



OPEN ACCESS

EDITED BY

Ana Corte-Real,
University of Coimbra, Portugal

REVIEWED BY

Chiara Pennisi,
Kore University of Enna, Italy
João Mendes-Abreu,
University of Coimbra, Portugal

*CORRESPONDENCE

Elizabeth B. Klerman
✉ ebklerman@hms.harvard.edu

RECEIVED 28 May 2025

ACCEPTED 11 August 2025

PUBLISHED 16 September 2025

CITATION

Wang W, Abbaspour S, Blumenthal KG,
Hashimoto DM, Robbins GK and
Klerman EB (2025) Utilizing group-based
models to identify adverse event patterns
after an intervention.
Front. Med. 12:1637091.
doi: 10.3389/fmed.2025.1637091

COPYRIGHT

© 2025 Wang, Abbaspour, Blumenthal,
Hashimoto, Robbins and Klerman. This is an
open-access article distributed under the
terms of the [Creative Commons Attribution
License \(CC BY\)](#). The use, distribution or
reproduction in other forums is permitted,
provided the original author(s) and the
copyright owner(s) are credited and that the
original publication in this journal is cited, in
accordance with accepted academic
practice. No use, distribution or reproduction
is permitted which does not comply with
these terms.

Utilizing group-based models to identify adverse event patterns after an intervention

Wei Wang¹, Sara Abbaspour², Kimberly G. Blumenthal³,
Dean M. Hashimoto⁴, Gregory K. Robbins⁵ and
Elizabeth B. Klerman^{1,2*}

¹Division of Sleep and Circadian Disorders, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, United States, ²Division of Sleep, Department of Neurology, Massachusetts General Hospital, Boston, MA, United States, ³Division of Rheumatology, Allergy and Immunology, Department of Medicine, Massachusetts General Hospital, Harvard Medical School, Boston, MA, United States, ⁴Workplace Health, Mass General Brigham, Boston, MA, United States, ⁵Division of Infectious Diseases, Massachusetts General Hospital, Boston, MA, United States

Background: Standard adverse event (AE) monitoring only records whether events occur after the intervention, and not whether these events vary over time.

Objective: To test whether there were statistically distinct time-varying trajectories of AE (e.g., "side effects") after an intervention and identify characteristics of individuals associated with these patterns.

Design: Group-based trajectory models applied to an observational study of individuals who received one or two doses of a mRNA COVID-19 vaccine (i.e., the intervention).

Participants: 50,484 healthcare personnel who received their vaccinations within the Mass General Brigham (MGB) healthcare system.

Interventions: Vaccination.

Main measures: Allergic and non-allergic AE for 1–3 days after each of two COVID-19 vaccinations.

Key results: Trajectories models identified distinct groups with different trajectories after intervention: two groups after the first vaccination and five groups after the second vaccination. These groups differed by demographics, age, prior prescription for epinephrine auto-injectors, prior COVID-19 history, time-of-day of vaccination, and vaccine manufacturer.

Conclusion: Several different time-based trajectories after the intervention (e.g., first two COVID-19 vaccinations) were noted; individuals in these groups varied by demographic and clinical criteria. These time-based methods may be able to identify groups at higher risk of future adverse reactions, provide a basis for future studies of the physiology underlying these risk differentials, and improve counseling surrounding interventions associated with AEs. We suggest that trajectory-based methods be added to post-intervention surveillance.

KEYWORDS

side-effects, adverse events, trajectory analyses, COVID-19 adverse effects monitoring, vaccines, adverse effects following immunization (AEFI)

Introduction

Clinical signs and symptoms, including adverse events (e.g., “side effects,” adverse effects following immunization (AEFI)) from drugs, vaccines, and other interventions, may occur in clusters and vary over time. Standard adverse event monitoring, however, only records whether an event has occurred after the intervention (1). With access to longitudinal data from a large cohort of employees from a single healthcare system after their first and second doses of COVID-19 vaccinations, we could utilize time-varying group-based analytic models to define distinct groups of individuals with differing time-varying vaccine AEs. We hypothesized that different time-varying clusters would be found and that these clusters would have differential risk factors (e.g., sex, age, history of allergy). We propose that trajectory of symptoms could be useful for future AE studies after any intervention (e.g., drug, vaccine, surgery); different risk factors and trajectories over time could enable targeting physiological mechanisms, be used for patient and clinician education, and be used for patient planning (e.g., schedule intervention when fewer AEs are expected), which could improve compliance and reduce hesitancy for the intervention.

