studies and chart review validation. Because the presentation and definition of long COVID in children is still evolving, we think the computable phenotype algorithm does offer a consistent measure that can be applied across our EHR network. In addition, the possibility of incomplete infection records may decrease the sensitivity of detecting probable long COVID cases using the computable phenotype algorithm. In a prior study, the virus testing data from EHRs were compared against institutional registry data, and the EHRs were found to be more accurate through chart reviews. Additionally, home test-positive patients typically inform hospitals when seeking further medical care. Last, given the rapid evolution of the pandemic, an updated analysis is desirable in the future to assess the findings in the context of the current circulating variants.

In summary, this study of nationally representative pediatric cohorts in the U.S., revealed protective effectiveness of the BNT162b2 vaccine on long COVID as well as body-system-focused PASC condition clusters. The findings suggested the vaccine's predominant benefit in protecting against long COVID outcomes is its capacity to reduce the risk of SARS-CoV-2 infection. This study profoundly enriches our understanding of the BNT162b2 vaccine's impact on long COVID risks within the U.S. pediatric demographic, emphasizing the critical role of vaccination and infection prevention in public health policymaking.

Data sharing statement
The data is not publicly available due to privacy concerns. The data that support the findings of this study may be available through submitting PEDSnet Study Approval applications (https://pedsnet.org/research/accessing-pedsnet/request-pedsnet-study-approval/). All requests will be evaluated by the Steering Committee and processed by the Coordinating Center.

Declaration of interests
Dr. Mosa is a board member of i2b2tranSMART Foundation.

Dr. Jhaveri is a consultant for Gilead, AstraZeneca, Sanofi, and Seqiru; receives grants from GSK, CDC, and AHRQ; holds an editorial stipend Editor-in-Chief of the Journal of the Pediatric Infectious Diseases Society (JPIDS); and receives royalties from UpToDate.

Appendix A. Supplementary data
Supplementary data related to this article can be found at https://doi.org/10.1016/j.eclinm.2024.102962.

References
1 Davis HE, McCorkell L, Vogel JM, Topol EJ. Long COVID: major findings, mechanisms and recommendations. *Nat Rev Microbiol.* 2023;21(3):133–146.

2   Davis HE, Assaf GS, McCorkell L, et al. Characterizing long COVID in an international cohort: 7 months of symptoms and their impact. *EClinicalMedicine*. 2021;38.

3   Thaweethai T, Jolley SE, Karlson EW, et al. Development of a definition of postacute sequelae of SARS-CoV-2 infection. *JAMA*. 2023;329(22):1934–1946.

4   Lopez-Leon S, Wegman-Ostrosky T, Perelman C, et al. More than 50 long-term effects of COVID-19: a systematic review and meta-analysis. *Sci Rep*. 2021;11(1):16144.

5   Groff D, Sun A, Ssentongo AE, et al. Short-term and long-term rates of postacute sequelae of SARS-CoV-2 infection: a systematic review. *JAMA Netw Open*. 2021;4(10):e2128568.

6   Nasserie T, Hittle M, Goodman SN. Assessment of the frequency and variety of persistent symptoms among patients with COVID-19: a systematic review. *JAMA Netw Open*. 2021;4(5):e2111417.

7   Rao S, Lee GM, Razzaghi H, et al. Clinical features and burden of postacute sequelae of SARS-CoV-2 infection in children and adolescents. *JAMA Pediatr*. 2022;176(10):1000–1009.

8   Frenck Jr RW, Klein NP, Kitchin N, et al. Safety, immunogenicity, and efficacy of the BNT162b2 Covid-19 vaccine in adolescents. *N Engl J Med*. 2021;385(3):239–250.

9   Walter EB, Talaat KR, Sabharwal C, et al. Evaluation of the BNT162b2 Covid-19 vaccine in children 5 to 11 years of age. *N Engl J Med*. 2022;386(1):35–46.

10 Flaxman S, Whittaker C, Semenova E, et al. Assessment of COVID-19 as the underlying cause of death among children and young people aged 0 to 19 years in the US. *JAMA Netw Open*. 2023;6(1):e2253590.

11 Chemaitelly H, AlMukdad S, Ayoub HH, et al. Covid-19 vaccine protection among children and adolescents in Qatar. *N Engl J Med*. 2022;387(20):1865–1876.

12 Cohen-Stavi CJ, Magen O, Barda N, et al. BNT162b2 vaccine effectiveness against omicron in children 5 to 11 years of age. *N Engl J Med*. 2022;387(3):227–236.

13 Florentino PTV, Alves FJO, Cerqueira-Silva T, et al. Vaccine effectiveness of CoronaVac against COVID-19 among children in Brazil during the Omicron period. *Nat Commun*. 2022;13(1):4756.

14 Olson SM, Newhams MM, Halasa NB, et al. Effectiveness of BNT162b2 vaccine against critical Covid-19 in adolescents. *N Engl J Med*. 2022;386(8):713–723.

