Post Hoc Ergo Propter Hoc?

Side Effect Misattribution and Vaccine Hesitancy

Svenja Miltner* Jonatan Riberth† Anton Arbman Hansing‡

This version: January 16, 2026

Abstract

Side effect concerns are the most widely cited causes of vaccine hesitancy. We study how personal exposure to suspected vaccine adverse events affects future vaccine uptake in the context of the COVID-19 pandemic. In March 2021, several European countries suspended the AstraZeneca vaccine after indications of vaccine-induced thrombosis. Although clinically confirmed cases were extremely rare, extensive media coverage may have exaggerated the link between COVID-19 vaccines and blood clots. Governments face a difficult trade-off: acting quickly is vital to prevent harm, yet premature interventions risk amplifying fears and discouraging vaccination. This paper speaks to the scope of these costs by studying how salient personal experiences lead individuals to overestimate the likelihood of adverse events and infer—often incorrectly—that they developed a vaccine-related symptom. Using new Swedish register data on reported adverse events, vaccination uptake, and healthcare use, together with matching methods, we estimate the effects of exposure to blood clots unlikely to have been caused by COVID-19 vaccination. We find, first, that such exposure increases hesitancy toward subsequent doses, plausibly due to misguided attribution. Second, effects are strongest for blood-clot-related conditions—the ones most similar to vaccine-induced thrombosis—where uptake falls by 4 percentage points, suggesting an important role for health authorities and media in shaping perceptions. When clinicians report the events as suspected side effects, subsequent uptake drops by a further 13 percentage points, indicating that cues from healthcare providers also shape perceived risk and influence future vaccine decisions.

Keywords: COVID-19, Adverse Events, Vaccine Hesitancy

*FU, Freie Universität Berlin. E-mail: s.miltner@fu-berlin.de.

†IIES, Stockholm University. E-mail: Jonatan.Riberth@iies.su.se.

‡IIES, Stockholm University. E-mail: Anton.Hansing@iies.su.se.

The authors are grateful to Arash Nekoei, Torsten Persson, Anna-Mia Ekström, Tove Fall, Olof Östergren, Natalia Danzer, Peter Mohr, participants at the 2025 SWECOV workshop and the fall PhD workshop at Freie Universität Berlin. This research was conducted within SWECOV—the Swedish Register-based Research Program on COVID-19—hosted by Stockholm University and financed by Riksbankens Jubileumsfond (RIK21-0004).

1 Introduction

In spite of its central importance to public health, vaccination continues to provoke a lasting and polarizing debate over safety and urgency. Vaccine hesitancy—the delay or refusal of available and effective vaccination—remains a critical challenge in containing communicable diseases, and is often driven by fears of harmful side effects (Jones, 2020). During the COVID-19 pandemic, scientists developed and deployed vaccines at unprecedented speed. At the same time, individuals had to navigate a high-stakes, emotionally charged decision in a complex information environment flooded with conflicting claims. The World Health Organization warned of an infodemic: an "overabundance of information, both online and offline," which plausibly made it harder to assess vaccine risks and benefits and may have fueled hesitancy.

A central challenge in pandemic management is communicating risk transparently while acting under uncertainty. Clinicians must report suspected adverse events even when causality is unproven, and regulators may take precautionary steps such as temporary restrictions or suspensions. These actions are essential for safety monitoring, but they also create salient cues that can be hard for the public to interpret—potentially raising perceived risks and reducing vaccine uptake.

We bring this trade-off to data in the context of the COVID-19 pandemic by studying how personal exposure to health events after vaccination affects subsequent vaccine uptake. In early 2021, reports linking the AstraZeneca COVID-19 vaccine to a rare type of blood clot triggered widespread media coverage and suspensions across Europe. Using rich Swedish register data on individual level vaccine decisions and diagnoses, we show that experiencing an unrelated blood clot soon after vaccination substantially increases dropout from the vaccine schedule. Effects are augmented when the blood clot occurs shortly after vaccination and when a clinician reports the event as a suspected side effect. Other acute conditions—whether plausibly vaccine-induced or not—do not show comparable effects.

When people face a new risk—such as whether to get vaccinated during the COVID-19 pandemic—they rely on both personal experiences and public information. We hypothesize that public information shapes how people interpret symptoms after vaccination. We hypothesize that highly visible safety signals—such as clinician reporting of suspected side effects or temporary pauses of drugs—shift attention toward adverse outcomes and change how people interpret symptoms that occur after vaccination. In particular, when one condition becomes the focus of public discussion, people may be more likely to attribute later, unrelated symptoms to the vaccine, especially if they resemble the condition under scrutiny.

Sweden commenced vaccination against COVID-19 in late 2020 and predominantly administered two mRNA vaccines and one adenoviral vaccine by AstraZeneca. In March 2021, reports from several countries raised the suspicion of a possible link between the AstraZeneca COVID-19 vaccine and Thrombosis with Thrombocytopenia Syndrome (TTS), an extremely rare but severe blood clotting condition. Over the course of a few days several countries suspended the AstraZeneca vaccine temporarily for safety concerns. Though the incidence of TTS was no higher than about 10 per million administered doses, media coverage could have contributed to public skepticism about vaccine safety related to blood clots more generally—conditions affecting upwards of 0.5% of the Swedish population annually. AstraZeneca had been used in Sweden for three months by the time of vaccine suspension on March 16th, making up for a total of 200,000 doses

administered. The suspension was eventually lifted and vaccination resumed on March 25 for individuals above the age of 65, until the vaccine was finally phased out by July 2021.

We are interested in how the exposure to blood clots—unrelated to the vaccine—shapes perceived risks about COVID-19 vaccines. Our primary outcome of interest is the continuation of COVID-19 vaccine schedules, i.e. the uptake of at least one subsequent vaccine dose. Our empirical strategy compares individuals with the same number of doses, vaccine brands, and month of vaccination, where some experienced a blood clot soon after vaccination while others did not. The credibility of this design stems from comparing individuals who all initiated vaccination and from a matching strategy that pairs each treated individual with a control sharing the same vaccination history, age, and similar socioeconomic and health characteristics that may influence both blood clot risk and vaccine uptake.

Our findings suggest that individuals generalize signals about specific but salient adverse events to other similar diagnoses and drugs. In particular, we find that experiencing a blood clot—unrelated to the COVID-19 vaccine—within six weeks of COVID-19 vaccination reduces uptake of the next dose by 4 percentage points. Given a baseline uptake of 94.5%, this implies a doubling of the dropout risk from the vaccination program. We conclude, that the distinct pathology of vaccine-induced blood clot does not prevent individuals from attributing unrelated symptoms to the vaccine they received. We also find moderate spillover effects on close family members. When we combine the response of the diseased individual with these spillovers, roughly half of the overall increase in vaccine hesitancy comes from the individual who experienced the blood clot, while the other half comes from close family members who subsequently discontinue vaccination. We find little support that a general misattribution mechanism could be at play, where unrelated acute events are routinely attributed to recent vaccination. In particular, for other acute conditions that could plausibly have been attributed to the vaccine, we find only very modest effects on vaccine schedule adherence. By contrast, the response to blood clots is sizable, consistent with misattribution that is activated by external cues through intense public attention.

We find that effect size is greatest for blood clots suffered soon after the vaccination. We argue that this is explained by individuals being more likely to draw the causal attribution of the link, rather than incapacitation etc. Furthermore, when a clinician reports the blood clot event as a side effect of the vaccination—a choice that is plausibly communicated to the patient—-the vaccine uptake is reduced by another 13 percentage points. The total effect on these individual is a close to five-fold increase in dropout rates, compared to matched controls. This suggests that clinician reporting, much like media coverage, cues individuals to believe that the side effect was real.

Strikingly, we find that the effect is of similar magnitude across vaccine brands, despite no suspected link for vaccines other than AstraZeneca. We infer that, even if prevailing information suggests a specific vaccine-adverse event relationship, individuals do not discriminate across vaccine brands in their risk perception of adverse events when deciding to complete their vaccine schedule.

Using proxies for the perceived benefits of vaccination, we find smaller effects among individuals with higher expected benefits from vaccinating. This pattern is consistent with vaccination decisions reflecting a trade-off between perceived benefits and perceived costs in the form of side effects. It is less consistent with alternative explanations in which uptake falls mechanically due to incapacitation, or in which the

blood-clot episode triggers a non-compensatory reaction such as generalized distrust under which high-benefit individuals would be no less likely to drop out.

Finally, we find no heterogeneity with respect to several proxies for health literacy; individuals with higher health literacy are equally likely to attribute blood clots to the vaccine they received.

Our paper speaks to both behavioral economics and public health. First, we relate to a literature on how individuals use past experiences in decision making (Ashraf et al., 2024; Bordalo et al., 2022; Malmendier et al., 2021). We contribute to this literature by examining emotionally charged, high-stakes vaccination decisions taken in the context of a pandemic, where individual choices carry implications for public health. Consistent with the previous work on experience-based decision making, we document how salient adverse health events in one domain spill over to future choices in other domains.

