Postacute sequelae of SARS-CoV-2 in the population: Risk factors and vaccines

Aditi Bhargava

aditi.bhargava@ucsf.edu

UCSF https://orcid.org/0000-0003-1334-0517

Sabra Inslicht

UCSF

Article

Keywords:

Posted Date: March 17th, 2025

DOI: https://doi.org/10.21203/rs.3.rs-6175467/v1

Additional Declarations: There is **NO** Competing Interest.

Abstract

The contribution of pre-existing conditions to severe versus mild postacute sequelae (PASC) of SARS-CoV-2 in the population is lacking. Here, we evaluated reproductive and other PASC side-by-side in unvaccinated and vaccinated individuals after 1st SARS-CoV-2 infection. In an online global survey of 7,541 individuals from 95 countries, high grade fever (> 102˚F)/ hospitalization after a first SARS-CoV-2 infection were more likely to be reported by vaccinated males than unvaccinated males (13.64% vs. 8.34%; p = 0.0483; HR = 1.63 [95% CI: 1.008, 2.65]). Women reported experiencing more frequent PASC than men. More women than men reported vaccine-associated adverse events (AEs) after the 1st dose (60.85% vs 48.79%, p < 0.01). Vaccine-associated hospitalization was reported by 6.24% SARS-CoV-2 naïve respondents versus 1.06% of unvaccinated after 1st SARS-CoV-2 infection. Pre-existing thyroid disease, osteoporosis, and autoimmune disorders were more prevalent in women, whereas more men reported back problems, elevated cholesterol, and hypertension; several pre-existing conditions posed ≥ 2.0 relative risk of developing severe vs mild COVID-19. Individuals aged < 55 years had an absolute risk of 6.01%, whereas individuals ≥ 55 years had an absolute risk of 11.69% of getting severe vs mild COVID-19 disease. Vaccinated women reported significantly greater menstrual cycle-associated reproductive PASC compared to unvaccinated women. Vaccinated men reported hormonal changes and sexual dysfunction as reproductive PASC compared to unvaccinated men. A detailed and thorough follow-up is needed to better understand if pre-existing health conditions and/or vaccine-associated AEs exacerbated COVID-19 sequelae.

Main

Pre-existing health conditions and comorbidities likely contribute to severe versus mild postacute sequalae (PASC) of SARS-CoV-2 infections. PASC is a collection of symptoms with sequelae across many organ systems with a subset of PASC likely manifest differently by sex, age, and pre-existing health conditions (risk factors). While age (> 65 yrs), male sex, body weight, and smoking are some known risk factors, health-related pre-existing conditions that might be a risk factor for severe vs mild PASC have not been systematically interrogated. The impact of COVID-19 vaccines on reproductive health has been the subject of significant research and discussion, especially the changes observed in reproductive health related to menstrual cycles and hormonal changes.

Vaccines for COVID-19 were developed with the help of project “warp speed”. Due to the unprecedented nature of the pandemic, the clinical trials that were conducted to test the efficacy and safety of mRNA COVID-19 vaccines deviated from the typical clinical trial process, in that within two months of initiating Phase 1–3 trials, the placebo arm was folded into the vaccine arm[1]. While millions of doses have been administered; vaccine safety and effectiveness remain equivocal. Many initial reports of published vaccine safety and high efficacy were funded by the manufacturers, including Pfizer and Novartis, and based on Phase 1–3 clinical trial data for less than six months rather than real world data[2,3]. COVID-19 vaccine effectiveness in the real world for the 2023–2024 monovalent XBB.1.5 vaccine stands at ~ 50% (95% CI = 44–55%). Protection against intensive care unit admission and death had an effectiveness at

~ 40% (95% CI = 16–58%) 4–6 months after the dose[4]. Its effectiveness in immunocompromised individuals ≥ 18 years against COVID-19–associated hospitalization was 38% in the 2 months following receipt of an updated vaccine[5]. Triple and quadruple mRNA vaccinated individuals were unable mount any significant antibody responses to newly emerging SARS-CoV-2 variants in the summer of 2024, with vaccinated individuals exhibiting antibody evasion properties[6].

In a global study entitled **p**eople's **r**esp**o**nse to **v**accine **e**ffectiveness and **s**afety (PROVES), we surveyed both vaccinated and unvaccinated people about their pre-existing health conditions, mental health, COVID symptoms, and vaccine-associated adverse events (AEs). While the survey had many components and objectives that are beyond the scope of the data described here, the subset of objectives presented in this study are as follows: first, to better understand COVID-19 vaccine hesitancy, we analyzed 5,562 freeform responses from the unvaccinated respondents using machine learning and large language models. Second, we aimed to compare PASC in vaccinated versus unvaccinated men and women systematically. Third, we evaluated the contribution of pre-existing conditions diagnosed by a medical professional on relative risk of severe versus mild PASC in the cohort. Fourth, we compared vaccine-associated AEs reported by COVID-19 naïve individuals to PASC reported after 1st SARS-CoV-2 infection in unvaccinated individuals to understand benefits versus risk of COVID-19 vaccines. Finally, we focused on reproductive PASCs in vaccinated and unvaccinated individuals after 1st SARS-CoV-2 infection.

Results

Demographics and characteristics

In the PROVES cohort, 7541 individuals from 95 different countries participated in the survey between March 2022 and April 2024; 61.70%, 12.50%, 5.72%, and 5.19% of responses were from the United States of America, Canada, the United Kingdom of Great Britain and Northern Ireland, and Australia, respectively followed by New Zealand (2.00%) and Mexico (1.35%) and 89 other countries (Fig. 1a and Supplementary Table 1). The median age of vaccinated and unvaccinated men and women were 57 and 55 years, respectively (Table 1). Nearly equal numbers of vaccinated and unvaccinated individuals reported being infected once (≥30%), twice (4.0%), whereas ≥60% reported never being tested positive for SARS-CoV-2 (Fig. 1b). The average number of SARS-CoV-2 infections did not differ by vaccination status or sex with average number of infections being ~0.4 ($p=0.655$ for sex and $p=0.35$ for vaccination status). Unvaccinated men and women reported a higher percentage of pre-existing conditions compared to their vaccinated counterparts (Table 1). Vaccinated men and women were more likely to report poor sleep quality (24.26% vs 17.98 (Men), 30.12% vs 21.48% (Women), $p=0.039$ for sex and $p=0.015$ for vaccination status) and trouble concentrating (35.69% vs 26.91% (Men), 49.44% vs 36.52% (Women)) compared to unvaccinated counterparts as PASC. Men reported higher consumption of alcohol per day in the past month (from the date of taking the survey) than women with unvaccinated men reporting the highest and vaccinated women the lowest consumption rate (8 vs 5 days, Table 1).