Methods

In this observational study, data were collected from 50,484 healthcare personnel who received one or two doses of either the Pfizer or Moderna mRNA COVID-19 vaccine between December 2020 and April 2021 within the Mass General Brigham (MGB) healthcare system. Demographic data, allergy history, prior epinephrine auto-injector prescription, prior COVID-19 history, vaccine manufacturer (Pfizer or Moderna), and time-of-day of vaccination data were retrieved from the MGB electronic health record’s COVID-19 Datamart. Self-reported symptoms for 1–3 days after each vaccination were collected using a REDCap survey that was sent daily for 3 days after each vaccination. Detailed information about the REDCap AE survey has been previously reported (2, 3).

Group-based trajectory models (4) were utilized to identify patterns of participant-reported post-vaccination AEs over the 3 days following each vaccination dose. Participants exhibiting similar patterns of daily AEs were clustered into distinct subgroups using logistic modeling of dichotomous outcomes conditioned on group membership. The optimal number of groups was determined using Akaike (AIC) and Bayesian (BIC) information criteria. A second set of logistic regression models was then employed to explore the associations between these trajectory subgroups and potential predictors. Any symptoms of rash, hives, swelling, and wheezing were grouped together as any allergic-type AEs and coded as a binary outcome (yes, no). Symptoms such as new headache, fatigue, joint pain, muscle pain, and fever were grouped as “any non-allergic AE” and coded as (lower/higher severity, none). Allergic-type and non-allergic AEs were further combined into a single outcome: any allergic-type and/or any non-allergic AEs (yes or no). If on any day after vaccination, a participant did not respond to at least one of the 9 symptom questions (e.g., rash, hives, swelling, and wheezing, new headache, fatigue, joint pain, muscle pain, or fever), the AE outcome for that day was coded as “Missing.” Potential factors considered in the models included age group (18–40, 41–60; >60), gender, race/ethnicity groups (White/non-Hispanic, Non-white/non-Hispanic, Any race-Hispanic (“Hispanic”), Any race/other Ethnicity [including not

reported]), epinephrine auto-injector prescription, vaccine manufacturer (Pfizer, Moderna), time-of-day of vaccination (6:00–10:59, 11:00–15:59, 16:00–21:59), prior allergy history (yes, no), and any prior COVID-19 diagnosis or positive test (yes, no). Adjusted Odds Ratios (ORs) of any AEs and their 95% confidence intervals (CIs) were reported for each predictor in the multivariable logistic regression models. We examined the relationships both after the first vaccination and the second vaccination doses. All trajectory analyses were repeated using two approaches: (1) using all available data, assuming data were missing at random (“Full”); and (2) using “Complete-case” subsets, where there was no missing data for all 3 days following the first vaccination dose and all 6 days following both doses (3 days after each dose). No imputation methods were used, as the longitudinal data patterns were complex and there were no reasonable assumptions about the missing data distribution that could have led to unreliable estimates and artificially inflated precision. All analyses were performed using SAS Version 9.4. Group-based trajectory analyses were conducted using PROC TRAJ.

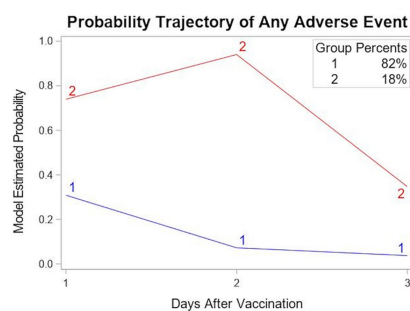
These analyses were conducted under MGB Human Subjects (IRB) Protocol 2021P001080. The IRB waived the requirement for informed consent.