Second, we contribute to a literature on how individuals infer causal relationships from the temporal proximity of events, even when the evidence of an underlying link is weak or absent. While well established in cognitive psychology—often described as an "illusion of causality"—this phenomenon and its consequences have received comparatively little attention in economics. A notable exceptions is Malmendier and Tate (2005), who show how CEOs misattribute firm performance to their own actions. In a similar vein, Espín-Sánchez et al. (2023) find that churches strategically time prayers to periods when rainfall is more likely, causing the congregation to attribute rainfall to prayer rather than to natural causes. Our focus is on the extent of such misattribution in the healthcare domain. This is closely related to nocebo responses in the medical literature, where expectations about adverse effects lead individuals to experience (psychosomatic) symptoms even in the absence of an active ingredient. Our setting differs in that we study causal attribution following salient realized adverse events, rather than symptoms generated by expectations alone. That said, expectations may still shape attribution: individuals with negative expectations about a vaccine may be more likely to interpret a realized health event as vaccine-related. Consistent with the broader role of attribution and expectations in health, a substantial share of participants in the placebo arms of COVID-19 vaccine trials report adverse events following vaccination (Haas et al., 2022). Pertinently, Asan et al. (2024) show that patient reports of severe headache (a common presentation of nocebo symptoms) increased substantially after media coverage of rare vaccine-induced blood-clot cases and the associated warning symptoms in Germany.

We are not the first to study how reports linking the AstraZeneca COVID-19 vaccine to blood clots shaped vaccine attitudes during the pandemic. Agosti et al. (2022) and Deiana et al. (2022) analyze vaccine attitudes around the suspension of AstraZeneca. Using aggregate-level data across Europe, they provide evidence that the episode, and particularly the subsequent retraction, fueled skepticism extending beyond the AstraZeneca vaccine itself. While these studies document cross-vaccine spillovers, our individual-level data allows us to exploit personal exposure to blood clot events as a source of variation in salience.

The paper proceeds as follows. Section 2 provides a brief background on vaccine induced blood clots. Section 3 describes the data, and Section 5 explains the empirical strategy. In Section 6 we document our findings, while Section 7 concludes.

2 Vaccine-Induced Blood Clots

Sweden rolled out its first COVID-19 vaccines in December 2020 and, by February 2021, had deployed vaccines from all three major suppliers in Europe: BioNTech/Pfizer, Moderna, and Oxford–AstraZeneca. Soon after market access the EMA received an unexpected number of reports of severe side effects following vaccination with AstraZeneca. Typically, patients would show an unusually low count of blood platelets, internal bleeding and blood clots forming at uncommon sites of the body, such as the brain's dural venous sinuses. By March 16 three Northern European countries suspended AstraZeneca's COVID-19 vaccine Vaxzevria alongside Sweden within a week's time.[1] Pending a formal investigation of the events, the EMA announced that the overall benefits of vaccination still outweighed risks of suffering a rare blood clot event. In consequence, the Swedish health care agency resumed vaccination with AstraZeneca for people aged 65 and over on 25 March. The EMA finally confirmed a causal link between the Vaxzevria vaccine and the extremely rare and potentially fatal condition known as vaccine-induced thrombosis with thrombocytopenia (VITT). Thereafter, demand for AstraZeneca's vaccine in Sweden collapsed, and by May 2021 it was phased out as the rollout continued almost entirely with mRNA vaccines.

While the pathological mechanism of VITT is still under investigation, current studies suggests that VITT is caused by an autoimmune response to components specific to adenoviral vaccines. This is in line with the observation, that the adverse events are predominantly observed in patients vaccinated with J&J and AstraZeneca, but not with common mRNA vaccines. Additional research concluded that the AstraZeneca vaccine is not linked to a general increase in the propensity of thrombotic events overall but specifically elevates the risk of rare TTS.

Note, that the unusual clinical picture of VITT, including diagnosis criteria of platelet formation and widespread locations of clotting, made it possible to distinguish from other common types of blood clots early on. On aggregate the hematological condition was diagnosed in 1:150 000 individuals vaccinated with AstraZeneca. To benchmark this against other common types of blood clots: Pulmonary embolism, another severe clotting event, has an incidence rate of 167:100 000 among the Swedish population (Johansson et al., 2014).

Evidence on who was at risk of developing VITT is limited. UK safety monitoring suggested the event was more common among younger adults (about 1 in 50,000 under 50 vs 1 in 100,000 over 50), and early case reports seemed to involve more women, but later work stresses that these patterns may partly have reflected differential rollout and reporting rather than personal risk factors. Overall, studies have not identified individual characteristics that predicts VITT (Klok et al., 2022).

Experiencing a blood clot does not constitute a medical contraindication to COVID-19 vaccination. For example, public health authorities did not advise individuals with a history of common blood clots, such as deep-vein thrombosis following surgery or during pregnancy, to avoid vaccination. However, individuals who developed a blood clot soon after vaccination with AstraZeneca, were recommended to complete their vaccine routine with an mRNA vaccine. Furthermore, the general scientific consensus is that any vaccination

[1] The vaccine held market access authorization for the entire European Union and several other countries. Globally, in March 2021, more than as 20 countries halted the administration of the vaccine within a matter of days.

against COVID-19 is still highly beneficial even amidst the risk of general thrombotic events. In particular, studies have shown that the risk of suffering a blood clot is significantly higher following infection than following vaccination (Zhao et al., 2024; Katsoularis et al., 2022).

Our empirical strategy requires that reports of vaccine-related blood clots were salient, particularly for individuals who personally experienced such an event. In particular, we require that public reporting interacts with personal exposure: public reporting must shift behavior more for those who experience a blood clot than for otherwise similar individuals who do not. A necessary precondition is that the controversy received broad public attention. Previous studies document a spike in public interest around the suspension, using Google searches as a proxy for awareness (Deiana et al., 2022; Agosti et al., 2022), suggesting that the adverse events were part of a public debate.[2] In Sweden, we observe a similar pattern: Google searches related to blood clots spiked in March, when national health authorities announced the suspension of AstraZeneca (see fig. B1).

3 Data

3.1 Administrative Records

We combine data from several Swedish administrative sources covering the entire Swedish population. The data is accessed through the Swedish Register-based Research Program on COVID-19 (SWECOV). The permission to use the data is obtained from Sweden's Ethical Review Authority (permit numbers 2021-02225, 2022-013550-02, 2022-06118-02 and 2024-02342-02).

Reported Side Effects We draw on individual-level records of suspected adverse drug reactions (ADRs) reported to Sweden's national spontaneous-reporting system throughout the COVID-19 pandemic, obtained from the Swedish Medical Products Agency between the years of 2005 and 2022. Each entry contains the date of onset and exact notification dates of a reaction, the reporter type (a health-care professional or the individual), the suspected drug link (ATC-code and drug brand name), MedDRA reaction codes, and the EudraVigilance seriousness classification. While spontaneous systems are known to under-report, they are intended to capture any physiological effect judged unrelated to a drug's therapeutic purpose.[3]

COVID-19 vaccines The Public Health Agency provides us with data on all COVID-19 vaccinations in Sweden between January 2020 and March 2023. These records include the date of administration for each dose and the vaccine brand and manufacturer, allowing us to construct our measure of COVID-19 vaccine hesitancy.

Deaths, diagnoses & drugs We use the Swedish Death Register to identify individuals who died, either due to their blood clot conditions or because of other reasons. Furthermore, we exploit data on healthcare visits from the patient register administrated by the National Board of Health and Welfare. This data entails detailed

[2] The authors use key words such as "thrombosis" and "AstraZeneca vaccine."

[3] See García-Abeijón et al. (2023) and Hazell and Shakir (2006) for systematic literature reviews on the nature of under-reporting of adverse events.

information on all specialist care visits within the period of 2005–2022, including the date and diagnoses. We use this data to identify cases of blood clot, as well as other conditions, based on the ICD-10-SE classification.

Socioeconomic & demographic characteristics Finally, we access information from Statistics Sweden's registers on basic socioeconomic and demographic characteristics, such as age, sex, region, income.

In addition, we rely on three sources to create proxies for health literacy. (1) From the military archive, we gather data from military tests that assess cognitive ability. (2) We use data on educational attainment and occupations to pin down individuals who have a university degree or (3) who have a healthcare practitioner in their family.

3.2 Sample and Key Variables

Blood clots We are interested in the effect of experiencing a blood clot post-vaccination. To this end, we first restrict our sample to individuals who have received at least one COVID-19 vaccine. This leaves us with individuals who have a basic willingness to get vaccinated. We further restrict the sample to individuals who were alive by 2023. We use a broad definition of blood clot diagnoses, including conditions that were highly unlikely to have been caused by the vaccine. Instead, the set of diagnoses included are meant to reflect blood clot related conditions that the individual might have inferred to be due to the vaccine. This allows us to draw conclusions about the generalization of the specific severe side effect of VITT to other blood clot diseases. We consider individuals with a relevant blood clot diagnosis within 42 days of their first or second COVID-19 vaccination.[4] We exclude individuals diagnosed with a blood clot in the five years preceding the COVID-19 pandemic. This allows our treatment to more closely mirror the actual VITT cases, which were unlikely to occur disproportionately among individuals with a prior history of blood clots. This leaves us with a total of 22,957 individuals who developed a blood clot after COVID-9 vaccination.