Vaccine hesitancy: People's response to COVID-19 mRNA vaccines

Many countries mandated COVID-19 vaccines requiring vaccination for entry into the workplace, for international travel, and playing professional sports. However, no such mandates at the federal level were imposed in the US. Despite mandates, adoption of COVID-19 vaccines was met with wide-spread refusal even by those who were previously vaccinated for other indications. In the PROVES cohort, 80.07% of the participants were unvaccinated, underscoring the need to be heard. To better understand the reasons for vaccine refusal, participants were asked "For which of the following reasons did you not get vaccinated? Select all that apply" with options "medical condition, religious beliefs, and other (write-in)". Of note, most of the COVID-19 unvaccinated participants had received more than one vaccine prior to the COVID-19 pandemic with only 2 respondents who were never vaccinated since childhood, hence the unvaccinated individuals in this Cohort cannot be labeled as "anti-vaxxers" *per se*. Among those who responded, more women than men refused the vaccine for a pre-existing medical condition (17.94% vs 9.13%, $p<0.001$, Fig. 1c) and religious reasons (25.06% vs 17.58%, $p<0.01$, Fig. 1c). The Mistral large language model (LLM) was integrated into the $ollama$ Python library to analyze freeform "other" responses for vaccine refusal from 3,756 unvaccinated individuals. Several people provided multiple reasons for vaccine refusal resulting in a total of 7254 freeform responses. After iterative training, we were able to group these responses into 7 main subcategories besides the already predefined medical and religious beliefs categories. In the "other reasons category", more men than women (92.85% vs 88.80%, $p<0.001$) refused the vaccine with the most frequent write-in reason being (paraphrased) "safety concerns, not enough data on the mRNA vaccines/ experimental and gene therapy" (3033 responses, Fig. 1d and see https://aseesa-inc.github.io/WorlMap/ for interactive maps) followed by "mistrust of Pharma and the government" (1354 responses). The third most frequent other reason for refusing was "freedom of choice/personal choice" (691 responses); fourth reason was "naturally immune" due to prior infection (317 responses) followed by "didn't want to/was not needed/I am healthy" in the 5th place (90 responses). The 6th subcategory was allergies and/or adverse reaction or injuries from other vaccines (52 responses). In the 7th subcategory, several women declined the vaccine as they were pregnant, breast feeding, or planning on becoming pregnant (25 responses, Fig 1d). These concerns were mainly voiced from the United States, Canada, the United Kingdom, and Australia with responses from several other countries as shown in Supplementary Table 1 (See https://aseesa-inc.github.io/WorlMap/ for interactive maps).

Vaccine hesitancy is often (wrongly) associated with education level. In our cohort, education levels for college degree or two-year college (11.38% vs 9.37%, $p<0.05$), bachelor's degree (31.94% vs 28.40%, $p<0.05$), and post graduate degree (12.60% vs 17.94%, $p<0.01$) differed significantly between the unvaccinated versus the vaccinated, respectively, whereas no differences were present for Master's degree or other categories (Fig. 1e). Equal percentages of the unvaccinated and vaccinated individuals were employed or retired at the time of taking the survey, whereas other categories such as disabled (2.08% vs 5.58%, $p<0.001$), unemployed (5.23% vs 3.23%, $p<0.01$), and something else (8.36% vs 5.22%, $p<0.001$) differed significantly between the unvaccinated and vaccinated, respectively (Fig. 1f).

COVID-19 disease severity in vaccinated and unvaccinated individuals

Vaccines are administered at scale to a largely healthy population to prevent serious disease, hospitalization, and death; while many vaccines such as the measles, chickenpox, and polio vaccines have been instrumental in reducing diseases severity and associated morbidity, the vaccines don't necessarily prevent transmission[7,8]. High grade fevers (>102ºF) and/or hospitalization after a first SARS-CoV-2 infection (referred to as severe COVID-19 disease in this study) were more likely to be reported by vaccinated males than unvaccinated males (13.64% vs. 8.34%; p = 0.0483; HR = 1.63 [95% CI: 1.008, 2.65]). More than 90% of women, regardless of vaccination status reported mild-to-moderate grade fever (<102ºF) (Fig. 2a). Regardless of vaccination status and biological sex, fatigue/tiredness followed by body and/or muscle ache, headache, and chills were the most frequently reported PASC after first SARS-CoV-2 infection. Women in general reported experiencing more frequent symptoms than men including change /loss of smell and taste, nausea, diarrhea, and congested nose after 1st SARS-CoV-2 infection (Fig. 2b, Extended data Fig. 1 and Extended Data Table 1).

PASC in vaccinated and unvaccinated individuals

More women reported vaccine-associated AEs than men after the 1st dose (60.85% vs 48.79%, p<0.01, Fig. 2c) and vaccinated individuals reported experiencing more frequent COVID-19 symptoms than unvaccinated respondents after 1st SARS-CoV-2 infection (Fig. 2d and Extended Data Fig. 2 and Extended Data Table 2). To evaluate how PASCs differed between the vaccinated and unvaccinated individuals after 1st SARS-CoV-2 infection, we next quantified the relative risk of experiencing specific PASC for vaccinated versus unvaccinated individuals. COVID-19-associated symptoms reported by >5% of the PROVES cohort respondents were analyzed. Cox proportional hazards models were employed to estimate the hazard ratios (aHRs) and their 95% confidence intervals (CIs) for each individual symptom (Extended Data Table 3). Vaccinated COVID-19 respondents were significantly less likely than unvaccinated COVID-19 respondents to develop the following reported post-COVID symptoms: fatigue/tiredness (aHR = 0.25), change/loss of smell and taste (aHR = 0.59 and 0.67, respectively), fever (aHR = 0.38), body and/or muscle aches (aHR = 0.35), nasal stuffiness (aHR = 0.84), chills (aHR = 0.62), and headache (aHR = 0.44), whereas vaccinated individuals were more likely to develop and reported eye soreness (aHR = 3.05), cough (aHR = 2.24), dyspnea (aHR = 3.61), trouble breathing (aHR= 1.85), sleeping more (aHR = 1.95), new confusion/inability to fully awake (aHR = 1.41), hospitalization (aHR = 1.25), intubation/ventilation (aHR = 1.15), and pneumonia (aHR = 1.4) amongst other reported symptoms (Fig. 2e and Extended Data Table 3).

Vaccine-associated adverse events versus PASC in unvaccinated

To ascertain vaccine benefits versus risk against symptoms experienced after infection with SARS-CoV-2, we compared vaccine-associated systemic AEs reported by vaccinated but COVID-19-naïve individuals after any dose to identical PASC experienced by unvaccinated but vaccine-naïve individuals after initial SARS-CoV-2 infection. Vaccine AE-associated hospitalization was reported by 6.24% (56/898)

individuals compared with 1.06% (25/2351) of unvaccinated individuals with COVID-19 (Fig. 2f, p=0.015). Difficulty sleeping, change in mood/impact on morale, heat/cold intolerance, earache/pain were all vaccine-associated AEs that were ≥1.5 log2 fold changed compared with PASCs in unvaccinated individuals (Fig. 2f and Table 2). Chills, fever, shortness of breath, cough AEs were less than those reported by unvaccinated COVID-19 individuals. However, in absence of an infection, symptoms such as cough, sore throat, nasal congestion, change/loss of smell or taste are not expected.

Sex differences in pre-existing health conditions before SARS-CoV-2 infection

Pre-existing conditions were defined as those diagnosed by a medical professional (Fig. 3a and Extended Data Table 4). Sex differences were found for pre-existing endocrine and autoimmune health conditions. Thyroid disease (12.61% F vs 2.8% M), autoimmune diseases (9.49% F vs 2.6% M), allergies (22.19% F vs 16% M), asthma (11.41% F vs 8.8% M), osteoporosis (5.228% F vs 0.428% M), irritable bowel syndrome (7.62% F vs 3.5% M,), GERD (8.62% F vs 4.5% M), and other skin disorders were reported by greater numbers of women, whereas back problems (14% M vs 11.38% F), elevated cholesterol (19% M vs 15.05% F), hypertension and blood pressure-related (24% M vs 12.99% F) pre-existing conditions were present in more men than women (Fig. 3b; error bars show 95% CI of HR).