Results

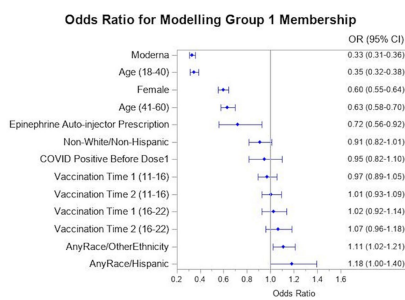
Data analysis of AEs after the first vaccination included 50,484 participants, of whom 34,803 had non-missing data for all three follow-up days. Among these, 50,270 participants received the same vaccine brand for both doses. The Complete-case subset for AEs following both doses consisted of 24,529 participants with no missing data across all six follow-up days. Results for the Full dataset (with missing data) are in the [Supplementary material](#).

After the first vaccination dose, two distinct groups were identified ([Figure 1A](#)) for the Complete-case subset. Group 1 members experienced minimal AEs only on the first day post-vaccination. In contrast, Group 2 members (18%) reported more severe symptoms on the second day but showed significant improvement by the third day. A forest plot for the adjusted odds ratios of any AEs based on the logistic regression model for the Group 1 membership is shown in [Figure 1B](#). Participants with a history of prior COVID infection, those in the younger age groups (18–40 or 41–60), females, and those with a prior prescription for an Epinephrine Auto-injector had a higher likelihood of belonging in Group 2 ([Figure 1C](#); note that [Figures 1B,C](#) are complementary since there are only two groups). Based on these results, individuals with the following characteristics would be more likely to have some symptoms on the first day after vaccination (i.e., Group 2 compared to Group 1), more symptoms on the second day, and fewer symptoms on day 3 than day 1: age 18–40 years (OR = 1.85, 95% CI = (1.68, 2.04)) or 41–60 years (OR = 1.48, 95% CI = (1.34, 1.64)), Female (OR = 1.75, 95% CI = (1.63, 1.88)), Non-White Non-Hispanic (OR = 1.26, 95% CI = (1.16, 1.36)) or Hispanic (OR = 1.43, 95% CI = (1.27, 1.62)), Epinephrine Auto-injector prescription (OR = 1.45, 95% CI = (1.24, 1.70)), history of COVID positive test before first vaccine dose (OR = 2.16, 95% CI = (1.97, 2.37)), and/or dosage with Moderna vaccine (OR = 1.20, 95% CI = (1.14, 1.27)).

After the second vaccine dose, five distinct groups were identified based on the AEs reported ([Figure 2A](#)) for the Complete-case subset. Group 1 members experienced minimal AEs at all time points. A forest plot for the adjusted odds ratios of any AEs based on the multivariable binary logistic regression model for the Group1 membership is shown



A



B

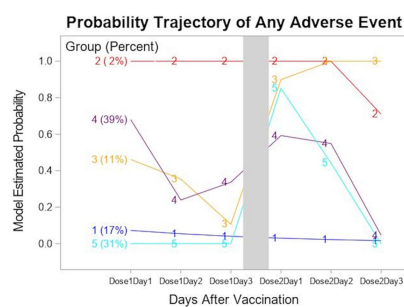
Odds Ratios for Modelling Group 2 Membership

Effect	OR (95% CI)
Age (18–40)	1.85 (1.68–2.04)
Age (41–60)	1.48 (1.34–1.64)
Female	1.75 (1.63–1.88)
Non-White/Non-Hispanic	1.26 (1.16–1.36)
AnyRace/Hispanic	1.43 (1.27–1.62)
AnyRace/OtherEthnicity	1.00 (0.93–1.07)
Epinephrine Auto-injector Prescription	1.45 (1.24–1.70)
COVID Positive Before Dose1	2.16 (1.97–2.37)
Moderna	1.20 (1.14–1.27)
Vaccination Time 1 (11–16)	1.02 (0.96–1.09)
Vaccination Time 1 (16–22)	0.96 (0.89–1.04)