In Table 1 we display the diagnoses that we consider in our broad definition of blood clots.[5] The displayed mortality rates refer to the share of individuals that are excluded from the sample because they pass away at some point after developing blood clot and before 2023.[6] As evident from column (4) in Table 1 the conditions differ remarkably in fatality but as a fair share of individuals have passed away within 1–2 years of developing either of the blood clots can all be classified as severe conditions. In parts of the analysis we consider spillover of diagnoses to family members and also include family members of individuals that passed away. As pointed out in Section 2, the conditions specified in our broad definition of blood clot did not serve as a contraindication for further vaccination; individuals were unlikely to be advised against taking the COVID-19 vaccine by medical professionals.

[4]Note, that the interval between the first and second dose was in general roughly six weeks. Hence we suspect that cases with vaccination pauses beyond the common six weeks are a group of selected individuals who are potentially already hesitant. While it is possible to include also individuals who developed blood clot after six weeks, it is hence more difficult to interpret the average effect in this group. Furthermore, the actual cases of vaccine induced blood clot typically happened within six weeks, making the blood clot cases under consideration more similar to vaccine induced blood clot.

[5]Throughout the remainder of the paper, we refer to the diagnoses listed in Table 1 collectively as "blood clots" for expositional convenience, although some of these diagnoses are not blood clots in a strict clinical sense.

[6]We abstain from using available information on the cause of death as the measure disregards that the primary cause of death may in fact be a comorbidity of the blood clot condition.

Table 1: Distribution of diagnoses among individuals developing a blood clot following vaccination

Diagnosis (ICD-10)	**Share (%)**	**Number of individuals**	**Mortality rate (% deceased before 2023)**
Acute myocardial infarction (I21)	30.6	7,027	22.0
Cerebral infarction or stroke (I63)	28.0	6,428	28.2
Phlebitis and thrombophlebitis (I80)	21.0	4,820	14.9
Pulmonary embolism (I26)	15.3	3,514	33.5
Other venous embolism and thrombosis (I82)	2.4	551	29.8
Arterial embolism and thrombosis (I74)	1.9	436	31.0
Portal vein thrombosis (I81)	0.8	184	49.0
Total	100	22,957	100

Notes: This table displays the prevalence of selected blood clot indications in the final sample of individuals living in Sweden in 2021-2023 (column 1 & 2). Diagnoses are included if recorded within 42 days of the first or second COVID-19 vaccination. Column 3 presents a measure of all-cause mortality for each diagnosis code referring to fatal cases within each category.

Out of the 22,957 individuals who developed blood clot after COVID-19 vaccination, 443 individuals in addition had serious adverse events for blood clot-like symptoms reported by medical professionals. The average age among individuals developing blood clot after COVID-19 vaccination is 74 years. The number of treated units is split equal between the first and the second dose (consistent with blood clot being unrelated to vaccination and showing no trend across time).

In Figure 1 we display the number of first-time blood clot diagnoses per month between 2019 and 2023. Ex post, there was no aggregate increase in blood clot incidence when COVID-19 vaccination was rolled out at scale. Blood clots are also common: each year, about 0.5% of individuals experience a first-time diagnosis. As a result, many blood clots occurring shortly after vaccination would mechanically be expected to be coincidental rather than vaccine-induced.

Ex ante, individuals could nevertheless worry that rare vaccine-related cases were underdetected or not yet reported. By March 2021, however, millions of AstraZeneca doses had already been administered worldwide and only a small number of suspected cases had been identified. Even allowing for reporting delays, contemporaneous evidence implied a very low probability that a newly developed blood clot was vaccine-induced—though this is precisely the kind of base-rate information that is difficult for laypeople to interpret in real time.

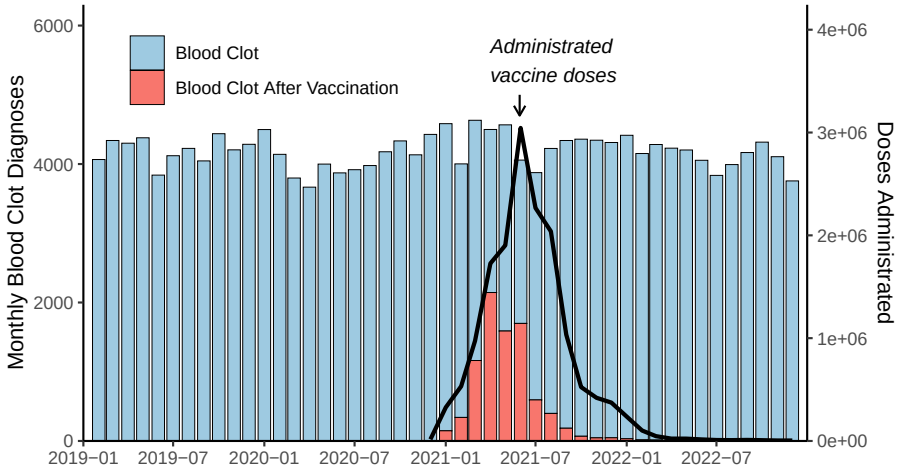


Figure 1: Timeline of Blood Clot Diagnoses

Notes: This figure displays first time blood clot diagnoses across time between 2019 and 2022. The red bars represents individuals that developed blood clot within six weeks after receiving either their first or second COVID-19 vaccine dose and hence constitute our main treatment group. The black line represents the monthly number of first and second COVID-19 vaccine doses administrated

Other conditions Apart from blood clot, we construct a group of *physical conditions*, consisting of all ICD-10 S-codes (Chapter XIX: Injury, poisoning and certain other consequences of external causes), which capture acute physical injuries to specific body regions. These diagnoses are unlikely to be attributed to the COVID-19 vaccine. We expect any effects among individuals with these conditions to, if anything, reflect an "incapacitation effect", e.g. the physical restrictions from breaking a leg make it more difficult get another dose of COVID-19. Finally, we construct a set of all *other severe acute conditions* consisting of all acute, non-physical conditions. These are similar in severity to blood clot and it is plausible that a lay person would have attributed the condition to the vaccine to a similar extent as blood clot, had blood clot not been made especially salient in the media. See Section D.1 for a more detailed description of how we select diagnoses for this category. In summary, we distinguish between three sets of diagnoses: (i) blood clots, which may be conflated with VITT; (ii) physical conditions, which are primarily incapacitating and unlikely to be attributed to the vaccine; and (iii) other severe acute conditions, which are comparable in severity to blood clots and could plausibly be attributed to the vaccine in the absence of heightened media salience.

Outcome variables We exploit the information on vaccination doses to define an outcome variable of subsequent vaccination, equal to one if an individual takes at least one more COVID-19 dose before March 2023, which marks the end of our data but also the end of the acute phase of the COVID-19 pandemic, independent of the brand of that next dose. We also study *switching*, defined as 1 if an individual continues vaccination but with a different brand than the brand that preceded the blood clot.

4 Conceptual Framework

Motivation The following section presents a simple model of how vaccination choices are impacted by experienced side effects. Individuals begin by choosing whether to get an initial vaccine dose, weighing benefits against the perceived probability and severity of side effects. Following the initial dose, they may experience a health event. Drawing on the type of illness, the time from vaccination to onset, and possible links drawn between vaccination and the symptoms by media and doctors, the individual estimates a probability that their ailment was caused by the vaccine. They then update their risk assessment of the vaccine before deciding whether to take the next dose.

The model, while simple, delivers predictions about which kinds of ailments and information treatments cause individuals to drop out of the vaccination program, as well as which individuals are most responsive. We will bring these predictions to the data.

Learning about vaccine risk Individuals are uncertain about the true probability of vaccine side effects $\theta \in [0, 1]$. Before vaccination, their belief over θ is captured by a shared prior $p(\theta)$. After receiving the first dose, individuals may or may not experience a vaccine-induced illness—$z_i \in \{0, 1\}$—where $z_i = 1$ with probability θ.

Importantly, z_i is unobserved by the individual. Instead, she observes a health signal y_i. The signal consists of (i) symptoms (or lack thereof), (ii) time from vaccination to onset, and (iii) cues from media or clinicians about the vaccine-illness link. The distribution of y_i depends on whether the illness was truly vaccine-induced ($z_i = 1$) or not ($z_i = 0$). We denote the respective conditional likelihood functions by g_1 and g_0. g_0 can be thought of as the distribution of health outcomes absent vaccination, whereas g_1 represents the individual's view of how vaccine side effects tend to present.

The individual uses the health signal y_i to form a posterior distribution over θ:

$$p(\theta|y_i) \propto p(\theta) \cdot (\theta \cdot g_1(y_i) + (1 - \theta) \cdot g_0(y_i)) \tag{1}$$

which is then used to decide whether to take the next dose.

Vaccination choice Beliefs about side effect risks, weighed against benefits, are used in vaccination choices. We denote the perceived benefits of vaccination by $B(x_i)$, where x_i is a vector of observable traits. Benefits include both health and non-health considerations.

The costs of vaccination consist of the perceived risk of side effect, $E[\theta]$, multiplied by a shared severity S. For the first dose, the expectation over θ is computed with regards to the prior distribution $p(\theta)$, whereas for the second, it is computed from the posterior $p(\theta|y_i)$. Finally, ε_i captures idiosyncratic attitudes toward vaccination not captured elsewhere. The perceived net benefit of vaccination is

$$u_i = B(x_i) - E[\theta] \cdot S + \varepsilon_i \tag{2}$$

We assume that individuals choose not to vaccinate if $u_i \leq 0$. If $u_i > 0$, individual i gets vaccinated at time

$$t_i = T_0(x_i) + \Delta_i \tag{3}$$

where $\Delta_i|u_i \sim G(\cdot|u_i)$ such that Δ_i is stochastically decreasing in u_i,[7] and $T_0(x)$ denotes the vaccine availability date for an individual with observables x.

