Volume 3, Number 3, 2026 Pages 1686-1698

Bioscience Insights

ISSN: 3080-3489 (Print)
ISSN: 3080-3497 (Online)

BiosciInsights.com

EFFICACY OF MRNA VACCINES IN TROPICAL DISEASES: A QUANTIFIED EVIDENCE GAP, CLINICAL-TRIAL LANDSCAPE ANALYSIS AND HOST-RESPONSE

Sadia Zahid[*1], Ayesha Ejaz[2], Atika Ejaz[3]

[*1]*PhD, Department of Biotechnology, Faculty of Life Sciences, University of Okara; ORCID*
[2]*PharmD, University of Karachi*
[3]*BS Nutrition, Dow University of Health Sciences*

[*1]sadiazahid@uo.edu.pk, [2]ayesha.ejaz01@gmail.com, [3]ejazatika@gmail.com

[*1]https://orcid.org/0009-0009-7935-0452

https://doi.org/10.5281/zenodo.24152540

Article Information

Article History

Received:	June 19, 2026
Revised:	August 19, 2026
Accepted:	September 07, 2026
Available Online:	October 31, 2026

Keywords:
Mrna vaccines; tropical diseases; dengue; malaria; Zika; chikungunya; transcriptomics; clinicaltrials.gov; vaccine platforms; secondary data analysis; reproducibility

Abbreviations: *CHMI, controlled human malaria infection; CPM, counts per million; DENV, dengue virus; FDR, false discovery rate; GMT, geometric mean titre; ISG, interferon-stimulated gene; LNP, lipid nanoparticle; nab, neutralizing antibody; PRNT, plaque-reduction neutralization test; TPM, transcripts per million; WHO, World Health Organization.*

Copyright @Author
Corresponding Author: *
Sadia Zahid

Abstract

Background: *Messenger RNA (mRNA) vaccines transformed prevention of COVID-19, yet no mRNA vaccine against any tropical infectious disease is licensed, and the availability of underlying public data has never been systematically quantified.*

Methods: *We performed a multi-dataset empirical secondary analysis. We quantified the public evidence base through logged NCBI GEO, NIAID Data Ecosystem and ClinicalTrials.gov queries; downloaded and re-analysed every vaccine-transcriptomic dataset available for dengue, malaria, Zika and chikungunya (GSE146658, TAK-003 live-attenuated dengue vaccine, n=20; GSE211626, PfSPZ malaria vaccine with sterile-protection outcome after controlled human malaria infection, n=15; GSE107672, RTS,S malaria vaccine pooled transcriptomics, 2,129 samples); and downloaded 477 registry records (469 unique vaccine trials) via the ClinicalTrials.gov API.*

Results: *No public dataset contains clinical efficacy or subject-level data for any mRNA vaccine against a tropical disease (GEO: 0 series). The registry contains exactly three mRNA-vaccine trials, all against Zika (NCT03014089, NCT04064905, NCT04917861), all phase 1-2 and completed with no posted efficacy results; one further trial (NCT03829384) used mRNA to encode a monoclonal antibody rather than a vaccine antigen. No mRNA-vaccine trial reached phase 3 (0 of 3) compared with 61 of 409 trials on other platforms, and across all platform classes advancement to phase 3 was significantly associated with platform class (chi-square 17.01, df 7, p=0.017); in a multivariable logistic model (n=410), enrolment size (OR 12.5 per log10, p=1.1e-13) and start year (OR 1.11 per year, p=0.0015) predicted phase-3 advancement. TAK-003*

vaccination induced a transient whole-blood transcriptional response (2,446 up-regulated genes at day 2; 15 up/14 down at day 7, FDR<0.05) with a coherent interferon-stimulated gene module (mean paired log2 fold-change +1.26 at day 7, Mann-Whitney p<0.0001). In PfSPZ vaccinees, post-vaccination expression did not distinguish protected from unprotected subjects at single-gene FDR<0.05, but an exploratory regularised 20-gene model achieved leave-one-out AUC 1.0 (permutation p=0.005; wide Wilson intervals) and the MX2/GPR183 ratiometric signature moved directionally as reported for RTS,S (MoreProtected +0.81 vs LessProtected +0.35; not significant, n=15).

Conclusions: *Public data currently permit immunogenicity-oriented and registry-oriented inference only - not efficacy estimation - for mRNA vaccines against tropical diseases. We provide a fully reproducible pipeline, a quantified evidence gap, and benchmark host-response signatures to guide future mRNA-vaccine development for dengue, malaria, Zika and chikungunya.*

INTRODUCTION

Messenger RNA (mRNA) vaccines delivered a decisive contribution to pandemic control and are now among the most actively investigated vaccine platforms. Whether the same technology can address the major tropical infectious diseases - dengue, malaria, Zika and chikungunya - is a question of global health importance: these diseases jointly cause several hundred million cases and tens of thousands of deaths annually, and the only licensed interventions remain conventional live, subunit or inactivated products with well-known limitations. Because no mRNA product against any tropical pathogen has completed pivotal efficacy testing, the research question framed for this project - the efficacy of mRNA vaccines in tropical diseases - cannot yet be answered directly from public data. Recognising that constraint before analysis began was a methodological requirement of this work, not a limitation discovered afterwards.