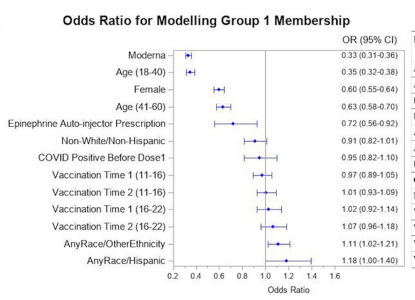
C

FIGURE 1

(A) Estimated probability and group membership using longitudinally self-reported AEs of any symptoms on days 1, 2, or 3 after vaccination dose 1 using the Complete-case dataset ($N = 34,803$). (B) Forest plot for the adjusted odds ratios of any AEs based on the multivariable binary logistic regression model for Group 1 membership. Reference values: no prior COVID positive diagnosis, age >60, Male, no Epinephrine Auto-injector, White/non-Hispanic race/ethnicity, Pfizer vaccine, and Vaccination time 6–11 am. (C) Adjusted odds ratios of any AEs based on the multivariable nominal logistic regression model for Group 2 membership using Group 1 as the reference group. Groups in bolded text were significant at $p < 0.05$. (B,C) Are complementary—with different reference groups.



A



B

Odds Ratios for Modelling Group 2–5 Membership

Effect	Group 2	Group 3	Group 4	Group 5
OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)
Age (18–40)	5.79 (3.78–8.86)	3.18 (2.73–3.71)	3.59 (3.14–3.91)	2.15 (1.93–2.40)
Age (41–60)	3.23 (2.18–4.96)	1.85 (1.59–2.16)	1.73 (1.55–1.92)	1.35 (1.21–1.51)
Female	3.93 (2.89–5.32)	2.50 (2.22–2.82)	1.80 (1.66–1.96)	1.33 (1.22–1.44)
Non-White/Non-Hispanic	1.92 (1.47–2.51)	1.22 (1.05–1.41)	1.15 (1.03–1.29)	0.97 (0.86–1.10)
AnyRace/Hispanic	1.44 (1.05–2.19)	1.04 (0.83–1.31)	0.90 (0.75–1.08)	0.69 (0.57–0.84)
AnyRace/OtherEthnicity	0.90 (0.68–1.18)	0.86 (0.76–0.97)	0.89 (0.82–0.98)	0.92 (0.84–1.02)
Epinephrine Auto-injector Prescription	2.20 (1.29–3.75)	1.75 (1.27–2.42)	1.61 (1.23–2.10)	0.95 (0.70–1.27)
COVID Positive Before Dose1	2.02 (1.46–2.80)	0.84 (0.67–1.04)	1.37 (1.17–1.60)	0.68 (0.57–0.81)
Moderna	3.23 (2.60–4.00)	3.82 (3.44–4.24)	2.51 (2.32–2.71)	3.45 (3.19–3.74)
Vaccination Time 1 (11–16)	0.90 (0.70–1.15)	1.05 (0.93–1.19)	1.00 (0.91–1.09)	1.08 (0.98–1.19)
Vaccination Time 1 (16–22)	0.89 (0.66–1.20)	1.04 (0.90–1.21)	0.87 (0.78–0.98)	1.10 (0.98–1.24)
Vaccination Time 2 (11–16)	1.24 (0.97–1.60)	1.22 (1.08–1.38)	1.01 (0.92–1.10)	0.91 (0.83–1.00)
Vaccination Time 2 (16–22)	1.36 (1.01–1.84)	1.27 (1.10–1.48)	0.91 (0.81–1.02)	0.87 (0.77–0.97)

C

FIGURE 2

(A) Estimated probability and group membership using longitudinally self-reported AEs of any symptoms on days 1, 2, and 3 after each of vaccination doses 1 and 2 using the Complete-case dataset ($N = 24,529$). (B) Forest plot for the adjusted odds ratios of any AEs based on the multivariable binary logistic regression model for Group 1 membership. Reference values as in Figure 1. (C) Adjusted odds ratios of any AEs based on the multivariable nominal logistic regression model for Group 2–5 membership using Group 1 as the reference group. Groups in bolded text were significant at $p < 0.05$.

in Figure 2B. Group 2 members (2% of the participants) consistently experienced AEs after both doses. Group 3 members experienced some AEs after the first dose and significantly worse AEs after the second dose. Group 4 members reported AEs on the first day after each dose and improved thereafter. Group 5 members had minimal AEs after the first dose but experienced more AEs on the first day after the second dose. Individuals of younger age, female gender, and those who received the Moderna vaccine are associated with Group 2 (all $p < 0.0001$), showing the highest likelihood of experiencing AEs throughout all days following vaccinations. Adjusted odds ratios of any AEs based on the multivariable logistic regression model for Group 2–5 membership using Group 1 as the reference group are shown in Figure 2C.