Model predictions The model delivers six predictions that we bring to the data. Predictions 1–4 relate to which kinds of health signals trigger vaccination dropout, whereas 5–6 concern who responds more.

1. *Attribution.* Health events following vaccination reduce subsequent doses only if they can plausibly be attributed to the vaccine.
2. *Timing.* A shorter time from vaccination to symptom onset increases the effect.
3. *Clinician cues.* Effects are larger when a clinician indicates a possible connection between the diagnosed illness and the vaccine.
4. *Media salience.* Illnesses linked to vaccination in media will have a greater effect than similarly plausible diagnoses without such coverage.
5. *Benefit gradient.* Individuals with lower predicted benefit from the vaccine respond more to health events.
6. *Hesitancy amplification.* Individuals who delay their first vaccination reduce their uptake of the next dose more in response to a health event.

The last two predictions share a common logic. Among those who took the first dose, individuals with lower $B(x_i)$ and/or higher Δ_i have lower u_i—closer to the threshold $u_i = 0$—and are therefore more likely to be tipped into dropping out upon suffering a negative health event.

5 Method

Challenges to Identification & Matching Our empirical strategy is based on comparing immunization outcomes among individuals who developed blood clot shortly after receiving the COVID-19 vaccine to individuals who received the same vaccine but did not develop a blood clot. This approach relies on the assumption that conditional on observable characteristics, developing blood clot is orthogonal to vaccine hesitancy. We assume that this conditional independence assumption holds also among individuals who were diagnosed and additionally have their doctor report the blood clot event as a severe side effect. The main threats to our identification strategy are that individuals may be differentially likely to (1) develop a blood clot in the first place, (2) get diagnosed with a blood clot and (3), conditional on being diagnosed with a blood clot, have their doctor report it as a side effect. In particular, individuals may be selected in terms of predispositions affecting the probability of developing blood clot in the first place as well the inclination to consult health care professionals. Furthermore, we cannot rule out that individuals who have a latent hesitancy to vaccinate press medical professionals to report their condition as a side effect.

To handle these endogeneity concerns, we deploy a two-step matching procedure. The key idea is to identify individuals who were similar along observable characteristics by the time an individuals is diagnosed with blood clot. To that end, we first identify control individuals in terms of the COVID-19 *vaccine history* including timing and the brand of previous vaccine doses; An individual that develops a blood clot after the

[7] Formally, for any u, u': $u' > u \implies G(\cdot|u)$ stochastically dominates $G(\cdot|u')$.

second dose is matched to an individual who received the same vaccine brand within the same month for the first and second dose. By matching on vaccine brand together with vaccination timing, we address what is perhaps the most important concern: individuals who are more worried about side effects selecting into vaccines perceived as safer and receive their doses later. We, furthermore, restrict the matching to individuals who by the time of developing blood clot, had not yet received their next dose. Within these vaccine history-cells, we deploy extreme gradient boosting to compute propensity scores along basic socioeconomic characteristics. Each treated unit is matched to their nearest untreated propensity score neighbor who had not yet taken their next dose.

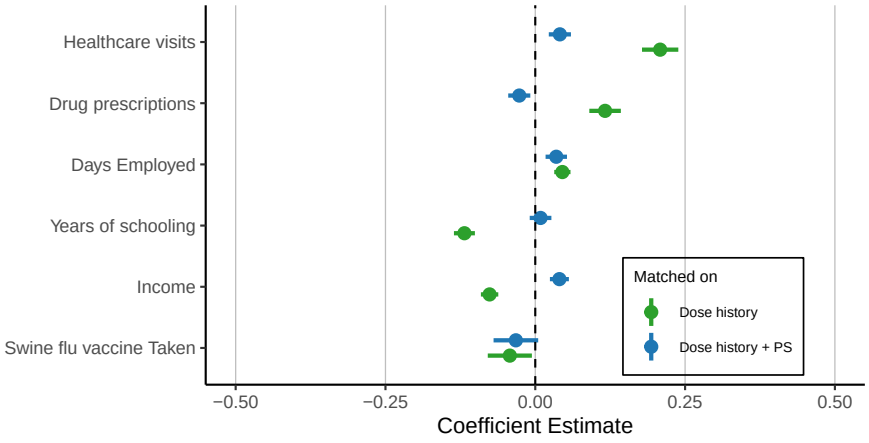


Figure 2: Balance in Pre–treatment Characteristics

Notes: This figure displays results from univariate regressions where the treatment (diagnosed with blood clot) is regressed on standardized covariates in the "Dose history" sample where we match on (i) vaccine brand(s) received, (ii) calender month. Furthermore, each treated individual is matched to a control individual who had not yet taken their next dose by the time the treated individual is diagnosed with the blood clot. "Dose history + PS-matching" is instead based on a sample where to each treated individual we match a control individual based on propensity scores.

As evident from Figure 2, matching solely on vaccine dose history leaves residual imbalance in terms of individual health characteristics (e.g.). Notably, while there is little difference in terms of socioeconomic characteristics, individuals in the treatment group appear to be sicker as measured by the number of drug prescriptions and healthcare visits. To approximate conditional independence, we match individuals on observable socioeconomic characteristics that jointly predict both vaccine hesitancy and blood clot onset. The propensity score matching does a good job at achieving balance, although some residual balance remains.

6 Results

The role of media salience In Figure 3 we display our main results for the different blood clot conditions along with estimates for physical conditions and other severe acute conditions.

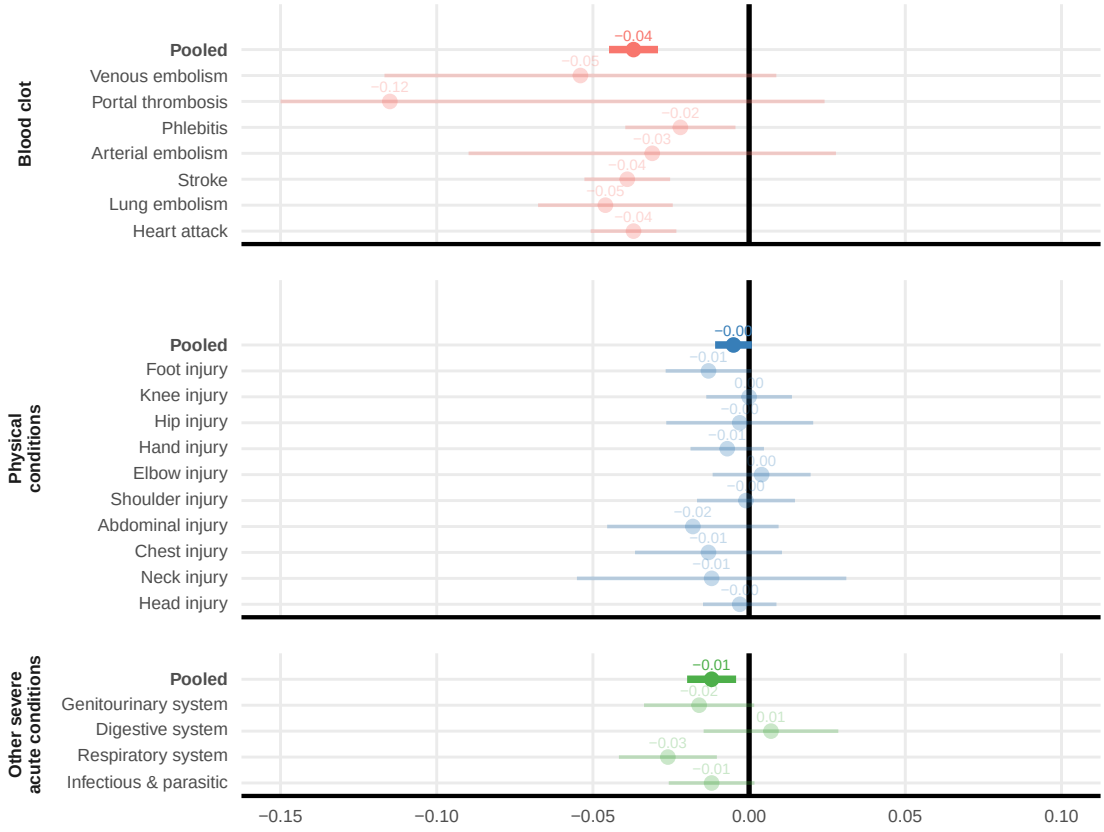


Figure 3: Vaccine Hesitancy and Media Exposure

Notes: This figure displays coefficient estimates of regressions run on matched sample where the dependent variable is equal to 1 if an individual takes additional COVID-19 vaccine doses and 0 otherwise.We consider three types of conditions: (i) Blood clot conditions that were unlikely to have been vaccine-induced but plausibly confounded to have been induced by the COVID-19 vaccine; (ii) Physical conditions that were highly unlikely to have been adverse events and where the individual is unlikely to infer that they they were adverse events; (iii) Other Severe Acute conditions conditions that were beforehand somewhat likely to have been adverse events from the COVID-19 vaccine but turned out to also be unlikely to have been vaccine-induced and where there was little to no public discussion revolving them being potential adverse event.