15 Sacco C, Del Manso M, Mateo-Urdiales A, et al. Effectiveness of BNT162b2 vaccine against SARS-CoV-2 infection and severe COVID-19 in children aged 5–11 years in Italy: a retrospective analysis of January–April, 2022. *Lancet*. 2022;400(10346):97–103.

16 Khan FL, Nguyen JL, Singh TG, et al. Estimated BNT162b2 vaccine effectiveness against infection with delta and omicron variants among US children 5 to 11 years of age. *JAMA Netw Open*. 2022;5(12):e2246915.

17 Wu Q, Tong J, Zhang B, et al. Real-world effectiveness of BNT162b2 against infection and severe Diseases in children and adolescents. *Ann Intern Med*. 2024. https://doi.org/10.7326/M23-1754. Published online January 9.

18 Brannock MD, Chew RF, Preiss AJ, et al. Long COVID risk and pre-COVID vaccination in an EHR-based cohort study from the RECOVER program. *Nat Commun*. 2023;14(1):2914.

19 Ayoubkhani D, Bermingham C, Pouwels KB, et al. Trajectory of long covid symptoms after covid-19 vaccination: community based cohort study. *BMJ*. 2022;377.

20 Antonelli M, Penfold RS, Merino J, et al. Risk factors and disease profile of post-vaccination SARS-CoV-2 infection in UK users of the COVID Symptom Study app: a prospective, community-based, nested, case-control study. *Lancet Infect Dis*. 2022;22(1):43–55.

21 Perlis RH, Santillana M, Ognyanova K, et al. Prevalence and correlates of long COVID symptoms among US adults. *JAMA Netw Open*. 2022;5(10):e2238804.

22 Tran VT, Perrodeau E, Saldanha J, Pane I, Ravaud P. Efficacy of first dose of covid-19 vaccine versus no vaccination on symptoms of patients with long covid: target trial emulation based on ComPaRe e-cohort. *BMJ medicine*. 2023;2(1).

23 Azzolini E, Levi R, Sarti R, et al. Association between BNT162b2 vaccination and long COVID after infections not requiring hospitalization in health care workers. *JAMA*. 2022;328(7):676–678.

24 Ballouz T, Menges D, Kaufmann M, et al. Post COVID-19 condition after Wildtype, Delta, and Omicron SARS-CoV-2 infection and prior vaccination: Pooled analysis of two population-based cohorts. *PLoS One*. 2023;18(2):e0281429.

25 Kuodi P, Gorelik Y, Zayyad H, et al. Association between BNT162b2 vaccination and reported incidence of post-COVID-19 symptoms: cross-sectional study 2020-21, Israel. *NPJ Vaccines*. 2022;7(1):101.

26 Nehme M, Vetter P, Chappuis F, Kaiser L, Guessous I. Prevalence of post-coronavirus disease condition 12 weeks after Omicron infection compared with negative controls and association with vaccination status. *Clin Infect Dis*. 2023;76(9):1567–1575.

27 Al-Aly Z, Bowe B, Xie Y. Long COVID after breakthrough SARS-CoV-2 infection. *Nat Med*. 2022;28(7):1461–1467.

28 Taquet M, Dercon Q, Harrison PJ. Six-month sequelae of post-vaccination SARS-CoV-2 infection: a retrospective cohort study of 10,024 breakthrough infections. *Brain Behav Immun*. 2022;103:154–162.

29 Strain WD, Sherwood O, Banerjee A, Van der Togt V, Hishmeh L, Rossman J. The impact of COVID vaccination on symptoms of long COVID: an international survey of people with lived experience of long COVID. *Vaccines (Basel)*. 2022;10(5):652.

30 Scherlinger M, Pijnenburg L, Chatelus E, et al. Effect of SARS-CoV-2 vaccination on symptoms from post-acute sequelae of COVID-19: results from the nationwide VAXILONG study. *Vaccines (Basel)*. 2021;10(1):46.

31 van der Maaden T, Mutubuki EN, de Bruijn S, et al. Prevalence and severity of symptoms 3 months after infection with SARS-CoV-2 compared to test-negative and population controls in The Netherlands. *J Infect Dis*. 2023;227(9):1059–1067.

32 Kahlert CR, Strahm C, Güsewell S, et al. Post-acute sequelae after severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection by viral variant and vaccination status: a multicenter cross-sectional study. *Clin Infect Dis*. 2023;77(2):194–202.

33 Arjun MC, Singh AK, Pal D, et al. Characteristics and predictors of Long COVID among diagnosed cases of COVID-19. *PLoS One*. 2022;17(12):e0278825.

34 Budhiraja S, Indrayan A, Mahajan M. Effect of COVID-19 vaccine on long-COVID: a 2-year follow-up observational study from hospitals in north India. *medRxiv*. 2022, 2022-2027.

35 Wynberg E, Han AX, Boyd A, et al. The effect of SARS-CoV-2 vaccination on post-acute sequelae of COVID-19 (PASC): a prospective cohort study. *Vaccine*. 2022;40(32):4424–4431.