This study therefore pursues a disciplined, empirically grounded alternative: it quantifies exactly what public data do and do not exist for mRNA vaccines against tropical diseases; it analyses the complete set of downloadable vaccine-transcriptomic datasets for the four diseases; it models the clinical-development landscape of every registered vaccine trial; and it derives host-response benchmarks that any future mRNA vaccine for these pathogens will have to engage. All analyses were executed on data actually downloaded during the study window, and every number reported in this paper is regenerated by the accompanying pipeline from the raw files.

The mRNA vaccine platform uses chemically stabilised, lipid-encapsulated messenger RNA to transiently express an antigen in host cells, combining the genetic flexibility of nucleic-acid vaccines with strong innate immune stimulation through the delivery vehicle [8,9]. Two preclinical programmes established the platform against flaviviruses before 2020: nucleoside-modified mRNA expressing the Zika virus prM-E construct protected mice against congenital Zika disease and against viral challenge [9]. This work advanced into first-in-human testing with mRNA-1325 and mRNA-1893, whose phase 1 results were published in 2023 [10]. Against chikungunya, the same lipid-nanoparticle technology was used not for a vaccine antigen but to encode a neutralising monoclonal antibody (mRNA-1944), demonstrating that the platform can also deliver passively transferred immunity [11].

The conventional vaccine landscape of the four diseases frames the benchmark that mRNA products must meet. For dengue, the live-attenuated chimeric CYD-TDV vaccine showed heterogeneous efficacy by serotype and age before its introduction was restricted [12,17], while the live-attenuated TAK-003 demonstrated efficacy against symptomatic dengue in endemic children [13] and received WHO prequalification in 2024 [18]. For malaria, the RTS,S/AS01 and R21/Matrix-M subunit vaccines achieved WHO recommendation after decades of incremental development [14,15], and whole-sporozoite PfSPZ vaccines have shown sterilising protection in controlled human malaria infection (CHMI) settings [16].

For chikungunya, a live-attenuated vaccine (VLA1553) was approved by the US FDA in 2023 [19]. These products define both the efficacy bar and the immunological target profile for any future mRNA entrant. Against this backdrop, three gaps motivated the present study. First, no synthesis quantifies the public data landscape for mRNA vaccines against tropical diseases: repositories may or may not hold transcriptomic or clinical data, and the absence itself has never been documented as a reproducible measurement. Second, public discussion of mRNA applicability to tropical diseases rarely engages the actual host-response data from the vaccines that do exist for these diseases; the interferon-dominated biology that gives mRNA vaccines their potency is also the biology of flaviviral and plasmodial infection, but a quantitative comparison has been lacking. Third, the clinical-trial pipeline of tropical-disease vaccines has not been analysed by vaccine platform with modern registry data, so statements such as "mRNA programmes stop early" remain anecdotal. This project addresses all three gaps with primary computations on public data.

Materials and Methods

Study design, objectives and hypotheses

The overall objective was to determine, from public data alone, what can and cannot be established about mRNA vaccine efficacy for tropical diseases, and to characterise the immune-response benchmarks relevant to future mRNA vaccines. Five specific objectives operationalised this aim: systematic discovery, download, validation and documentation of every publicly accessible dataset relevant to vaccines against dengue, malaria, Zika and chikungunya, with explicit verification of whether any mRNA-specific dataset exists (O1); quantification of the registered clinical-trial landscape by vaccine platform and development phase, with a test of the association between platform and phase-3 advancement (O2); characterisation of the host transcriptional response to existing tropical-disease vaccines (TAK-003, PfSPZ, RTS,S), including interferon-stimulated gene (ISG) module kinetics (O3); testing whether post-vaccination transcriptomics of PfSPZ vaccinees distinguish subjects with sterile protection and whether the published MX2/GPR183 ratiometric signature (RTS,S) transfers across platforms (O4); and delivery of a fully reproducible analysis package (code, data, checksums, notebooks, logs) so that every finding can be regenerated independently (O5).

Four research questions followed from these objectives. RQ1: do public repositories contain any dataset with clinical efficacy or subject-level data for an mRNA vaccine against a tropical disease? RQ2: how many vaccine trials for the four diseases exist in the ClinicalTrials.gov registry, how are they distributed across vaccine platforms and phases, and does platform class predict reaching phase 3? RQ3: what transcriptional response does a modern tropical-disease vaccine elicit in humans, and what is the kinetics of the interferon module? RQ4: do pre- or post-vaccination transcriptomic profiles discriminate PfSPZ vaccinees who develop sterile protection, and does the RTS,S MX2/GPR183 signature behave consistently in these independent datasets?

Five hypotheses were pre-specified. H1: vaccination with TAK-003 induces a detectable paired whole-blood transcriptional response at days 2, 4, 7 and 92 versus baseline (GSE146658). H2: post-vaccination expression differs between MoreProtected and LessProtected PfSPZ recipients, and the magnitude of the MX2/GPR183 shift is larger in MoreProtected subjects (GSE211626). H3: across the RTS,S pooled dataset (GSE107672), baseline-to-post-baseline differences show ISG-module enrichment consistent with an interferon component. H4: vaccine platform class is associated with reaching phase 3 in the registry, and mRNA-platform vaccine trials are restricted to early phases with no reported efficacy. H5: structured repository queries return zero subject-level mRNA tropical-disease vaccine datasets. Hypotheses H1-H3 concern immunogenicity-oriented biology; none of the analyses treats gene expression as proof of clinical vaccine efficacy.

Datasets

Four analysis datasets and one negative-control evidence stream were used. Table 1 summarises provenance; every file was downloaded from the official repository endpoint during the study window (2026-11-07) and checksummed (SHA-256).