Pre-existing conditions and relative risk (RR) of severe covid-19

Next, we calculated relative risk of severe versus mild COVID-19 disease with respect to each of these pre-existing conditions that were reported by at least 5% of the male of female participants. Presence of arthritis, gout, or chronic joint problems had the greatest relative risk (2.25) of getting severe versus mild COVID-19 disease (Fig. 3c), followed by osteoporosis (2.14), elevated cholesterol (1.656), and hormonal changes (1.46). Immunosuppressed and/or immunocompromised individuals had a RR of 1.49, whereas pre-existing blood clots in legs (5.51), lungs (5.41), deep vein thrombosis (3.25), liver diseases (3.29), nervous system disorders (3.48), renal and urinary disorders (3.02), stroke or transient ischemic attacks (2.70) and type 2 diabetes (2.15) had greater RR than individuals without these conditions, although these conditions were not present in high numbers in this cohort (Extended Data Table 5). Individuals aged 55 years or less had an absolute risk of 6.01%, whereas individuals ≥ 55 years had an absolute risk of 11.69% of getting severe vs mild COVID-19 disease.

Reproductive PASC

In our sample, 6.3% of females reported pre-existing menstrual cycle irregularities, whereas 5.93% of females and 0.891% of males reported hormonal changes as pre-existing conditions diagnosed by a medical professional. No reproductive adverse events were reported by 82.75% of unvaccinated females and 98.32% of males after COVID-19 illness vs 60.39% of vaccinated females and 93.62% of males (p=0.00016 and p=0.04, respectively) (Table 3). Hormonal changes and sexual dysfunction were reported by 2.13% vs 0% (p=0.011) and 3.4% vs 0.67% (p=0.04) of vaccinated men versus unvaccinated men after 1st SARS-CoV-2 infection (Table 3). Vaccinated women reported higher incidence of hormonal

changes, absence of periods, irregular or heavy periods, and hot flashes compared with unvaccinated women after 1st SARS-CoV-2 infection (Table 3).

Discussion

In this cross-sectional survey, the median age of the respondents was 55 years, whereas the average age was 47 years. While the average numbers of COVID-19 infections did not differ by sex and/or vaccination status, unvaccinated individuals and males reported more numbers of pre-existing conditions. Osteoporosis, arthritis, hormonal changes, high blood pressure and elevated cholesterol were some pre-existing conditions that showed ≥1.4 relative risk ratio of developing severe versus mild COVID-19 symptoms. Individuals <55 years old had an absolute risk of 6.01% and those ≥55 years of age had an absolute risk of 11.69% of developing severe vs mild COVID-19. Protection against intensive care unit admission and death had an effectiveness at ~40% (95% CI = 16%–58%) 4–6 months after additional doses of the 2023-1024 COVID-19 vaccines in adults ≥65 years of age[4].

When PASCs were compared between unvaccinated and vaccinated individuals after first SARS-CoV-2 infection, vaccinated individuals did not have any clear advantage. While vaccinated individuals were less likely to report body ache, chills, and loss of smell/taste, they were more likely to report new confusion, nasal stuffiness, sore throat, trouble sleeping and breathing. Risk reduction against moderate/severe COVID-19 in mRNA vaccinated individuals was 15.45% in <6 months vaccinated and 21.7% in >6 months vaccinated, whereas there was no benefit of mRNA or adenovirus COVID-19 vaccines on severe disease (Extended Data Table 6). The risk of PASC of SARS-CoV-2 infection remains high even in vaccinated individuals[9] and vaccination before infection only confers partial protection from SARS-CoV-2 sequelae[10]. Analysis of medical records from >400,000 veterans and >4.7 million uninfected contemporaneous individuals suggests that vaccinated individuals are less likely to report Long COVID compared with unvaccinated[9].

In our cohort, 56.30% and 45.95% of the individuals reported COVID-19 vaccine-associated AEs within 7 days of receiving the 1st or the 2nd dose, respectively. In a side-by-side comparison of vaccine AEs reported by SARS-CoV-2 naïve individuals as compared to PASC reported after 1st SARS-CoV-2 infection in unvaccinated individuals, 6.24% vaccinated (56/898) vs 1.06% unvaccinated (25/2351) individuals reported hospitalization. COVID-19 vaccines are known to be reactogenic[4]. In 2023–2024, ≥90% COVID-19 vaccine recipients (Pfizer, Moderna, or Novavax) reported non-serious adverse events, whereas ≥10% of the recipients reported serious adverse events within 7 days after vaccination to V-safe. Serious AEs resulted in individuals being unable to complete daily activities[11] and included a report of one of the following: death, life-threatening illness, hospitalization or prolongation of hospitalization, and permanent disability[11].

The human challenge study[12] that exposed healthy 36 volunteers aged 18-29 years with no prior evidence of COVID-19 infection or vaccination intranasally to a bolus of wild-type SARS-CoV-2 virus

(GBR/484861/2020) revealed that despite being exposed to large amounts of virus from an infected individual, only ~53% (18/34) of the people developed mild COVID-19[12]. These finding closely recapitulates the real-world data, where a significant subset of people despite being exposed to the virus, remained asymptomatic or had mild-to-moderate symptoms. In agreement with these findings, in the PROVES cohort, >90% of the individuals regardless of the vaccination status and with an average age of 47 years (median was 55 years), reported mild-to-moderate symptoms except for vaccinated men, where the prevalence of severe disease was 13.64%; male sex is a recognized risk factor for severe disease. Anosmia and poor sleep quality (dyssomnia) was reported by 67% of the infected participants. Runny nose, stuffiness, and headaches were the most common and frequently reported symptoms. Consistent with the findings of the human challenge studies, headache, stuffiness, loss of smell, and taste were also the most frequently reported symptoms from COVID by the PROVES study participants, regardless of vaccination status.

The human challenge study also revealed that while COVID-19 symptoms clear within few days to a week, SARS-CoV-2 virus could still be detected by RT-PCR in 85% of the individuals at the time of discharge (nearly 19 days after exposure) and in another 33% of the individuals 28 days post-exposure[12]. Viable virus could only be retrieved ~6.5 days (median) from nasopharyngeal swabs and the study found no correlation between viability and RT-PCR load[12]. Thus, use of viral load using RT-PCR in absence of any symptoms can be misleading. This study also found that naturally immune individuals had 4x higher titers of neutralizing antibodies than against spike protein. Thus, comparing titers of spike antibodies in the vaccinated and unvaccinated is not a valid comparison to predict immunity. Other studies have found that triple and quadruple COVID-19 mRNA vaccinated individuals were unable mount any significant antibody responses to recent SARS-CoV-2 variants, and exhibited antibody evasion properties[6] further suggesting limited effectiveness of mRNA COVID-19 vaccines.