We find an average reduction in uptake of 4 percentage points among individuals who experienced a blood clot soon after vaccination.[8] There is little apparent heterogeneity across specific blood clot diagnoses, although our data does not allow for sharp conclusions on this point.

We find negative effects also for the ex-ante plausible conditions, albeit smaller in magnitude. This difference is unlikely to reflect differences in severity: as shown in Table Table D1, these conditions are also severe and associated with high mortality rates. Taken together, the results suggest that the media salience of blood clots amplified hesitancy among individuals who experienced such events shortly after vaccination.

[8] In an alternative specification, we explicitly construct a control group where control units are drawn from a donor pool of individuals who developed blood clot *before* the COVID-19 pandemic. This approach directly creates comparable units in terms of propensity to develop blood clot. However, this group of individuals is smaller rendering it less feasible to target other imbalances in other characteristics such as socioeconomic characteristics. Either way, the approach yields identical results for blood clot diagnoses (results upon request).

At the same time, the negative coefficient for the ex-ante plausible conditions indicates that developing a condition shortly after vaccination —even one not causally related to the vaccine— can also reduce vaccine uptake.

We do not detect an effect for individuals experiencing a pure placebo diagnosis, such as a foot injury. This is in line with the notion that individuals can not plausibly connect their diagnosis to the vaccine and therefor do not adjust their beliefs towards the vaccine.

Spillover to family members In Table B1, we report the estimates for partners, siblings, and children of the diseased. The outcome variable is defined analogously to our previous approach, equaling 1 if a network member takes another vaccine dose after the focal individual within their family is diagnosed with a blood clot, and 0 otherwise. We report our estimates for the case that the focal individual died or lived on after experiencing a blood clot. Note that death is coded irrespective of cause. Overall, the estimated effects are small and negative, although the estimates for children are close to 0. Notably, the effect for partners of individuals with a non-fatal diagnosis is roughly half the size of the effect on the diseased themselves. The estimates for siblings are smaller in size potentially capturing that partners are more exposed to the struggles accompanying the diagnosis than other network members. This pattern suggests that the responses we document are not solely driven by individuals updating beliefs about their own predisposition to side effects, but more broadly by the inference that side effects are relatively common when a family member develops a blood clot shortly after vaccination. Note, that we find insignificant and small estimates for all but siblings in cases where the focal individual dies. Overall, our results bear relevance in the more general quantification of social costs of side effects.

Accounting for the spillovers to family members, the coefficient point estimates in Table B1 suggest that every blood clot is associated with 0.08 individuals dropping out of the vaccination schedule[9]. In other words, about half of the effect on stopping vaccination comes from the affected individual, and the other half comes from family members stopping vaccination.

6.1 What information is used?

Heterogeneity by time until blood clot Our explanation for why individuals become more hesitant after experiencing blood clot is that they form an assessment about the risk of developing side effects from the next dose. This in turn depends on an assessment of the likelihood that the blood clot that they developed was induced by their previous vaccine dose. The perceived link should be stronger for individuals who developed blood clot or any plausibly vaccine-induced condition closer in time to vaccination. We illustrate the results to testing this hypothesis in Figure 4, where we split the sample by the number of weeks elapsed between the COVID-19 vaccination and the blood clot incident. In **??** we display the corresponding results for the decision to switch vaccine brand rather than abstaining from vaccination completely.

[9]This number comes from multiplying the estimated effect sizes in Table B1 with the size of each network among the treated units (2.33 children, 2.1 siblings and 1 partner)

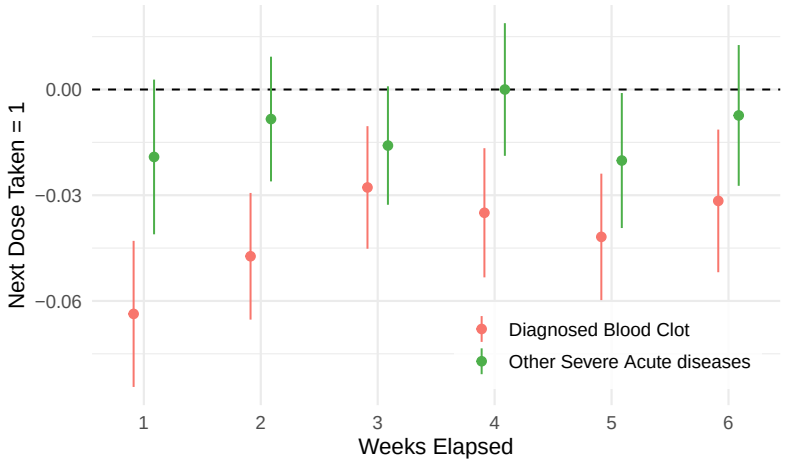
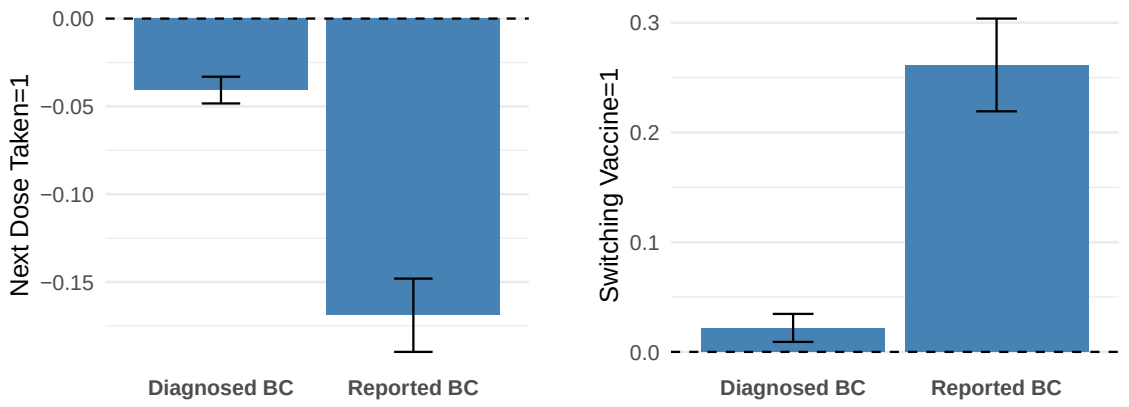


Figure 4: Heterogeneity by Time Elapsed Between Vaccine and blood Clot

Notes: This figure displays coefficient estimates of regressions run on matched sample where the dependent variable is equal to 1 if an individual takes additional COVID-19 vaccine doses and 0 otherwise.The sample is split by the number of weeks elapsed between the last COVID-19 vaccine dose and developing a condition.

We find that the estimates for blood clots are decreasing with latency period, suggesting individuals developing a blood clot condition long after vaccination are less hesitant to proceed with their vaccination schedule. If anything, there is less of a negative gradient for ex-ante plausible conditions, suggesting that for conditions that appeared even less plausibly linked to vaccination at the time (although all were very unlikely ex post), individuals who inferred a vaccine connection did so regardless of timing. Taken together, individuals appear somewhat rational in the sense that they are less likely to attribute a blood clot to the vaccine when it occurs long after vaccination. At the same time, the overall pattern supports the interpretation that the effects are driven by misattribution.

The role of doctors' communication A key question is whether misattribution arises merely from experiencing a negatie health event, or whether it requires additional confirmation from external sources. To shed further light on this, we consider effects among a subset of individuals who, on top of developing blood clot after COVID-19 vaccination, had it reported as a suspected side effect by a healthcare professional. Under the premise that doctors communicate their decision to report to the patient, we view doctors' reporting as confirming to the patient that the condition was a side effect. We find significantly larger effects in this subset with an average effect of 16 percentage points compared to the effect of just developing blood clot (4 percentage points). Our interpretation is that doctors report it as a side effect and in doing so cues the patient perception that the condition was vaccine-induced, increasing the perceived probability that the individual will develop side effects also from subsequent doses.



(a) Continuing Vaccination

(b) Switching Vaccines

Figure 5: Role of doctors' communication

Notes: This figure reports coefficient estimates from regressions estimated on the matched sample. Diagnosed BC indicates individuals diagnosed with a blood clot within six weeks of receiving a COVID-19 vaccine (or their matched controls), while Reported BC further restricts the sample to cases reported as suspected adverse events by a healthcare professional. **Panel (a)** uses an indicator for continued COVID-19 vaccination as the dependent variable. **Panel (b)** uses an indicator for switching vaccine brand as the dependent variable.