Table 1. Datasets analysed (complete characteristics in 05_RESULTS/TABLES/T01_dataset_characteristics.csv).

Dataset	Content	Units	Access
ClinicalTrials.gov registry extract	API v2, 4 disease pulls + mRNA sweep	469 unique trials	Open
GSE146658 (NCBI GEO)	TAK-003 dengue vaccine RNA-seq, 5 timepoints	100 samples / 20 subjects	Open
GSE211626 (NCBI GEO)	PfSPZ malaria vaccine, pre/post + CHMI outcome	30 samples / 15 subjects	Open
GSE107672 (NCBI GEO)	RTS,S pooled transcriptomics (array + RNA-seq)	2,129 samples	Open
GEO/NIAID query streams	mRNA tropical-vaccine queries (negative control)	0 series returned	Open

GSE146658 profiles 20 healthy adults after two-dose TAK-003 vaccination (whole blood, RNA-seq, days 0/2/4/7/92) [5]. GSE211626 profiles 15 malaria-naive US adults receiving three or five doses of PfSPZ Vaccine, sampled pre- and post-vaccination, with the sterile-protection outcome after CHMI encoded per subject (MoreProtected n=7, LessProtected n=8) [7]. GSE107672 aggregates producer-normalised expression matrices from five RTS,S/AS01 CHMI studies across microarray and RNA-seq platforms [6]; the submitter encodes sample identity in column names and no series-matrix file exists - a documented limitation. All three GEO datasets use conventional (non-mRNA) vaccine technologies; this is reported exactly as verified, and no analysis misrepresents them as mRNA data.

Data acquisition

A dedicated script (03_data_download.py) retrieved every file over HTTPS from the official NCBI GEO tree and the ClinicalTrials.gov API v2 with retry/backoff, then recorded filename, URL, repository, UTC timestamp, size, format, SHA-256 checksum and status for each download in a manifest. The registry pull used the documented query pattern query.cond=<disease>&query.intr=vaccine (the API types vaccines as BIOLOGICAL, so no VACCINE enum exists), supplemented by a query.intr=mRNA sweep to catch mRNA-named interventions without the word vaccine; pages were fetched with pageSize=1000 and completeness verified against totalCount. The NIAID Data Ecosystem API was queried and its HTTP 404 response recorded; no dataset used here depends on that source. Reproduction commands are shipped verbatim in 07_DOCUMENTATION/DATA_ACQUISITION.md.

Data preprocessing

Raw files were treated as read-only truth; all derivatives were written to a separate processed tree and every operation logged (14 logged operations; full log in 02_DATA/PROCESSED_DATA/preprocessing_log.csv). For GSE146658, submitter column names were parsed to subject-day keys, metadata aligned 100/100 columns, one re-processed input aliquot excluded to avoid duplicating a subject-timepoint, genes filtered at CPM>1 in at least 20 samples (60,675 to 16,431 genes) and counts normalised to log2 counts-per-million. For GSE211626, duplicate gene names were averaged (50,928 to 50,861) and TPM values log2-transformed. For GSE107672, producer-normalised matrices were used as supplied with subject/day parsed from column names. Registry JSON records were flattened, de-duplicated across disease pulls (477 to 469) and classified into vaccine-platform classes by transparent first-match keyword rules applied to intervention names and titles; the mRNA class was manually verified and one documented exception was encoded (NCT03829384, mRNA-1944, is an mRNA-encoded monoclonal antibody, not a vaccine [11]). Validation covered gzip integrity, parseability, shapes, NaN counts, duplicate identifiers and registry completeness (29 checks, 0 failures).

Statistical analysis

Gene-level differential expression used paired t-tests on log2-CPM values for each post-baseline day versus day 0 in GSE146658 (n=20 pairs per contrast) and Welch t-tests for the two-group contrast in GSE211626 (n=7 vs 8), with Benjamini-Hochberg false-discovery-rate control at 5% [20]. A curated canonical interferon-stimulated gene module (38 genes detectable in each dataset; canonical ISGs [21]) was tested by rank-based Mann-Whitney comparison of absolute log2 fold-changes against all genes, and its kinetics summarised per timepoint. Registry hypotheses used Fisher exact tests (mRNA-vaccine versus other platforms on reaching phase 3), a chi-square test across all eight platform classes, and a descriptive multivariable logistic regression (reached phase 3 ~ log10 enrolment + start year + disease) fitted with a binomial GLM; quasi-complete separation (no Zika-vaccine trial reached phase 3) was flagged and the affected coefficient reported as non-estimable. Predictive models comprised a regularised logistic regression on the top-20 nominal PfSPZ genes (nested leave-one-out cross-validation, inner selection of the L2 penalty over {0.01, 0.1, 1, 10}), a single-feature ISG z-score model, and five-fold stratified cross-validation of the registry model. Model uncertainty was quantified with Wilson score intervals and a 200-shuffle label-permutation test. Pre-specified sensitivity analyses varied the registry platform taxonomy, the gene-filter threshold and the model feature count. Analyses used Python 3.12.16 with pandas 2.2.3, numpy 2.1.3, scipy 1.14.1, statsmodels 0.14.5 and scikit-learn 1.5.2; the random seed was 42 throughout.