Pre-existing conditions that had a ≥2.0 relative risk of developing high grade fever or being hospitalization were- renal and urinary disorders (3.02), arthritis/joint pain (2.09) osteoporosis (2.02), certain nervous system disorders (3.48), blood clots in legs (5.51), type 2 diabetes (2.15), and liver diseases (3.29). Individuals with blood clots and deep vein thrombosis were particularly at higher relative risk (>5.0), whereas immunocompromised individuals had a relative risk of 1.49 in our cohort. High blood pressure and elevated cholesterol with a relative risk of 1.37 and 1.67, respectively were present in more men and could explain why men reported more severe PASC. Hypertension/high blood pressure was also a reported risk factor in the Yale IMPACT cohort[13] and individuals with hypertension/high blood pressure were more likely to be admitted to ICU than those without these pre-existing conditions[14].

In our study, sexual dysfunction was reported by higher percentages of vaccinated men and women and women also reported significantly higher percentages of irregular period and hormonal changes after 1st SARS-CoV-2 infection. Several studies have reported COVID vaccine-associated changes to the menstrual cycle and other reproductive side effects[15-17], women in our study cohort that were never exposed to SARS-CoV-2 also reported mRNA vaccine-associated changes in menstrual cycle. Previous

studies have found that 10.6% of individuals receiving 2 vaccine doses in the same menstrual cycle experienced a change in cycle length of >8 days vs 4.3% of unvaccinated cohort[15]. Data published in 2019 from Phase 1 mRNA vaccine trial by Moderna for influenza vaccine resulted reported nearly 124 unanticipated adverse events after intramuscular doses[18]. The most common unsolicited AEs were upper respiratory tract infection, back pain, pharyngitis, and oropharyngeal pain[18], the very symptoms experienced by many after SARS-CoV-2 infection. These findings underscore the impact of COVID-19 vaccines on reproductive health, needing further research to address if and how mRNA-based therapeutics might alter reproductive function. Individuals with certain pre-existing conditions such as high blood pressure and elevated cholesterol had a higher risk ratio of developing severe COVID-19 symptoms. It is unclear if it is the pre-existing condition and/or medications associated with these indications impact COVID-19 outcomes and need to be investigated.

Limitations of this study include reporting and misclassification bias and no information on degree of clinical disease severity or laboratory assessments were available. However, reporting bias can also be present in a hospital setting or in a doctor's office as they also rely on self-reports of symptoms. Sampling bias may be present, and only English speakers were surveyed. Both vaccinated and unvaccinated participants self-identified as SARS-CoV-2 positive, however, respondents provided information on test(s) used for positive diagnosis (RT-PCR, antigen, home test), some provided no information. Even so, diagnostic methods used then were approved under Emergency Use Authorization and had limitations; other studies that analyzed data from hospitalized patients after initial positive SARS-CoV-2 test, did not find detectable viral load in several patients in spite of being tested every 3 days[13,14]. Despite these limitations, the findings regarding COVID symptoms and PASC agree with those reported by the CDC between 2020-2022[19]. There are also several strengths in our study. We reached many respondents outside of the US with ~40% of participants being non-US based. Since the survey link was anonymous and could be shared by participants, the individuals who completed the survey did so based on recommendation. This survey collected responses from individuals regardless of vaccination and SARS-CoV-2 infection status, thereby allowing a side-by-side comparison of SARS-CoV-2 symptoms in vaccinated versus unvaccinated individuals. Pre-existing conditions were defined as those as diagnosed by a medical professional before COVID. We also compared only initial symptoms, thereby excluding bias arising from prior immunity to COVID-19. As the participants took this survey remotely and anonymously, there was no pressure for them to respond in one way or the other. Standard definitions were used, and write-in options were provided, when applicable. Information on pre-existing health conditions before any COVID-19 infection, prior vaccines taken, COVID-19 restrictions, previous and existing trauma were amongst some of the information that was collected.

Our study also collected vaccine hesitancy reasons in an unbiased, freeform manner. Vaccine acceptance is shown to be associated with education level in other studies[20], but despite ≥87% of the participants with some college degree in the PROVES cohort, hesitancy for COVID-19 vaccines, especially the mRNA vaccines, was high. People from various countries perceived COVID-19 vaccines, specifically the mRNA vaccines to be experimental, or as gene therapy, lacking safety data and rigor

expressed in >3,000 free-form responses. The second most freeform reason was cited as mistrust of Pharma and/or the government and politicization of COVID-19 vaccines (1,354 responses). Suppression of freedom of expression/personal choice (691 responses) was another concern expressed by the respondents. People naturally immune due to prior SARS-CoV-2 infection or antibodies against COVID-19 (317 responses), allergies or previous vaccine injuries to other vaccines (52 responses) were additional reasons for refusing the vaccine. Pregnancy-related issues, including breastfeeding and trying to conceive (25 responses) was also a major concern expressed. Some felt it was unnecessary given they were healthy and otherwise fit (90 responses). Our study highlights the importance of having an open dialogue and not being dismissive of people's concerns about vaccines and vaccine-associated adverse events in an otherwise healthy population.

Our study found that COVID-19 vaccines had limited effectiveness in protection against severe disease in the population surveyed and decreased only a subset of PASC, whereas reproductive PASCs were present in higher numbers in vaccinated individuals. People with high blood pressure, urinary/renal, osteoporosis, and several other pre-existing conditions were at >2 relative risk of experiencing severe COVID-19 symptoms, an often-overlooked factor. As more men than women reported high blood pressure and elevated cholesterol levels, these pre-existing conditions and/or medications could explain why individuals, particularly men, and not male sex, *per se*, are at higher risk of suffering from more severe COVID-19 disease.

Online Methods

Study Design and Participants

PROVES was a cross-sectional, self-report global survey conducted online between March 2022 and April 2024. This study was approved by UCSF's Institutional Review Board (IRB# 21-35226) and follows the guidelines of the STROBE Statement for cross-sectional studies[21]. Participants were recruited by emailing the UCSF IRB approved recruitment letter with an anonymous sharable Qualtrics link within the University of California system and other Universities listservers, emailing the link the colleagues worldwide, social media and telecommunication platforms such as Facebook, WhatsApp, and LinkedIn. Using the exponential non-discriminative snowball sampling method, participants were also encouraged to share the link to the recruitment letter with the link with their network of friends. Some physicians in private practice who took the survey informed as that they shared the link with their patients. Of the 7,541 individuals who participated in the survey between March 2022 and April 2024, 6,916 individuals provided their biological sex and vaccination status to meet the subgroup analysis criteria to be included in the analysis (Supplementary Fig. 1). Inclusion criteria: Adults over 18 years of age with working knowledge of English and the ability to complete the survey without help. Indication of sex and/or vaccination status. Exclusion criteria: Inability to read and write English, absence of information on biological sex and/or vaccination status.

Country of residence

The ggmap in R package studio was used to visualize and depict PROVES study participant based on their country of residence. To avoid skewness of data due to large number of participants from any given country, the values were converted to $\log_{10}$ scale. The countries in black and blue had highest participants and red with least. No individuals from countries shown in grey participated in the survey.