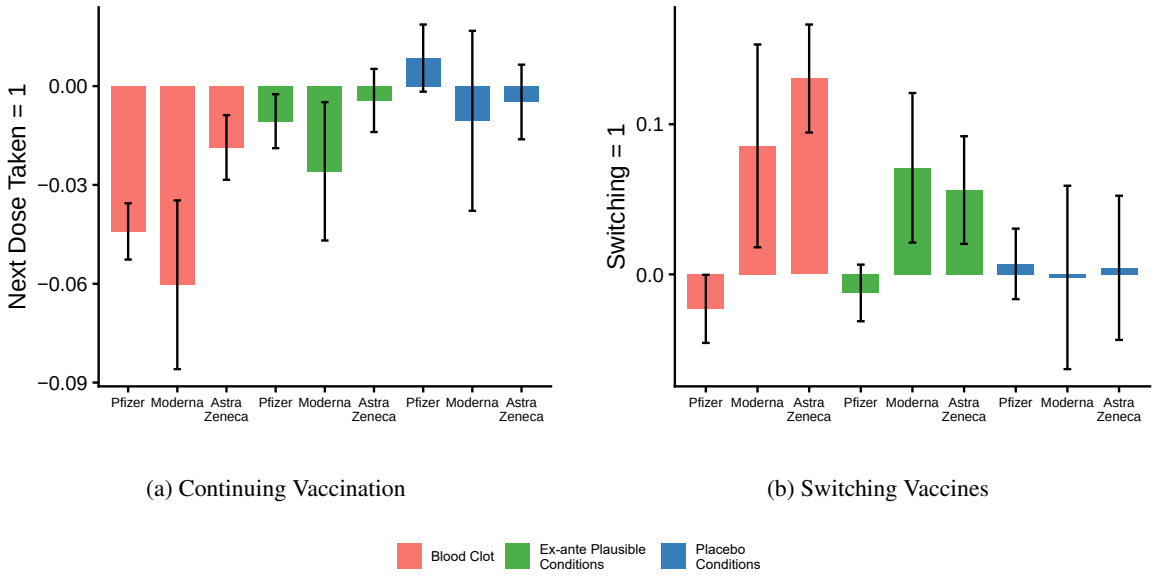
These results speak to an important policy trade-off. While doctors need to report suspected side effects to detect new ones, the act of reporting side effects also cause individuals to become more hesitant, despite the fact that the experienced symptom was unlikely to have been a side effect.

Heterogeneity across vaccine brands We exploit heterogeneity across vaccine brands to study whether individuals' responses to the VITT episode were narrowly targeted to AstraZeneca—the vaccine associated with VITT—or whether the shock spilled over to other brands as well. Spillovers would be consistent with imperfect attribution and "generalization" in risk beliefs, implying that the social costs of salient adverse events may extend beyond closely related conditions and also affect vaccines that were not implicated. To this end, we revisit our analysis across the three main vaccine brands administered in Sweden. We present the results of this exercise in Figure 6. To account for compositional differences in who receive the different brands, we reweight, observations by age, risk group status and whether or not they have a doctor in their family following DiNardo et al. (1996). We display these results in Figure B3.

For the decision to take the next dose, there is little heterogeneity across brands, suggesting that individuals who drop out of the vaccine schedule do not strongly distinguish which vaccine might have caused the adverse event.

However, many individuals respond by continuing vaccination but switching brand. Experiencing a blood clot after vaccination with AstraZeneca or Moderna significantly increases the probability of switching, with the largest effects for AstraZeneca. We do not detect comparable switching after a blood clot following Pfizer, consistent with Pfizer being perceived as the safest: while abstention effects look similar, the incentive to switch away from Pfizer is smaller.

Finally, switching effects are larger for AstraZeneca in the blood-clot analysis than in the ex-ante plausible analysis, whereas the difference is smaller for Moderna. This pattern is suggestive that media exposure



(a) Continuing Vaccination

(b) Switching Vaccines

Figure 6: Brand Heterogeneity

Notes: This figure displays coefficient estimates of regressions run on matched sample where the dependent variable is equal to 1 if an individual takes additional COVID-19 vaccine doses and 0 otherwise. Coefficient estimates are split up by the brand that the individual received prior to developing blood clot.

played a role in amplifying hesitancy and switching around AstraZeneca.

6.2 Who is affected?

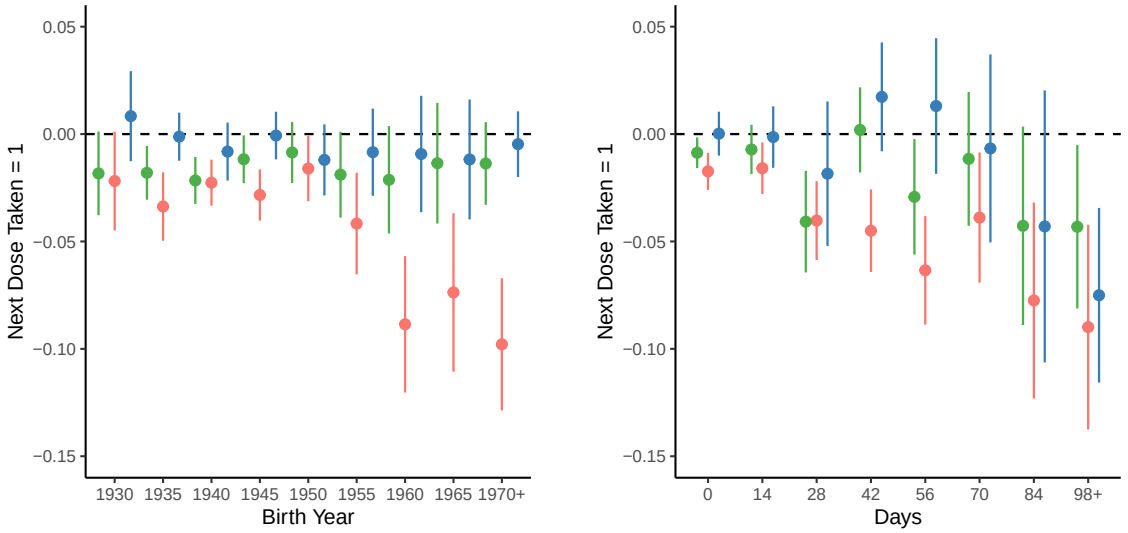
Heterogeneity in benefits of vaccination Abstaining from vaccination entails an excess risk of falling ill from COVID-19. Do individuals at higher risk of severe COVID-19 respond less to exposure to side effects? In Figure 7, we use two proxies for perceived benefits of vaccination: age and the time elapsed between when the vaccine became available to an individual and the timing of the first dose. We find that older individuals, and those who took the dose preceding the blood clot earlier, react less than younger individuals or those who took their first dose later.

For age, older individuals face a higher risk of severe COVID-19 and therefore have greater benefits from vaccination than younger individuals, who are on average more indifferent between vaccinating and not vaccinating in the absence of blood clot. A similar pattern holds for time since availability. Individuals who take the vaccine shortly after it becomes available are likely those with higher perceived benefits and are therefore less likely to stop immunization, compared to individuals who delay vaccination, for whom a shock to perceived costs is more likely to tilt the decision toward abstaining from future vaccination.

Combining the results from Figure 7 panel (a) and panel (b), there is much less of a gradient for other Severe Acute Diseases compared to blood clot, suggesting that these individuals do not trade off benefits and costs in the same way as those experiencing blood clot. As previously stated, the most plausible explanation is that the small negative effects that we find largely reflects an incapacitation effect that is independent of age or time until first dose.

Taken together, these results suggest that individuals trade off the benefits and costs of vaccination. While

the costs, in the form of perceived side-effect risk, are similar across individuals experiencing blood clot, perceived benefits vary, generating heterogeneity in the decision to stop vaccinating.



(a) Heterogeneity by age

(b) Time elapsed since first date available

Figure 7: Eliciting the perceived benefits of vaccination

Notes: This figure displays coefficient estimates of regressions run on matched sample where the dependent variable is equal to 1 if an individual takes additional COVID-19 vaccine doses and 0 otherwise. **Panel (a)**: Coefficient estimates split up by birth year. **Panel (b)** Coefficient estimates split up by weeks elapsed since first date the COVID-19 vaccine is available. Date of first availability is defined in Section D.2.

Heterogeneity in health literacy Finally, we are interested in heterogeneity by personal health literacy. This allows us to pin down the role of understanding cues, i.e. media communication on a rare side effect such as VITT, in the observed effects on vaccine adherence. For instance, having a doctor in the family may make it easier to understand the lack of clinical association between general blood clot conditions and the rare cases of VITT. Figure 8 reveals our main estimates by different measure of health literacy. We find no evidence of heterogeneity with respect to health literacy.

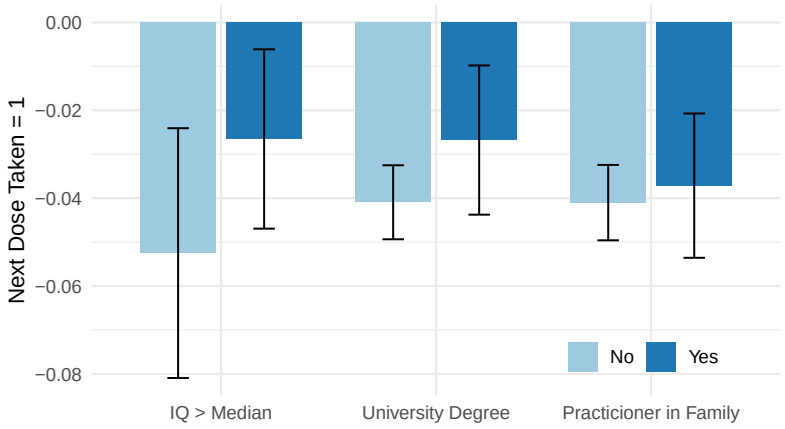


Figure 8: Heterogeneity in Health Literacy

Notes: This figure displays coefficient estimates of regressions run on matched sample where the dependent variable is equal to 1 if an individual takes additional COVID-19 vaccine doses and 0 otherwise.We split up three proxies for health literacy. *Doctor in family* is defined as "Yes" if an individual has a parent or a sibling with a medical degree or a nursing degree. *High Cognitive Ability* is based on cognitive tests completed by military conscripts. It is equal to "Yes" if an individual had above median score on the test compared to peers born the same year. *University Degree* is equal to "Yes" if an individual has at least a bachelor degree, corresponding to three years of higher education.