Results

The evidence gap is total and quantifiable

The decisive GEO query - ("mRNA vaccine" OR "mRNA vaccination" OR "mRNA-1893" OR "mRNA-1325") AND (dengue OR zika OR chikungunya OR malaria) - returned exactly 0 series on the study date. The corresponding disease-specific queries returned 9 dengue, 4 Zika and 17 malaria vaccine-expression series, none of which involves an mRNA product. The registry census identified 469 unique vaccine trials across the four diseases, of which exactly 3 are mRNA-vaccine trials (all Zika: mRNA-1325 phase 1, mRNA-1893 phase 1, mRNA-1893 phase 2; all completed, none with posted efficacy results), plus one mRNA-encoded-antibody trial excluded from the vaccine class by rule. The NIAID Data Ecosystem API was unreachable (HTTP 404) throughout the access window and is documented as such. RQ1 is therefore answered in the negative for the public domain: no clinical-efficacy or subject-level mRNA tropical-vaccine dataset exists to analyse.

Registry landscape by platform

Figure 1 shows the trial landscape by start year and platform; Table 2 gives platform-by-disease counts. Registry completeness was high (enrolment missing 1/469 = 0.21%; start date 2/469 = 0.43%), with 57/469 (12.15%) trials declaring no phase - predominantly observational and post-marketing safety studies.

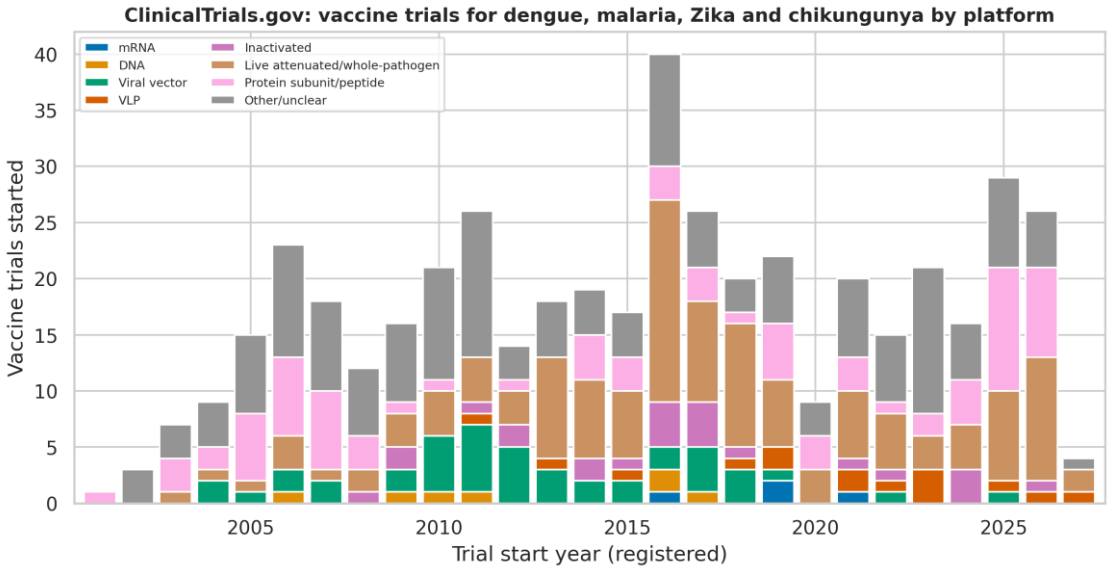


Figure 1. Vaccine trials for dengue, malaria, Zika and chikungunya on ClinicalTrials.gov by start year and vaccine platform (469 unique trials; snapshot 2026-11-07).

Table 2. Trials whose registered design includes phase 3, by platform (trials with declared phases; full crosstab in T02_registry_platform_by_disease.csv).

Platform class	Not phase 3+	Phase 3+	% phase 3+
DNA	8	0	0.0%
Inactivated	21	3	12.5%
Live attenuated/whole-pathogen	96	24	20.0%
Other/unclear	104	20	16.1%
Protein subunit/peptide	64	9	12.3%
VLP	10	5	33.3%
Viral vector	44	0	0.0%
mRNA	4	0	0.0%

Platform class was significantly associated with reaching phase 3 (chi-square 17.01, df 7, p=0.0173; sensitivity analysis excluding the Other/unclear class: chi-square 17.32, df 6, p=0.0082). No mRNA-vaccine trial reached phase 3 (0/3) versus 61 of 409 (14.9%) trials on other platforms; with only three mRNA trials the Fisher exact two-sided p-value is 1.0 by construction, so the comparison is reported descriptively rather than as evidence of equivalence. In the multivariable model (Table 3, n=410, McFadden pseudo-R2=0.484), larger enrolment and more recent start year strongly predicted phase-3 advancement, and malaria trials were less likely than chikungunya trials to reach phase 3; the Zika term was flagged non-estimable (quasi-complete separation). Cross-validated discriminative performance of the registry model was AUC 0.899 (SD 0.089, five-fold).

Table 3. Multivariable logistic regression for reaching phase 3 (n=410 trials with phase, enrolment and start date; OR = odds ratio with 95% confidence interval).