Artificial intelligence-assisted subcategorization of freeform responses for vaccine refusal

The methodology for this study involved the development and application of a large language model (LLM)-based pipeline to categorize free-text survey responses related to COVID-19 vaccine hesitancy. The dataset in .csv file format was loaded into a pandas DataFrame, with a focus on the “specified” column containing participant responses. A predefined set of seven (7) subcategories were established to capture common themes in vaccine refusal “other” category, including (1) safety concerns, not enough data on the mRNA vaccines/ experimental and gene therapy, (2) distrust/mistrust in pharmaceutical companies or government, (3) natural immunity due to prior infection, (4) perceived lack of need, freedom of choice, allergies or prior vaccine injury, and pregnancy-related concerns. To classify the responses, the Mistral LLM was integrated via the “Ollama” Python library. A custom prompt was designed to instruct the model to assign each response to one of the subcategories, emphasizing contextual understanding and synonym recognition (e.g., terms like “experimental” or “gene therapy” were mapped to safety concerns). The “categorize_response” function processed each response, handling empty or missing entries by classifying them as “Other”, which were then excluded from response count. The Mistral model generated predictions, which were stored in a new “Category” column in the DataFrame. The updated dataset was saved to a new .csv file, for further analysis. Preliminary validation and additional quality control was performed by inspecting the first few rows of the categorized DataFrame and by crossing checking using find function for subcategories that had fewer responses, such as pregnancy related. This approach leveraged the capabilities of the Mistral LLM for natural language understanding, combined with pandas providing a model for data mining of freeform responses.

Statistical analysis

Response Date Subgroups

Before conducting any analysis, the dataset was deidentified of any identifiers. To further preserve anonymity, we utilized the one-way hash (SHA-1) cryptographic function to dynamically assign survey respondents to eight distinct subgroups based on the specific survey date. The survey date, expressed as a string (e.g., “YYYY-MM-DD”) was fed into the SHA-1 algorithm, which generated a unique 160-bit hash value. To determine the subgroup assignment, we can either extract the last three bits of this hash, which can represent eight unique binary combinations (000 to 111) directly corresponding to the eight subgroups, or convert the hash into a large integer, divide it by eight, and use the remainder as the subgroup identifier. This approach leverages the deterministic nature of SHA-1 while ensuring a seemingly random subgroup distribution. The same survey date will always yield the same subgroup assignments, whereas the inherent unpredictability of the SHA-1 hash minimizes potential biases in the

subgroup allocation, enhancing the overall objectivity and robustness of the survey methodology. Responses were assigned to groups based on sex and/or vaccination status. Exact values are shown for discrete variables; for continuous variables the central tendency (median/average) is shown. Variability (error bars) was calculated as the 95% confidence interval (CI) of the hazard ratio (HR). Cox proportional hazards models were employed to estimate the hazard ratios (HRs) and their 95% confidence intervals (CIs) for each individual COVID-19 symptom after adjusting for vaccination status.

Biological sex was considered as an independent variable. Whenever possible, data was analyzed separately for males and females and referred to as men and women in this study. The one sample t-test was performed between sub-groups and pooled sub-groups of vaccinated and unvaccinated groups at 95% significance level. A two-tailed t-test for independent samples was performed to compare the subgroup values of two groups; this test was found to be slightly less permissive than a Chi-Squared test. Levene's test indicated that variances were generally equal between groups ($p > 0.05$). Shapiro–Wilk tests showed that subgroups were normally distributed, and Levene's test indicated that variances were generally equal between groups for Table 1 but not in others; thus, Welch's t-test was performed for data shown in other Tables. For Table 1, p-values were calculated for sex as the average of p[F-U vs M-U] and p[F-V vs M-V], and for vaccination status the average of p[F-V vs F-U] and p[M-V vs M-U]; F = female, M = male, V = Vaccinated, U = Unvaccinated. Relative Risk (RR) was calculated as $((s_1 / c_1)/t_1) / ((s_2 / c_2)/t_2)$ for severe cases s, mild/moderate cases c and surveillance time t (person-yrs). Data was analyzed and visualized (Figures and Tables) using Aseesa's custom data analysis service and R package was used for generation of Fig. 1a, Fig. 2e, Fig. 2f, Table 2, and Extended Data Table 3.

Bar graphs. Median central tendencies are shown with error bars representing interquartile range (IQR) and normalized to percent responses. The IQR, a measure of variability, is defined as the difference between the first and third quartiles of a data set. IQR is an appropriate measure of variability for data sets with outliers or skewed distributions and for ordinal variables. It's based on the middle half of the data, so it's not usually affected by outliers. When comparing two groups, a two-tailed *t*-test was used. P values: †, ††, †††: denoting $p < 0.01$, 0.001 and 0.0001, respectively.

Vaccine Efficiency (VE) Calculation

VE was calculated by including responses that provided an answer for Question ("How many times have you ever been diagnosed with or tested positive for COVID-19?"). Vaccinated individuals who did not provide vaccine Dose 1 date were excluded. Individuals with 0 number of COVID-19 (SARS-CoV-2) infections were excluded. For every included response, *unvaccinatedDuration* was calculated as: (SurveyDate – 2020) for unvaccinated responses, (VaccineDose1Date – 2020) for vaccinated responses. *VaccinatedDuration* was calculated as

(SurveyDate – vaccineDose1Date) for vaccinated responses, zero otherwise. *unvaccinatedDuration* was added to the unvaccinated person-year pool, and MIN(*vaccinatedDuration*, 180) days was added to the

vaccinated_near person-year pool. If *vaccinatedDuration* ≥ 180 days, *vaccinatedDuration* –180 days was added to the vaccinated_far person-year pool.

To calculate the Incidence Rate Ratio (IRR), which is the ratio of the rate (number per 1000 person-years of follow-up) of confirmed cases of COVID-19 in the vaccinated group to the corresponding rate in the unvaccinated group. Since vaccine efficacy is known to decline within 6 months, we calculated IRR separately for individuals vaccinated < 6 months ago versus unvaccinated and for individuals vaccinated ≥ 6 months ago versus unvaccinated.

We defined January 1, 2020, as the start of the surveillance period and calculated person-years of follow-up in the control group as follows: For unvaccinated duration: SurveyDate – January 1, 2020. For vaccinated duration: VaccineDose1Date – January 1, 2020. Person-years of follow-up in the vaccinated groups was calculated as follows: Vaccinated < 6 months: MIN(180 days, SurveyDate – VaccineDose1Date), Vaccinated ≥ 6 months: MAX(0, SurveyDate – VaccineDose1Date – 180 days). Calculate IRR = cases/(personYears*1000)

Vaccine efficacy (< 6 mo) = $-(IRR_{vaccinated < 6\ mo}/IRR_{unvaccinated} - 1)$

Vaccine efficacy (≥ 6 mo) = $-(_{IRRvaccinated \geq 6\ mo}/IRR_{unvaccinated} - 1)$

Declarations

Data availability: Due to IRB restrictions, GDPR and privacy concerns, individual survey responses are not publicly available. Any researcher interested in further analysis can request data from the authors which will be shared as per UCSF IRB requirements. Not all data from the survey responses has been analyzed; only analyzed deidentified data can be shared as needed upon request.

Acknowledgments

This work was supported by donor funds (AB). We thank all the respondents worldwide for their time in taking the PROVES survey. The authors thank Johannes Knapp and Akshay Uttarkar for their help with Statistical analysis and Dr. Michael Acree for his critical reading of the statistical methods and the manuscript. The authors thank Connie Fee for help with the Qualtrics survey.