In this analysis of data from both doses, individuals in Group 1 were more likely to report having no symptoms after both vaccinations. Group 1 characteristics included age over 60 years, Male, Hispanic, no Epinephrine Auto-injector prescription, and/or dosage with Pfizer vaccine. Compared with Group 1 individuals, those more likely to have symptoms on all 6 days of symptom monitoring (days 1–3 after each

dose) (Group 2) were more likely to have the following characteristics: ages 18–40 (OR = 5.79, 95% CI = (3.78, 8.86)) or 41–60 (OR = 3.23, 95% CI = (2.10, 4.98)) years, Female (OR = 3.93, 95% CI = (2.89, 5.32)), Non-White/Non-Hispanic (OR = 1.92, 95% CI = (1.47, 2.51)), Vaccine dose 2 time 4–10 pm (OR = 1.36, 95% CI = (1.01, 1.84)), Epinephrine Auto-Injector Prescription (OR = 2.20, 95% CI = (1.20, 3.75)), COVID Positive before Dose 1 (OR = 2.02, 95% CI = (1.45, 2.80)) and/or dosage with Moderna vaccine (OR = 3.23, 95% CI = (2.60, 4.00)). For individuals with time varying probability of AEs (Groups 3, 4, 5), the different characteristics were Race and Ethnicity, history of COVID positivity before Dose 1, and/or vaccine time.

These patterns remained similar for most subgroups in analyses using the Full dataset (see Supplementary material).

Discussion

We identified multiple groupings of time-based symptom trajectories after COVID-19 vaccination (i.e., an intervention) with

different demographic and/or clinical characteristics for the groups. Our work is innovative in demonstrating that the *multiple* time-scale-based trajectory analyses yield additional information that may be clinically actionable, including (i) identifying those individuals with little risk of AE and those at higher risk – and the times of that increased risk—who might benefit from additional counseling or monitoring, and/or (ii) counseling individuals about when symptoms/signs may appear or lessen. The method applies to both allergic and non-allergic reactions, an area currently underrepresented in the literature, highlighting the need for such versatility in drug and vaccine safety, quality control, and patient compliance. Group-based trajectory methods provide an innovative way to monitor daily AEs across multiple doses of a drug or vaccine. One study of weight loss after medication in children performed trajectory analyses and found changes related to the drug and age of the child (5). Another study used trajectory analyses to monitor compliance with medication use (6). A study of response to multiple COVID-19 vaccines used a similar latent class analysis to compare adverse events occurring any time after each vaccine (7), not the time course of adverse events following each vaccine. Collection of the data needed for group-based trajectory analyses can be done semi-automatically using email links, text messages, and phone apps. This study had high response rates among healthcare providers (2); another study reported high response rates by patients who received automated text message for monitoring adverse events after immunization (8).

Limitations

There are multiple limitations to this first demonstration of this methodology: (i) The AEs data were collected from an observational trial. Future work should include trials with appropriate randomization. (ii) The cohort consisted only of health care employees in the northeastern US, and thus the study findings related to specific AE trajectories may not be generalizable to other populations. Study populations may vary in reporting bias; highly motivated individuals may be more likely to report positive, negative, or no outcomes. Future work should be conducted in other populations. (iii) The data were collected after response to one potential intervention; future work should be conducted with other interventions (e.g., medications, medical devices, surgery).

Conclusion

By analyzing trajectory patterns over time (rather than only the presence/absence of an adverse event at a fixed time), these methods may identify groups at higher risk of future adverse reactions, provide additional physiological information that may be useful for future basic and clinical studies of AEs, and improve counseling surrounding interventions associated with AEs. This approach allows for real-world post-marketing monitoring to track the evolution of AEs over time for single or multiple interventions (e.g., chemotherapy). Particularly useful in unpredictable, idiosyncratic reactions (in addition to relatively predictable and/or allergic

reactions), this work may help optimize the timing of medications, vaccinations, or other interventions and serve as a versatile tool for pharmaceutical trial design, medication administration, and patient safety.