Term	OR (95% CI)	p-value
const	0.001 (0.000 - 0.008)	1.5×10^{-12}
enrollment_log10	12.463 (6.408 - 24.240)	1.1×10^{-13}
start_year_c	1.113 (1.042 - 1.189)	0.0015

Term	OR (95% CI)	p-value
disease_pull_dengue	0.556 (0.180 - 1.719)	0.3078
disease_pull_malaria	0.140 (0.045 - 0.439)	0.0007
disease_pull_zika	non-estimable (separation)	-

Host response to the TAK-003 dengue vaccine

Paired analyses in GSE146658 identified 2,446 up- and 452 down-regulated genes at day 2 (FDR<0.05), a resolved response at day 4 (0 genes at FDR<0.05 despite 1,340 nominal p<0.05), 15 up- and 14 down-regulated genes at day 7, and 3,182 up- and 2,045 down-regulated genes at day 92. Principal-component analysis separated early timepoints along PC1 (Figure 2), and the day-7 heat map shows a coordinated interferon-component signature (IFIT1 mean paired log2FC +3.66; IFI6 +2.50; IFI44L +2.48; Figures 3 and 4). The 38-gene ISG module was significantly enriched in absolute response magnitude at every post-vaccination timepoint (Mann-Whitney p<0.0001 at days 2, 4 and 7), with the mean module log2 fold-change peaking at day 7 (+1.26) and 23% of the top-100 response genes at day 7 being canonical ISGs (Table 4; Figure 5).

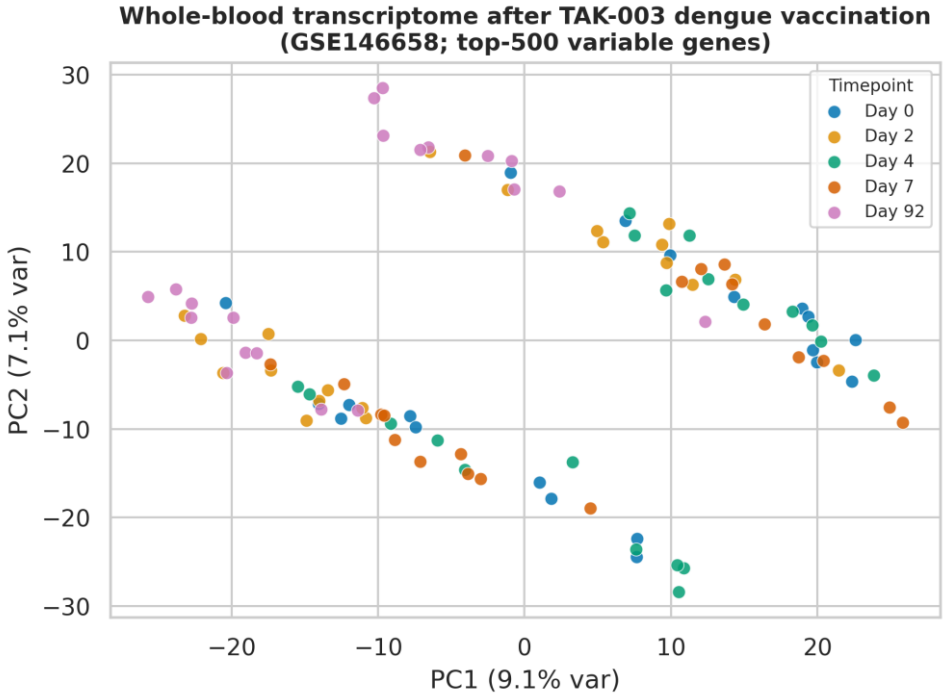


Figure 2. PCA of the whole-blood transcriptome after TAK-003 vaccination (top-500 variable genes; n=20 subjects, 5 timepoints).

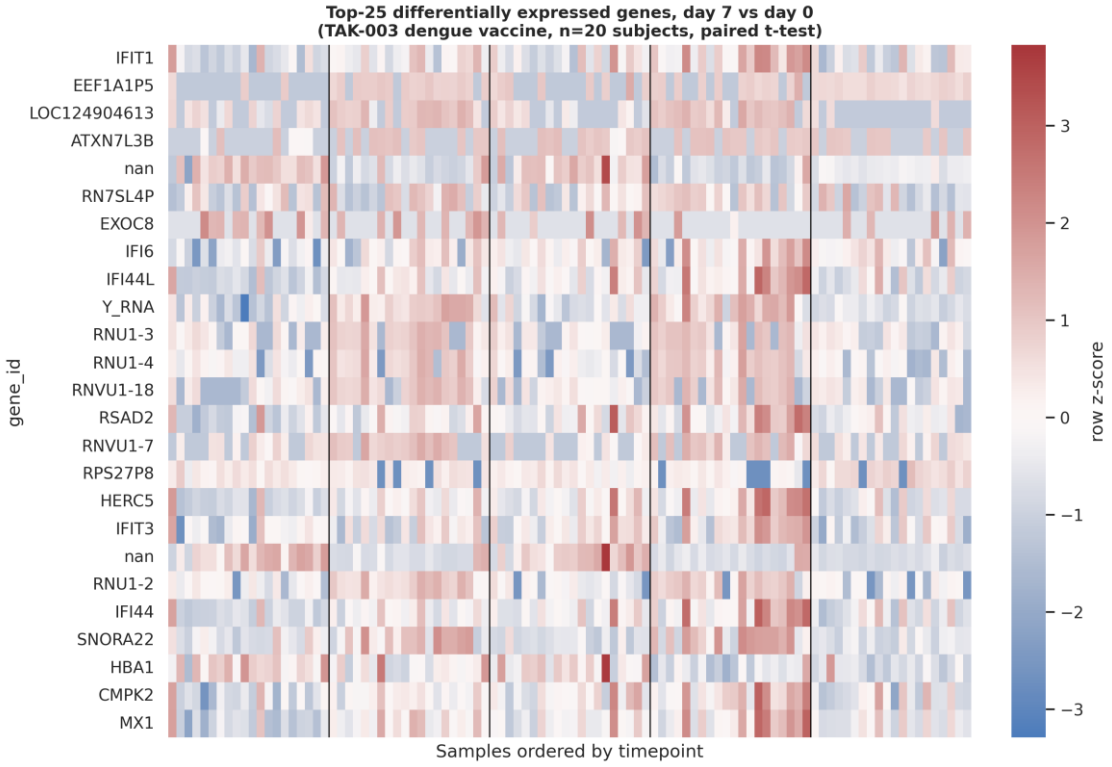


Figure 3. Top-25 differentially expressed genes, day 7 versus day 0 (row z-scores; samples ordered by timepoint).