7 Conclusions

This paper shows how experiencing a health condition shortly after vaccination—combined with informational cues such as media salience around side effects or a clinician's decision to file a side-effect report—can lead individuals to falsely attribute the condition to the vaccine. Theoretically, our findings highlight how people generalize across health domains: effects spill over not only to blood-clot diagnoses that differ from the specific adverse event of concern, but also across vaccine brands, even when only one brand was associated with elevated risk. Taken together, our findings point to a surprisingly large degree of misattribution, with individuals generalizing from a rare and vaccine brand-specific adverse event to other blood-clot diagnoses and to vaccines for which no elevated risk was suspected.

From a policy perspective, the results suggest that media salience is costly. These costs arise as soon as news about a potential side effect breaks: the withdrawal of the AstraZeneca vaccine shortly thereafter did not prevent a persistent increase in hesitancy among those who experienced health events in the same period. The evidence on clinician side-effect reporting also points to an important trade-off. Reporting is essential for detecting genuine adverse reactions, yet the act of reporting may unintentionally reinforce patients' beliefs that their condition was vaccine-induced, leaving a lasting impact on their willingness to continue vaccination.

References

Agosti, F., V. Toffolutti, N. Cavalli, S. Nivakoski, M. Mascherini, and A. Aassve (2022). Information and vaccine hesitancy: Evidence from the early stage of the vaccine roll-out in 28 european countries. *Plos one 17*(9), e0273555.

Asan, L., J. Kleine-Borgmann, B. Bozkurt, B. Frank, M. Köhrmann, C. Kleinschnitz, and U. Bingel (2024). Media coverage of covid-19 vaccination-associated cerebral venous sinus thrombosis was followed by a surge in emergency presentations due to headache–observations from a university hospital in germany. *Frontiers in Psychiatry 15*, 1378472.

Ashraf, N., G. Bryan, A. Delfino, E. Holmes, L. Iacovone, C. Meyer, and A. Pope (2024). Learning to see a world of opportunities. Technical report, Working Paper. Accessed November 22, 2024. https://navaashraf.com/wp . . . .

Bordalo, P., J. J. Conlon, N. Gennaioli, S. Y. Kwon, and A. Shleifer (2022, 08). Memory and probability*. *The Quarterly Journal of Economics 138*(1), 265–311.

Deiana, C., A. Geraci, G. Mazzarella, and F. Sabatini (2022). Perceived risk and vaccine hesitancy: Quasi-experimental evidence from italy. *Health Economics 31*(6), 1266–1275.

DiNardo, J., N. M. Fortin, and T. Lemieux (1996). Labor market institutions and the distribution of wages, 1973-1992: A semiparametric approach. *Econometrica 64*(5), 1001–1044.

Espín-Sánchez, J.-A., S. Gil-Guirado, and N. Ryan (2023, June). Praying for rain. NBER Working Paper 31411, National Bureau of Economic Research.

García-Abeijón, P., C. Costa, M. Taracido, M. T. Herdeiro, C. Torre, and A. Figueiras (2023, July). Factors associated with underreporting of adverse drug reactions by health care professionals: A systematic review update. *Drug Safety 46*(7), 625–636. Epub 2023 Jun 6.

Haas, J. W., F. L. Bender, S. Ballou, J. M. Kelley, M. Wilhelm, F. G. Miller, W. Rief, and T. J. Kaptchuk (2022). Frequency of adverse events in the placebo arms of covid-19 vaccine trials: a systematic review and meta-analysis. *JAMA network open 5*(1), e2143955–e2143955.

Hazell, L. and S. A. Shakir (2006). Under-reporting of adverse drug reactions: a systematic review. *Drug Safety 29*(5), 385–396.

Johansson, M., L. Johansson, and M. Lind (2014). Incidence of venous thromboembolism in northern sweden (veins): a population-based study. *Thrombosis journal 12*(1), 6.

Jones, S. P. (2020, April). Imperial college london yougov covid data hub. https://www.coviddatahub.com. Online resource; accessed via https://www.coviddatahub.com.

Katsoularis, I., O. Fonseca-Rodríguez, P. Farrington, H. Jerndal, E. H. Lundevaller, M. Sund, K. Lindmark, and A.-M. Fors Connolly (2022). Risks of deep vein thrombosis, pulmonary embolism, and bleeding after covid-19: nationwide self-controlled cases series and matched cohort study. *BMJ 377*.

Klok, F. A., M. Pai, M. V. Huisman, and M. Makris (2022). Vaccine-induced immune thrombotic thrombocytopenia. *The Lancet Haematology 9*(1), e73–e80.

Malmendier, U., S. Nagel, and Z. Yan (2021). The making of hawks and doves. *Journal of Monetary Economics 117*, 19–42.

Malmendier, U. and G. Tate (2005, November). Ceo overconfidence and corporate investment. *The Journal of Finance 60*(6), 2661–2700.

Zhao, L., H. Tran, M. Risk, and G. Nair (2024). Risk of blood clots after covid-19 vaccination and infection: a risk-benefit analysis. *Research square*, rs–3.

Appendix A. Medical Background

Pathology of vaccine induced blood clots Vaccine-induced immune thrombotic thrombocytopenia (VITT), also known as thrombosis with thrombocytopenia syndrome (TTS), is an extremely rare adverse event that has been observed following viral vector vaccines, in particular AstraZeneca and Johnson & Johnson. No cases of VITT were documented for the mRNA vaccines Pfizer and Moderna. The blood clots commonly form in unusual sites of the body (mainly cerebral or sinus vein thromboses) and are associated with a low count of blood platelets, unlike for other blood clotting diagnoses. VITT is an immune reaction to PF4 (platelet factor 4), a protein stored in blood cells, the so called platelets, and commonly released in case of injury. PF4 binds to negatively charged molecules, such as heparin, and thereby moderates the wound repair process in case of injury by fostering the blood to clump. In case of VITT, the body produces antibodies against PF4, which attach to the protein and thereby form an immune complex. These immune complexes activate platelets causing them to hyperactivate PF4 release and clump together. For this reason, VITT is unusually associated with the symptom of *low* platelet count, which is in stark contrast to other blood clot conditions. Note, that the pathology is similar to Heparin-induced thrombocytopenia (HIT). This condition of low platelets that is also associated with blood clots, caused by a blood thinning medication.

Treatment guidelines

Appendix B. Additional Descriptive Statistics & Additional Results

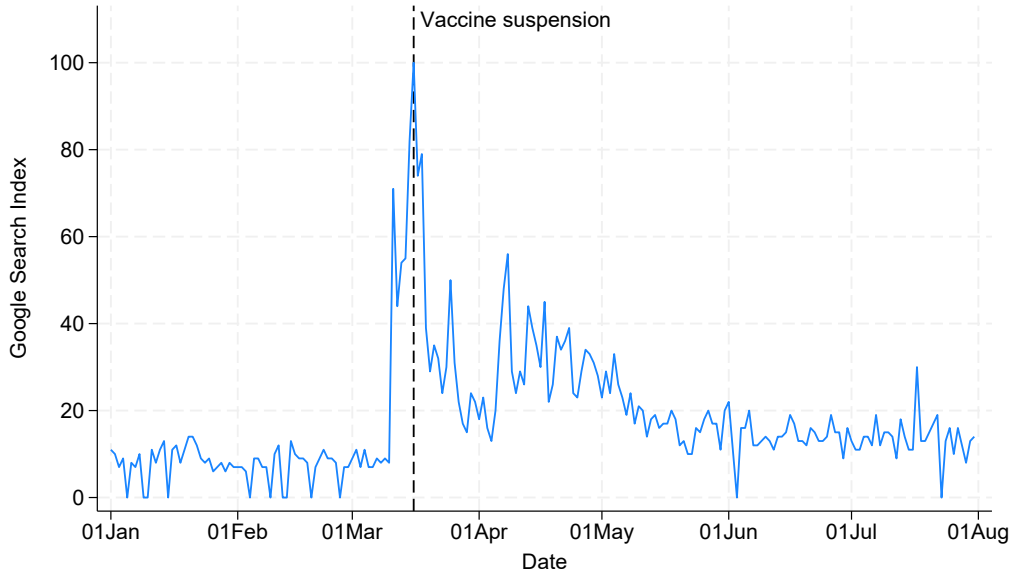


Figure B1: Timeline of Google Searches

Notes: This figure displays the google search index for Sweden using the keyword ”blodpropp” between January 1st and August 31st 2021. The frequency of searches are indexed to the time of highest search volume (= 100). The vertical line represents the suspension of AstraZeneca vaccine in Sweden on March 16 2021. Last accessed