Author contributions

The study was conceptualized by A.B and designed by A.B. and S.S.I. Data analysis and visualization was performed by A.B., J.K., and A.U. Data verification was done by A.B. Investigation, project administration, and supervision was done by A.B, and S.S.I. Writing of the original draft was by A.B. Both A.B and S.S. I. contributed to the final review, editing, and the final approval.

Ethics declarations

AB is founder of Aseesa Inc. The study was funded by donor funds to AB. Authors declare that they have no competing interests

References

1. Pfizer. in *VACCINES AND RELATED BIOLOGICAL PRODUCTS ADVISORY COMMITTEE BRIEFING DOCUMENT* (Food and Drug Administration, chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/https://www.fda.gov/media/144246/download, 2020).
2. Heath, P. T. *et al.* Safety and Efficacy of NVX-CoV2373 Covid-19 Vaccine. N Engl J Med 385, 1172–1183 (2021). https://doi.org/10.1056/NEJMoa2107659
3. Thomas, S. J. *et al.* Safety and Efficacy of the BNT162b2 mRNA Covid-19 Vaccine through 6 Months. N Engl J Med 385, 1761–1773 (2021). https://doi.org/10.1056/NEJMoa2110345
4. Roper, L. E. *et al.* Use of Additional Doses of 2024–2025 COVID-19 Vaccine for Adults Aged ≥ 65 Years and Persons Aged ≥ 6 Months with Moderate or Severe Immunocompromise: Recommendations of the Advisory Committee on Immunization Practices — United States, 2024. MMWR Morb Mortal Wkly Rep 73, 1118–1123 (2024). https://doi.org/10.15585/mmwr.mm7349a2
5. Link-Gelles, R. *et al.* Interim Effectiveness of Updated 2023–2024 (Monovalent XBB.1.5) COVID-19 Vaccines Against COVID-19-Associated Hospitalization Among Adults Aged >/=18 Years with Immunocompromising Conditions - VISION Network, September 2023-February 2024. MMWR Morb Mortal Wkly Rep 73, 271–276 (2024). https://doi.org/10.15585/mmwr.mm7312a5
6. Wang, Q. *et al.* Alarming antibody evasion properties of rising SARS-CoV-2 BQ and XBB subvariants. *Cell* 186, 279–286 e278 (2023). https://doi.org/10.1016/j.cell.2022.12.018
7. Fine, P. E. Herd immunity: history, theory, practice. Epidemiol Rev 15, 265–302 (1993). https://doi.org/10.1093/oxfordjournals.epirev.a036121
8. Solans, L. & Locht, C. The Role of Mucosal Immunity in Pertussis. Front Immunol 9, 3068 (2018). https://doi.org/10.3389/fimmu.2018.03068
9. Xie, Y., Choi, T. & Al-Aly, Z. Postacute Sequelae of SARS-CoV-2 Infection in the Pre-Delta, Delta, and Omicron Eras. N Engl J Med 391, 515–525 (2024). https://doi.org/10.1056/NEJMoa2403211
10. Al-Aly, Z., Bowe, B. & Xie, Y. Long COVID after breakthrough SARS-CoV-2 infection. Nat Med 28, 1461–1467 (2022). https://doi.org/10.1038/s41591-022-01840-0
11. Duffy, J. (ed National Center for Emerging and Zoonotic Infectious Diseases) (Website, 2024).
12. Killingley, B. *et al.* Safety, tolerability and viral kinetics during SARS-CoV-2 human challenge in young adults. Nat Med 28, 1031–1041 (2022). https://doi.org/10.1038/s41591-022-01780-9
13. Lucas, C. *et al.* Longitudinal analyses reveal immunological misfiring in severe COVID-19. Nature 584, 463–469 (2020). https://doi.org/10.1038/s41586-020-2588-y
14. Bhargava, A. & Knapp, J. D. Immunological Misfiring and Sex Differences/Similarities in Early COVID-19 Studies: Missed Opportunities of Making a Real IMPACT. *Cells* 12 (2023).

https://doi.org/10.3390/cells12222591

15. Edelman, A. *et al.* Association between menstrual cycle length and covid-19 vaccination: global, retrospective cohort study of prospectively collected data. BMJ Med 1 (2022). https://doi.org/10.1136/bmjmed-2022-000297
16. Male, V. Menstrual changes after covid-19 vaccination. BMJ 374, n2211 (2021). https://doi.org/10.1136/bmj.n2211
17. Wali, R., Alhindi, H., Saber, A., Algethami, K. & Alhumaidah, R. The Effect of COVID-19 Vaccine on Women's Reproductive Health: A Cross-Sectional Study. Cureus 15, e40076 (2023). https://doi.org/10.7759/cureus.40076
18. Feldman, R. A. *et al.* mRNA vaccines against H10N8 and H7N9 influenza viruses of pandemic potential are immunogenic and well tolerated in healthy adults in phase 1 randomized clinical trials. Vaccine 37, 3326–3334 (2019). https://doi.org/10.1016/j.vaccine.2019.04.074
19. Kompaniyets, L. *et al.* Post-COVID-19 Symptoms and Conditions Among Children and Adolescents - United States, March 1, 2020-January 31, 2022. MMWR Morb Mortal Wkly Rep 71, 993–999 (2022). https://doi.org/10.15585/mmwr.mm7131a3
20. Lee, K. W. *et al.* COVID-19 vaccine hesitancy and its associated factors in Malaysia. PLoS One 17, e0266925 (2022). https://doi.org/10.1371/journal.pone.0266925
21. Vandenbroucke, J. P. *et al.* Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration. PLoS Med 4, e297 (2007). https://doi.org/10.1371/journal.pmed.0040297

Tables

Table 1 to 3 are available in the Supplementary Files section.