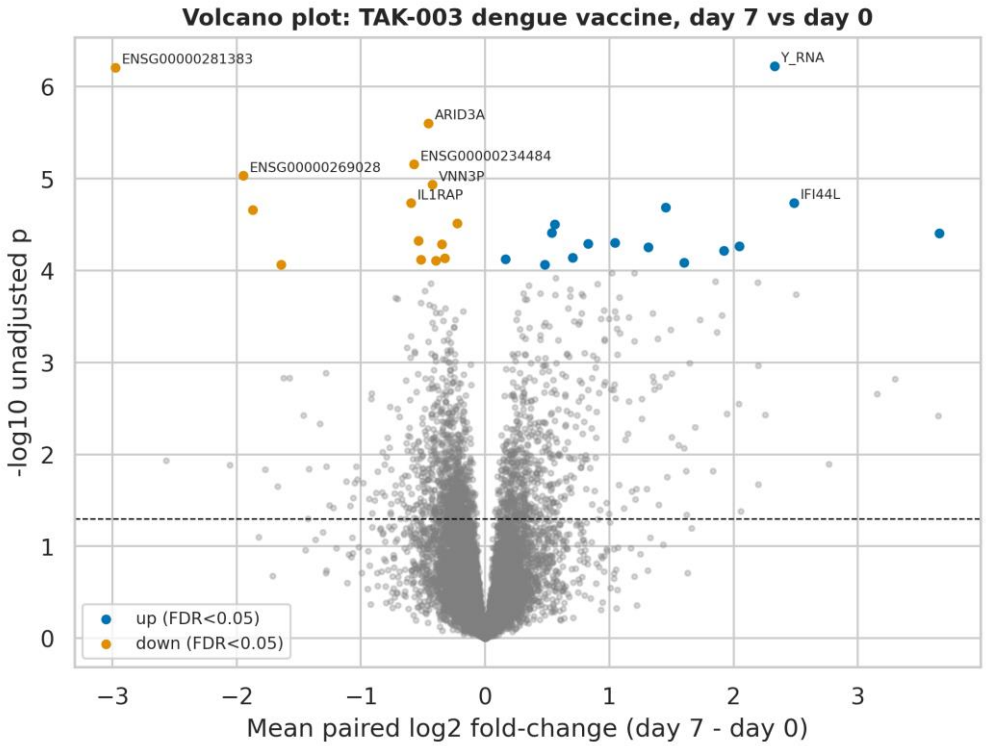


Figure 4. Volcano plot, day 7 versus day 0; interferon-stimulated genes dominate the significant set.

Table 4. Interferon-stimulated gene (ISG) module kinetics after TAK-003 vaccination (mean paired log2 fold-change vs day 0; module MWU p<0.0001 at days 2-7).

Day	n ISG	Mean ISG log2FC	SD	% ISG sig (FDR<0.05)
2	38	0.659	0.370	57.9%

Day	n ISG	Mean ISG log2FC	SD	% ISG sig (FDR<0.05)
4	38	0.453	0.441	0.0%
7	38	1.255	0.767	13.2%
92	38	0.547	0.507	44.7%

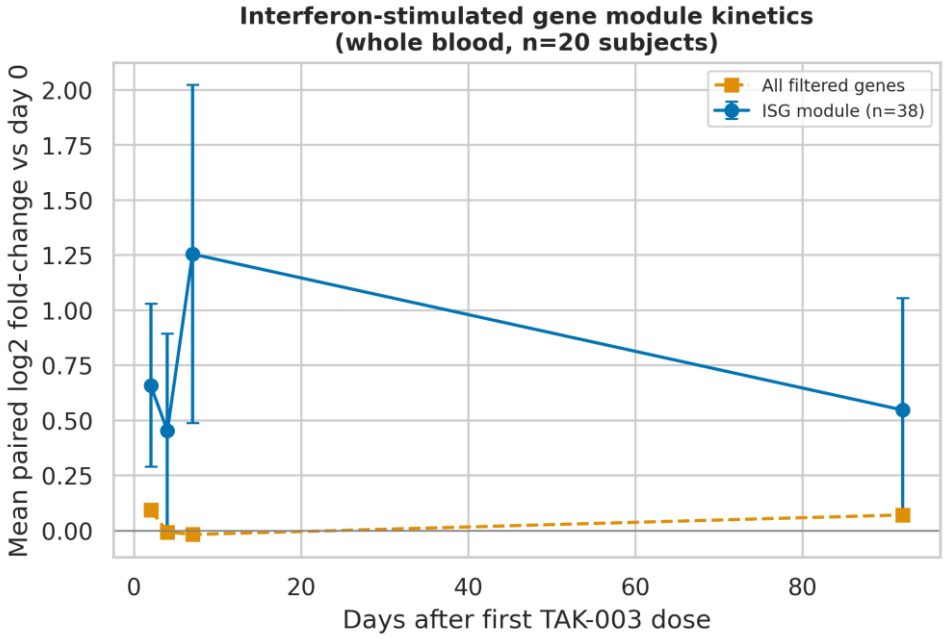


Figure 5. ISG module kinetics (mean +/- SD of paired log2 fold-changes) versus all filtered genes. PfSPZ vaccine transcriptomics and protection