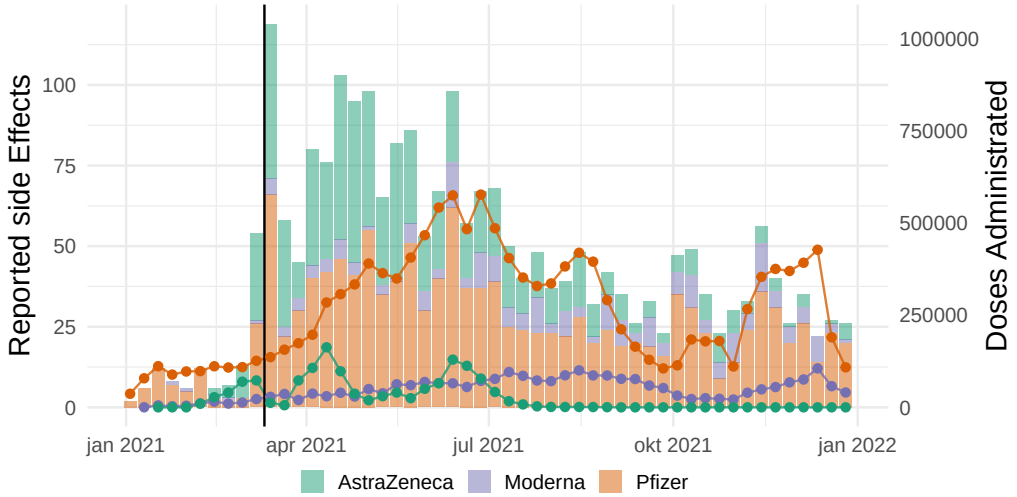


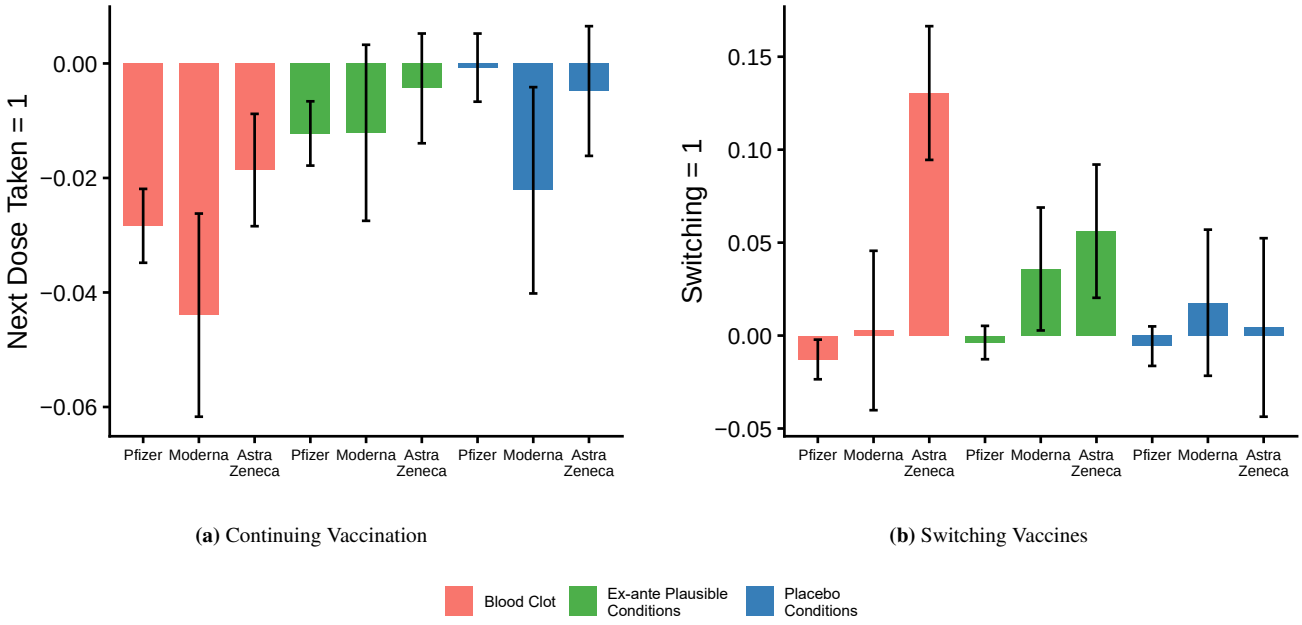
Figure B2: The Blood Clot Episode: Blood Clot Reports Over Time

Notes: This figure displays reported reported side effects for blood clot symptoms across time (bars). In addition, number of COVID-19 doses administrated are displayed (lines). The black line represents the first report of a potential link between AstraZeneca and an excess risk of developing blood clot.

Table B1: Effect of Blood Clot on Network Members' Vaccination Decisions

	Children	Partners	Siblings
Focal individual died = 0			
Treated	0.003	-0.019	-0.013
	(0.003)	(0.003)	(0.004)
n obs	84106	24590	59477
Focal individual died = 1			
Treated	0.005	-0.009	-0.018
	(0.004)	(0.006)	(0.008)
n obs	29190	6431	16155

Notes: This table reports estimates of the effect of exposure to a family member developing blood clot shortly after vaccination on the vaccination uptake of network members. Each column corresponds to a separate regression. The dependent variable is equal to 1 if an individual takes at least one COVID-19 vaccine dose after the family member is diagnosed with blood clot. Standard errors are clustered at the match-pair level.



(a) Continuing Vaccination

(b) Switching Vaccines

Figure B3: Brand Heterogeneity, DFL reweighted

Notes: This figure displays coefficient estimates of regressions run on matched sample where the dependent variable is equal to 1 if an individual takes additional COVID-19 vaccine doses and 0 otherwise. coefficient estimates are split up by the brand that the individual received prior to developing blood clot.

Appendix C. Description of Matching

Propensity scores are computed using an XGBoost gradient boosting model with a binary logistic objective. Categorical variables are one-hot encoded, including missing-value indicators. The model is trained using 5-fold stratified cross-validation. To prevent overfitting, within each fold of cross-validation an early stopping rule avoids overfitting by stopping the training procedure after 10 rounds of no performance improvement on the test data set. We set key hyperparameters as follows: a learning rate (η) of 0.1, a maximum tree depth of 4, a minimum child weight of 1, a subsample fraction of 0.8, and a column subsampling fraction of 0.8. These settings provide moderate regularization and control overfitting while capturing nonlinear relationships.

The large number of potential matches in combination with the tight blocking conditions makes the task to find nearest neighbors without replacement challenging. Instead, we implement matching with replacement. Given the large number of candidate control individuals, this is unlikely to be a concern.

Appendix D. Data & Variable Definitions

D.1 Defining ex-ante plausible conditions

To identify ex-ante plausible conditions, we begin by selecting diagnoses that are sufficiently common to generate meaningful statistical power—those with at least 1,000 cases per year on average between 2015 and 2022 (596 ICD codes). We then restrict to diagnoses for which more than half of cases originate in specialized care (93 codes), ensuring that the conditions reflect medically significant events rather than routine primary-care contacts. Next, we exclude ICD chapters that are clearly unsuitable for our context: circulatory diseases targeted by specific treatments (chapter I), external causes of injury (chapters S–Y), non-diagnoses such as symptoms or health-status factors (chapters R, U, Z), pregnancy-related and congenital conditions (chapters O, P, Q), cancers (chapters C, D), and mental-health conditions (chapter F), leaving 31 codes. Finally, we remove a small set of individual diagnoses that are either dominated by COVID-specific patterns (e.g., certain respiratory infections), reflect chronic or lifestyle conditions (e.g., diabetes complications, COPD, nutritional deficiencies, alcohol-related liver disease), represent post-surgical complications, or are too closely related to blood clots to serve as a meaningful comparison. This results in a focused set of 19 ICD codes that plausibly could be mistaken for vaccine-related by a non-expert but are not mechanistically linked to COVID-19 vaccination.

Table D1: Distribution of diagnoses among individuals developing a blood clot following vaccination

Diagnosis (ICD-10)	**Share (%)**	**Number of individuals**	**Mortality rate (% deceased before 2023)**
Infectious agents (B95–B98)	27.6	9,359	24.2
Other bacterial diseases (A30–A49)	22.6	7,650	32.6
Appendicitis (K35–K38)	16.4	5,552	1.2
Acute kidney failure (N17–N19)	8.5	2,881	41.1
Bacterial pneumonia (J15)	7.1	2,415	36.0
Pleural effusion (J90)	6.0	2,040	47.0
Unspecified kidney failure (N19)	2.6	882	40.6
Peritonitis (K65–K67)	2.4	825	18.5
Respiratory failure (J96)	1.6	546	58.1
Pulmonary oedema (J81)	1.5	497	46.6
Aspiration pneumonitis (J69)	0.9	317	60.8
Pneumococcal pneumonia (J13)	0.7	248	18.4
Respiratory failure, unspecified (J96.9)	0.7	235	22.6
Hepatic failure (K72.9)	0.4	138	57.0
Pyothorax / empyema (J86)	0.3	101	22.6
Others	0.6	201	45.3
Total	100.0	33,887	26.3

D.2 Defining date of availability

We adopt a data-driven approach to determine when the vaccine first becomes available to an individual, defining availability dates separately for each birth-year–by–healthcare-region combination. For each birth-year × region cell, let $f(i)$ denote the vaccination date of the i-th individual to be vaccinated (excluding healthcare workers). We define the date of availability as $f(i^*)$, where

$$i^* = \arg \min_{i \leq N-50} \{f(i+50) - f(i)\}.$$

Intuitively, the availability date is the vaccination date at which the time span between individual i and the individual vaccinated fifty places later is minimized.