Figures

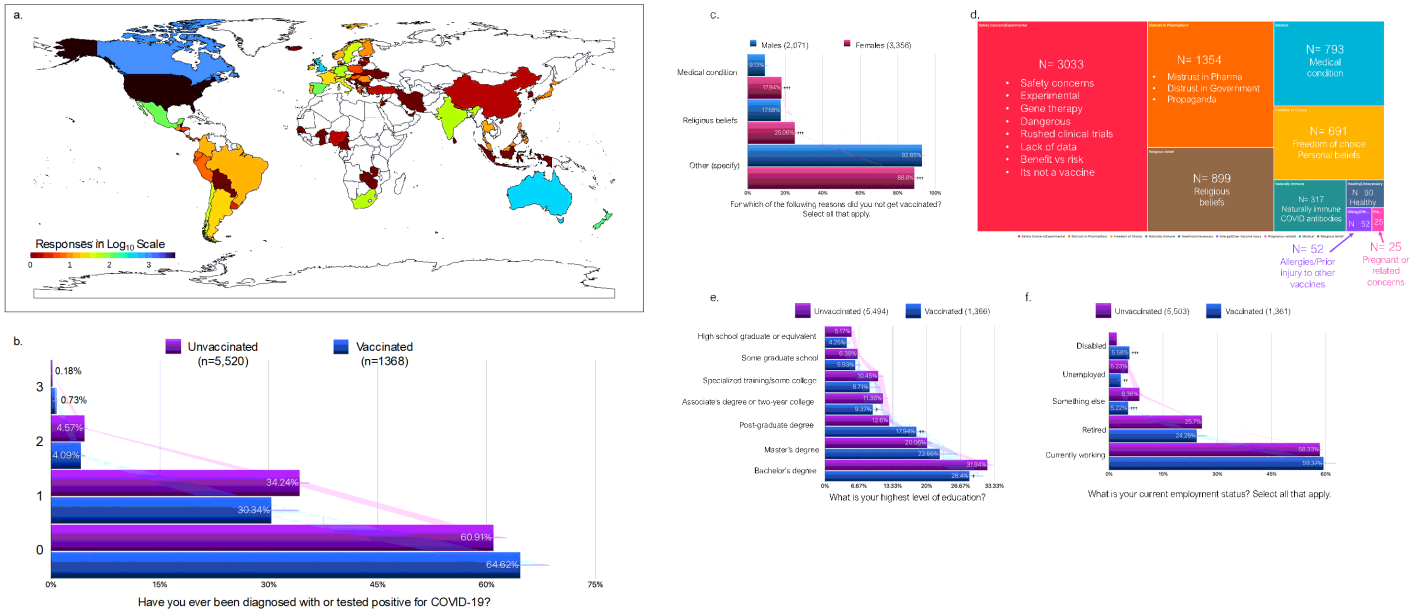

Figure 1

Participants demographics and characteristics. (**a**). World map showing PROVES participants (>7,100 individuals) from 95 countries. The ggmap in R package studio was used to visualize and depict PROVES study participant based on their country of residence at the time of taking the survey. To avoid skewness of data due to large number of participants from any given country, the values were converted to $\log_{10}$ scale. The countries in black and blue had highest participants and red with least. No individuals from countries shown in grey participated in the survey. (**b**) Equal percentages of unvaccinated and vaccinated participants reported being uninfected (0 times ≥60%), once (≥30%), and twice (≥4%). More vaccinated individuals reported ≥3 infections (0.73%) vs unvaccinated (0.18%, p=ns). (**c**) **COVID-19 vaccine hesitancy reasons.** Participants were asked reasons for refusing the COVID-19 vaccines. Participants selected multiple responses as shown (n=2,071 male and 3,356 female). More women than men refused the vaccine for medical (17.94% vs 9.13%, p<0.001) and religious beliefs (25.06% vs 17.58%, p<0.001), whereas more men than women refused for “other” reasons (92.85% vs 88.80%, p<0.001).

(**d**) Majority of individuals provided multiple reasons for vaccine refusal. Of the total 7254 responses, 5572 were extracted from the “other” freeform category using machine learning large language model (LLM). The Mistral LLM was integrated via the **“**Ollama**”** Python library, and a custom prompt was designed to instruct the model to assign each response to one of the seven predefined subcategories, emphasizing contextual understanding and synonym recognition. Treemap showing freeform responses for 7 subcategories along with number of responses for medical conditions and religious beliefs. The responses grouped in 7 major subcategories were defined based on freeform responses that included: (1) Safety concerns: mRNA vaccines are experimental/gene therapy; lack safety data and rigor (n=3,033 responses), (2) Mistrust of Pharma and/or the government and politicization of COVID-19 vaccines (n=1,354 responses), (3) Freedom of choice, personal beliefs, coercion (691 responses), (4) Natural

immunity from prior SARS-CoV-2 infection or presence of antibodies against COVID-19 (n=317 responses), (5) Am healthy and don't feel the need (n=90 responses), (6) Allergies and prior injuries with vaccines including other vaccines (n=52 responses), and (7) Pregnancy-related issues, including breastfeeding and planning to become pregnant (n=25 responses). Please see link for interactive figure for all categories and subcategories. See https://aseesa-inc.github.io/WorlMap/ for interactive maps. (**e**) Education level, and (**f**) Employment status of the vaccinated and unvaccinated participants. Bar graphs show median central tendencies with error bars representing interquartile range (IQR) and normalized to percent responses. A two-tailed *t*-test was used to determine significance. P values based on a two-tailed t-test: †, ††, †††: denoting p<0.01, p < 0.001 and 0.0001 versus previous group, respectively.

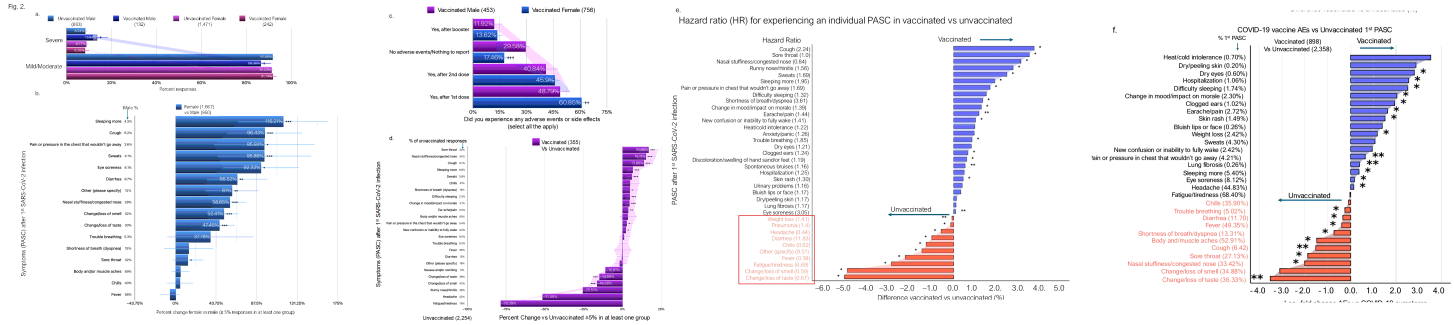

Figure 2

COVID-19 severity and PASC after 1st SARS-CoV-2 infection. (**a**) High-grade (≥ 102ºF) fever/hospitalization or moderate-to-mild fever (>99ºF-101.9ºF) in unvaccinated and vaccinated men and women; vaccinated men exhibited more severe symptoms (including hospitalization) then unvaccinated men (13.64% vs 8.34%, p<0.05). (**b**) More females than males reported COVID-19 symptoms (PASC) after 1st SARS-CoV-2 infection. Percent change in females versus males are shown for ≥5% of the responses in either group. Y-axis shows % responses in males. Females (n=1,667), and males (n=950). P values: *, **, ***: denoting p<0.01, p < 0.001 and 0.0001. For exact p values and $\log_2$ fold change, see accompanying Extended Data Table 1. Bar graphs show median central tendencies with error bars representing interquartile range (IQR) and normalized to percent responses. A two-tailed *t*-test was used to determine significance. P values based on a two-tailed t-test: †, ††, †††: denoting p<0.05, p < 0.01 and 0.001 versus previous group, respectively.