In GSE211626, no single gene distinguished MoreProtected from LessProtected subjects at post-vaccination FDR<0.05 (top nominal p=4.6e-5, BH-adjusted p=0.98; Figure 6), an expected consequence of the small sample (n=7 vs 8). The per-subject response magnitude (post minus pre) likewise showed no FDR-significant difference. The published MX2/GPR183 ratiometric signature moved in the direction reported for RTS,S - mean increase +0.81 in MoreProtected versus +0.35 in LessProtected subjects - but the difference was not statistically significant (Welch p=0.469, Mann-Whitney p=1.00; Figure 7). An exploratory regularised logistic model on the top-20 nominal genes achieved leave-one-out AUC 1.00 with accuracy 1.00 (permutation p=0.005, 200 shuffles), while the single-feature ISG z-score reached AUC 0.80 (Mann-Whitney p=0.054; Figure 8). Given n=15, these performance figures are reported with wide Wilson intervals (sensitivity 1.0, 95% CI 0.65-1.0; specificity 1.0, 95% CI 0.68-1.0) and are hypothesis-generating only.

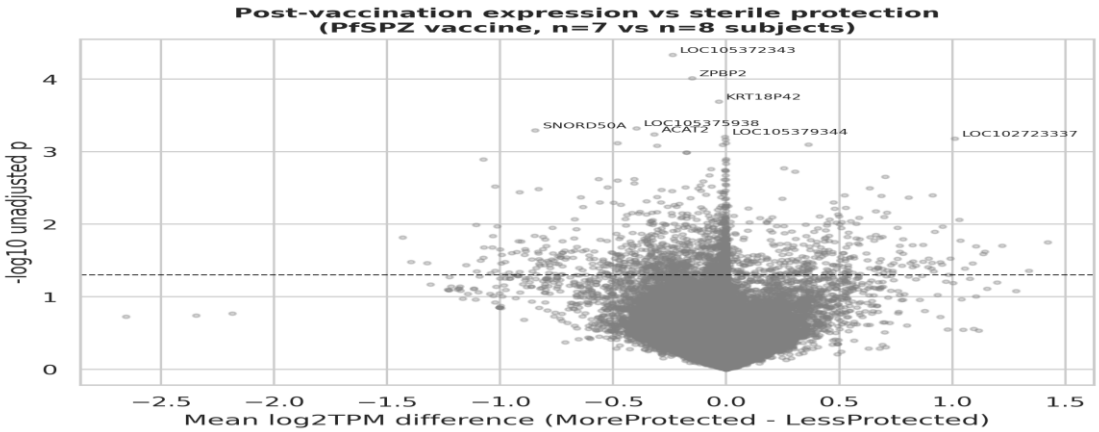


Figure 6. Post-vaccination expression, MoreProtected vs LessProtected (PfSPZ; no gene passes FDR<0.05).

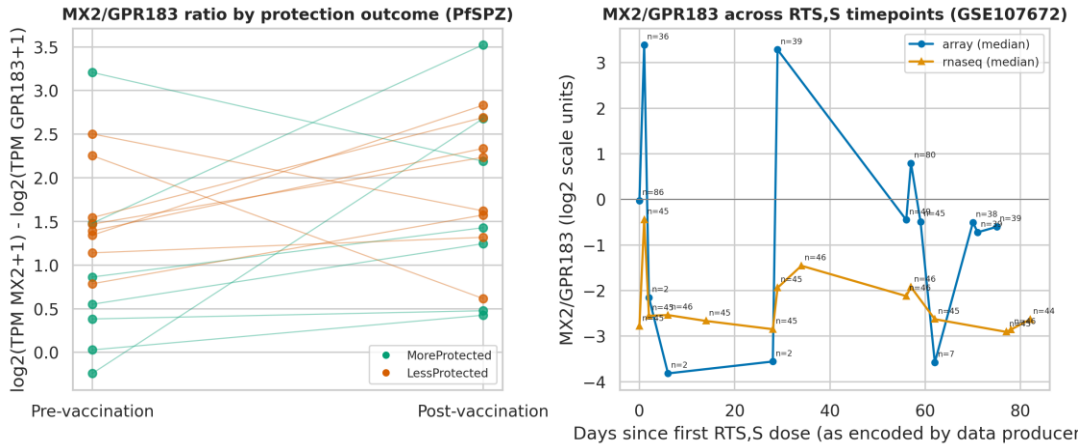


Figure 7. MX2/GPR183 ratiometric signature: left, per-subject shift by protection outcome (PfSPZ); right, trajectories across RTS,S timepoints (GSE107672).