36 Català M, Mercadé-Besora N, Kolde R, et al. The effectiveness of COVID-19 vaccines to prevent long COVID symptoms: staggered cohort study of data from the UK, Spain, and Estonia. *Lancet Respir Med*. 2024;12(3):225–236. https://doi.org/10.1016/S2213-2600(23)00414-9.

37 Hernán MA, Monge S. Selection bias due to conditioning on a collider. *BMJ*. 2023. https://doi.org/10.1136/bmj.p1135. p1135.

38 RECOVER. RECOVER program. recovercovid.org. Accessed September 5, 2023.

39 Forrest CB, Margolis PA, Bailey LC, et al. PEDSnet: a national pediatric learning health system. *J Am Med Inf Assoc*. 2014;21(4):602–606.

40 VanderWeele TJ. Mediation analysis: a practitioner's guide. *Annu Rev Public Health*. 2016;37(1):17–32. https://doi.org/10.1146/annurev-publhealth-032315-021402.

41 Imai K, Keele L, Tingley D. A general approach to causal mediation analysis. *Psychol Methods*. 2010;15(4):309.

42 VanderWeele T. *Explanation in causal inference: methods for mediation and interaction*. Oxford University Press; 2015.

43 Miao W, Geng Z, Tchetgen Tchetgen EJ. Identifying causal effects with proxy variables of an unmeasured confounder. *Biometrika*. 2018;105(4):987–993. https://doi.org/10.1093/biomet/asy038.

44 Dukes O, Shpitser I, Tchetgen EJT. *Proximal mediation analysis*; 2021. Published online September 24 http://arxiv.org/abs/2109.11904.

45 Razzaghi H, Forrest CB, Hirabayashi K, et al. Vaccine effectiveness against long COVID in children. *Pediatrics*. 2024. https://doi.org/10.1542/peds.2023-064446. Published online January 16.

46 PCORnet. PCORnet common data model (CDM). https://pcornet.org/wp-content/uploads/2023/04/PCORnet-Common-Data-Model-v61-2023_04_031.pdf. Accessed September 5, 2023.

47 OHDSI. OHDSI OMOP common data model (CDM). https://ohdsi.github.io/CommonDataModel/. Accessed September 5, 2023.

48 Lorman V, Rao S, Jhaveri R, et al. Understanding pediatric long COVID using a tree-based scan statistic approach: an EHR-based cohort study from the RECOVER Program. *JAMIA Open*. 2023;6(1):ooad016.

49 Tian Y, Schuemie MJ, Suchard MA. Evaluating large-scale propensity score performance through real-world and synthetic data experiments. *Int J Epidemiol*. 2018;47(6):2005–2014.

50 Suchard MA, Schuemie MJ, Krumholz HM, et al. Comprehensive comparative effectiveness and safety of first-line antihypertensive drug classes: a systematic, multinational, large-scale analysis. *Lancet*. 2019;394(10211):1816–1826.

51 Hong G, Deutsch J, Hill HD. Ratio-of-mediator-probability weighting for causal mediation analysis in the presence of treatment-by-mediator interaction. *J Educ Behav Stat*. 2015;40(3):307–340.

52 Huber M. Identifying causal mechanisms (primarily) based on inverse probability weighting. *J Appl Econom*. 2014;29(6):920–943.

53 Sturmer T, Webster-Clark M, Lund JL, et al. Propensity score weighting and trimming strategies for reducing variance and bias of treatment effect estimates: a simulation study. *Am J Epidemiol*. 2021;190(8):1659–1670.

54 Stefanski LA, Boos DD. The calculus of M-estimation. *Am Stat*. 2002;56(1):29–38.

55 Austin PC. Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. *Stat Med*. 2009;28(25):3083–3107.

56 Araf Y, Akter F, Tang Y, et al. Omicron variant of SARS-CoV-2: genomics, transmissibility, and responses to current COVID-19 vaccines. *J Med Virol*. 2022;94(5):1825–1832.

57 Link-Gelles R, Levy ME, Natarajan K, et al. Estimation of COVID-19 mRNA vaccine effectiveness and COVID-19 illness and severity by vaccination status during Omicron BA. 4 and BA. 5 sublineage periods. *JAMA Netw Open*. 2023;6(3):e232598.

58 Yousaf AR, Mak J, Gwynn L, et al. *COVID-19 mRNA vaccination reduces the occurrence of post-COVID conditions in US children aged 5-17 Years following omicron SARS-CoV-2 infection, july 2021-September 2022*, 1935. In: *Open forum infectious Diseases*. vol. 102023;vol. 10. Oxford University Press US; 2023. ofad500-2466.

59 Morello R, Mariani F, Mastrantoni L, et al. Risk factors for post-COVID-19 condition (Long Covid) in children: a prospective cohort study. *EClinicalMedicine*. 2023;59:101961. https://doi.org/10.1016/j.eclinm.2023.101961.

60 Camporesi A, Morello R, La Rocca A, et al. Characteristics and predictors of Long Covid in children: a 3-year prospective cohort study. *EClinicalMedicine*. 2024;76:102815. https://doi.org/10.1016/j.eclinm.2024.102815.