Data availability statement

The data analyzed in this study is subject to the following licenses/restrictions: Raw data will not be shared because they were originally collected from healthcare personnel for monitoring post-vaccination. Requests to access these datasets should be directed to ebkclerman@mgh.harvard.edu.

Ethics statement

The studies involving humans were approved by MassGeneralBrigham Human Research Committee. The studies were conducted in accordance with the local legislation and institutional requirements. The Human Research Committee waived the requirement for informed consent.

Author contributions

WW: Methodology, Formal analysis, Conceptualization, Software, Visualization, Writing – original draft, Writing – review & editing. SA: Data curation, Writing – review & editing. KB: Methodology, Writing – review & editing. DH: Methodology, Writing – review & editing. GR: Methodology, Writing – review & editing. EK: Resources, Project administration, Methodology, Conceptualization, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. KB was partially supported by CDC 6NU50CK000622-02-05 (Shenoy).

Conflict of interest

WW declares an affiliation with the National Sleep Foundation. KB declares consulting fees from Denali Therapeutics, royalties from UpToDate, and a grant from Thermo Fisher Scientific. GR declares clinical support from Citius Pharmaceuticals. EK declares the following: (i) consulting fees from Circadian Therapeutics, the National Sleep Foundation, and Yale University Press; (ii) honoraria from the American Academy of Sleep Medicine Foundation, the Sleep Research Society, and the Sleep Research Society Foundation; (iii) travel support from the European Biological Rhythms Society, Lighten up/EPFL Pavilions, the Sante Fe Institute, the Sleep Research Society, the Society for Research in Biological Rhythms, and the World Sleep Foundation; and (iv) their partner is the founder and CSO of Chronsulting.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The authors declare that no Gen AI was used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

References

1. Suku CK, Hill G, Sabblah G, Darko M, Muthuri G, Abwao E, et al. Experiences and lessons from implementing cohort event monitoring programmes for antimalarials in four African countries: results of a questionnaire-based survey. *Drug Saf.* (2015) 38:1115–26. doi: 10.1007/s40264-015-0331-7
2. Shenoy ES, Wickner PG, West LR, Banerji A, Blumenthal KG, Centi AJ, et al. Symptom monitoring after coronavirus disease 2019 (COVID-19) vaccination in a large integrated healthcare system: separating symptoms from severe acute respiratory coronavirus virus 2 (SARS-CoV-2) infection. *Infect Control Hosp Epidemiol.* (2021) 43:1439–46. doi: 10.1017/ice.2021.449
3. Abbaspour S, Robbins GK, Blumenthal KG, Hashimoto D, Hopcia K, Mukerji SS, et al. Identifying modifiable predictors of COVID-19 vaccine side effects: a machine learning approach. *Vaccine.* (2022) 10:1747. doi: 10.3390/vaccines10101747
4. Nagin DS, Jones BL, Passos VL, Tremblay RE. Group-based multi-trajectory modeling. *Stat Methods Med Res.* (2018) 27:2015–23. doi: 10.1177/0962280216673085
5. Pozzi M, Pisano S, Marano G, Carnovale C, Bravaccio C, Rafaniello C, et al. Weight-change trajectories of pediatric outpatients treated with risperidone or aripiprazole in a naturalistic setting. *J Child Adolesc Psychopharmacol.* (2019) 29:133–40. doi: 10.1089/cap.2018.0092
6. Frank AS, Lupattelli A, Matteson DS, Nordeng H. Maternal use of thyroid hormone replacement therapy before, during, and after pregnancy: agreement between self-report and prescription records and group-based trajectory modeling of prescription patterns. *CLEP.* (2018) 10:1801–16. doi: 10.2147/CLEP.S175616
7. Yamamoto C, Kobashi Y, Kawamura T, Nishikawa Y, Saito H, Oguro F, et al. Group of longitudinal adverse event patterns after the fourth dose of COVID-19 vaccination with a latent class analysis. *Front Public Health.* (2024) 12:1406315. doi: 10.3389/fpubh.2024.1406315
8. Leeb A, Regan AK, Peters IJ, Leeb C, Leeb G, Effler PV. Using automated text messages to monitor adverse events following immunisation in general practice. *Med J Aust.* (2014) 200:416–8. doi: 10.5694/mja13.11166

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmed.2025.1637091/full#supplementary-material>