Vaccine-associated adverse events. (**c**) More women than men (60.85% vs 48.79%, p<0.05) reported experiencing COVID-19 vaccine-associated AEs after the 1st dose, whereas nearly equal numbers of men and women reported AEs after the 2nd and the booster dose. (**d**) COVID-19 symptoms (PASC) reported by individuals unvaccinated and vaccinated respondents after 1st SARS-CoV-2 infection (initial symptoms only). Only individuals who confirmed being vaccinated prior to SARS-CoV-2 infection were considered for this comparison. % of unvaccinated responses are shown on the Y-axis. P values (Welch's t-test) *, **, ***: denoting p<0.01, p < 0.001 and 0.0001 vaccinated versus unvaccinated, respectively. For exact p values and $\log_2$ fold change for each PASC, see Extended Data Table 2. n/group is shown in parenthesis. Bar graphs show median central tendencies with error bars representing interquartile range

(IQR) and normalized to percent responses. A two-tailed *t*-test was used to determine significance. P values based on a two-tailed t-test: †, ††, †††: denoting p<0.01, p < 0.001 and 0.0001 versus previous group, respectively.

(**e**) Waterfall plot showing hazard ratios (HR) for experiencing various PASC in vaccinated individuals versus unvaccinated after initial SARS-CoV-2 infection. Only individuals who confirmed being vaccinated prior to SARS-CoV-2 infection were considered for this comparison. Cox proportional hazards models were employed to estimate the hazard ratios (HRs) and their 95% confidence intervals (CIs) for each individual COVID-19 symptom after adjusting for vaccination status. Blue bars represent increased HR for PASC, and red bars (and red boxed symptoms) represent decreased HR for PASC experienced by vaccinated vs unvaccinated. Accompanying Extended Data Table 3 shows upper and lower limits for confidence interval, difference (– values= >HR in vaccinated and + values >HR in unvaccinated), HR and exact p values. *:p≤0.05; **:p≤0.01. (**f**) To evaluate the benefits versus risk of COVID-19 vaccines, vaccine-associated AEs and side effects reported after any dose of vaccine were compared side-by-side with COVID-19 symptoms experienced by unvaccinated individuals after 1st SARS-CoV-2 infection. Vaccine AE-associated hospitalization was reported by 6.24% vs 1.06% (p=0.0151) of unvaccinated individuals reported hospitalization after SARS-CoV-2 infection. Difficulty sleeping (9.91% vs 1.74%, p=0.0157), change in mood/impact on morale (9.35% vs 2.30%, p=0.012), Earache/pain (8.35% vs 2.72%, p=0.033) were some symptoms that were experienced in greater numbers as AEs by vaccinated, whereas vaccines were effective in reducing change/loss of smell (4.01% vs 34.67%, p=0.009) and taste (3.23% vs 37.35%, p=0.055). See accompanying Table 2 for details. Vaccinated n=898, unvaccinated n=2,351.

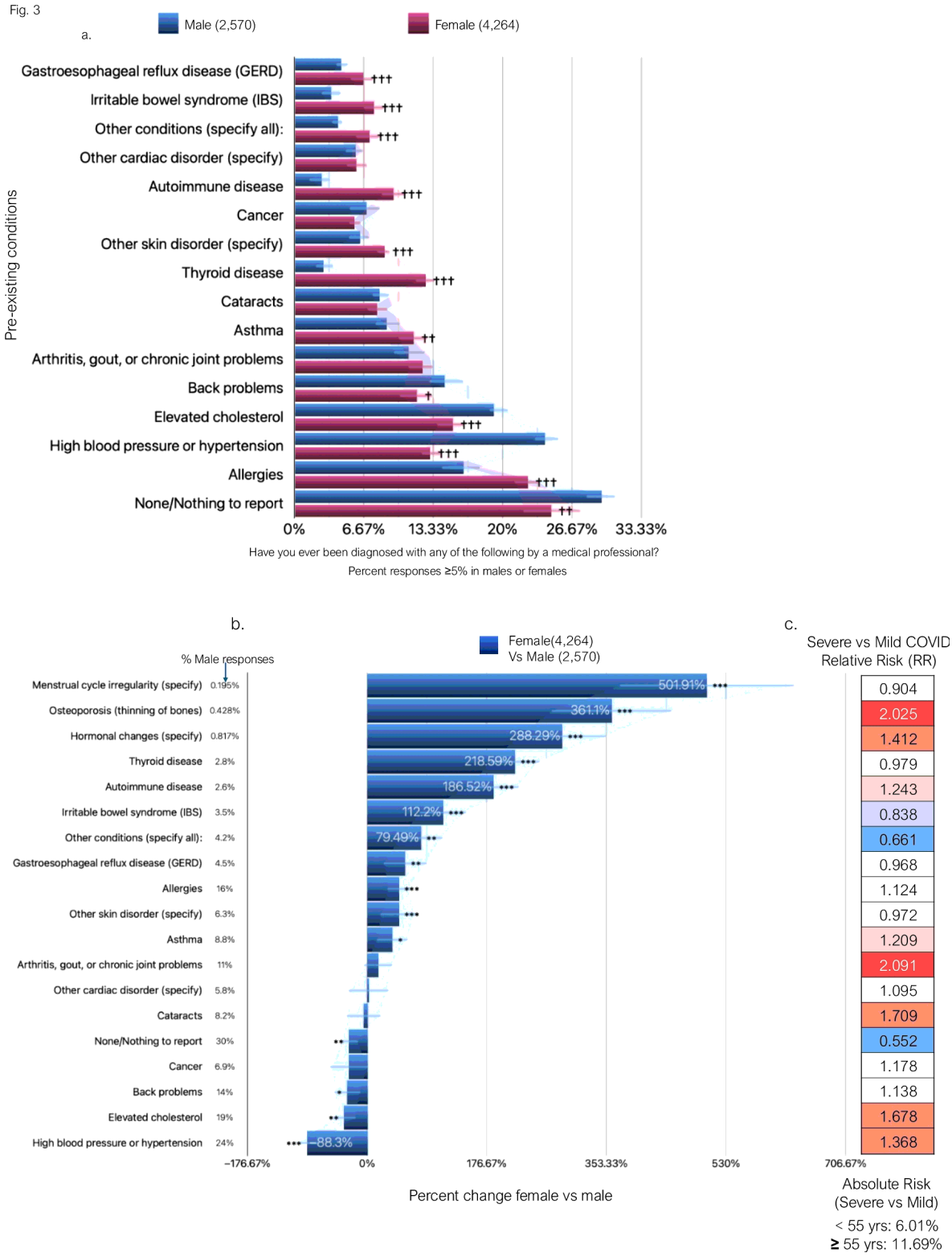

Figure 3

Pre-existing conditions and relative risk for experiencing severe vs mild COVID-19 disease. (**a**) Bar graph showing prevenance of pre-existing conditions in males and females regardless of vaccination status (>5% of responses are reported). See accompanying Extended Data Table 4 for exact p values and Log_2 fold change in pre-existing conditions in females versus males. (**b**) Percent change in pre-existing conditions in females versus males and (**c**) relative risk (RR) for each of these pre-existing conditions is

shown. Relative Risk was calculated as $((s_1 / c_1)/t_1) / ((s_2 / c_2)/t_2)$ for severe cases *s*, mild/moderate cases *c* and surveillance time *t* (person-yrs). See accompanying Extended Data Table 5 for relative risks for indications >1%. Percent labels indicate the percentage of males affected by the respective symptom. Error bars show 95% CI of the HR. Only symptoms reported by >5% of males or females are shown. P values based on a two-tailed t-test. *, **, ***: denoting $p<0.01$, $p < 0.001$ and 0.0001 females versus males, respectively.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Table13.docx
- SupplementaryFigandTable.docx
- ExtendedData16.docx