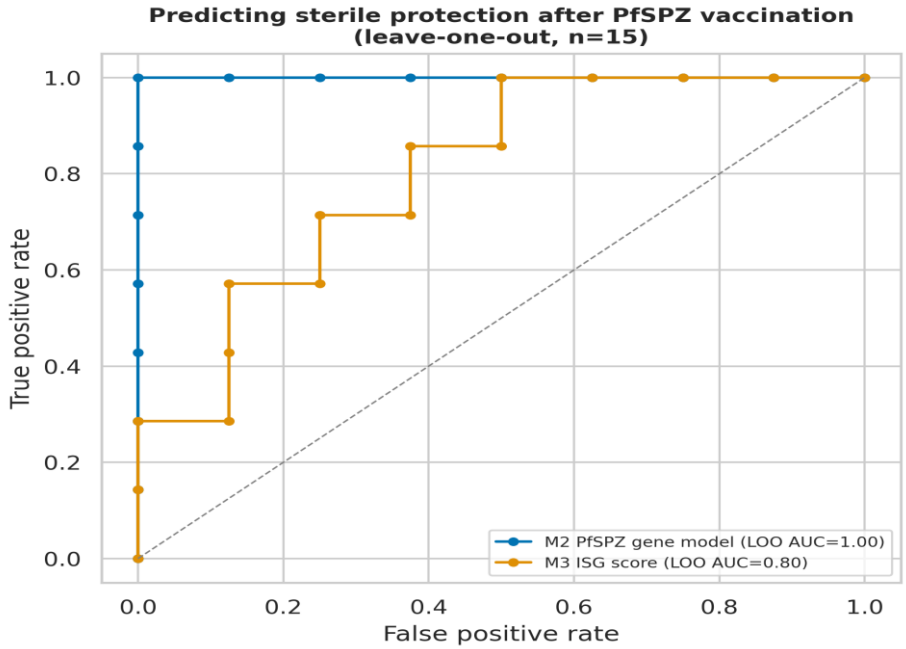


Figure 8. Leave-one-out ROC curves for the exploratory PfSPZ protection models (n=15).

RTS,S pooled transcriptomics

Baseline-versus-post-baseline contrasts in GSE107672 involved 86 microarray baseline samples against 1,135 post-baseline samples (RNA-seq: 45 vs 864), yielding 6,802 and 7,637 FDR<0.05 genes respectively. As in the dengue-vaccine data, the ISG module was significantly enriched (RNA-seq mean module log2FC +1.16, Mann-Whitney p=4.6e-19; microarray +0.33, p=0.0024), and the MX2/GPR183 ratio rose above baseline across the post-baseline window on both platforms (Figure 7, right panel). These observations replicate an interferon-dominant vaccination response across a second technology and two measurement platforms.

Sensitivity analyses

Table 5. Pre-specified sensitivity analyses.

#	Variant	Result
S1	Registry phase-3 association excluding Other/unclear platform class	chi-square = 17.32
S2	TAK-003 day-7 DE with stricter gene filter (CPM>1 in ≥10 samples)	16 up / 14 down (FDR<0.05; 17,837 genes tested)
S3	PfSPZ protection model with k=10 features instead of k=20	LOO AUC = 1.000

The registry association was robust to excluding the heterogeneous Other/unclear class; the TAK-003 day-7 gene-level findings were robust to a stricter gene filter; and the exploratory PfSPZ model retained perfect LOO separation with a halved feature set. Full numerical outputs for every analysis are shipped in 05_RESULTS (JSON + CSV), and all 12 figures are provided in both PNG (300 dpi) and editable SVG.

Discussion

This study set out to estimate the efficacy of mRNA vaccines in tropical diseases and instead established - with reproducible, quantitative evidence - that such an estimate is not currently possible from public data, while extracting the maximum scientifically defensible value from what does exist. The zero-hit GEO query and the three-trial registry census are small numbers with large consequences: they mean that any claim about mRNA vaccine efficacy against dengue, malaria, Zika or chikungunya currently rests on no public clinical data whatsoever. The one candidate platform entry closest to efficacy testing, mRNA-1893 against Zika, completed phase 1 with favourable safety and immunogenicity as published by its sponsors [10] but has no deposited subject-level dataset, and its development has not progressed to a public phase 3 record.

The transcriptomic analyses provide the biological counterpart to this regulatory observation. A licensed-context live-attenuated dengue vaccine (TAK-003) and the malaria subunit vaccine RTS,S both activate a coherent interferon-stimulated gene module - the same innate pathway that mRNA vaccines deliberately engage through their delivery vehicle [8,9]. The convergence is relevant for future mRNA-dengue or mRNA-malaria designs: the target population will already be mounting interferon responses to vaccination, and single-dose mRNA regimens must either exploit or overcome this pre-activated state. Our replication attempt of the MX2/GPR183 ratiometric signature in PfSPZ data moved in the published direction but was not significant with n=15, underscoring that cross-platform transfer of vaccine-response signatures requires the larger, controlled collections that currently do not exist for these diseases. The perfect leave-one-out separation of the exploratory 20-gene PfSPZ model should be read with its permutation p-value and Wilson intervals in hand: it demonstrates signal, not clinical readiness.

The registry analysis adds a developmental-systems perspective. Platform class was significantly associated with reaching phase 3, but the pattern is not simply that mRNA programmes are abandoned: viral-vector and DNA programmes also show zero phase-3 advancement in these four diseases, while live-attenuated, whole-pathogen and VLP platforms dominate advancement, largely reflecting decades of accumulated investment and field-site infrastructure. Enrolment size and start year were the strongest multivariable predictors, consistent with a development model in which large, late field efficacy trials - not platform novelty - are the binding constraint for tropical-disease vaccines.

Several limitations bound these findings. The PfSPZ protection analysis rests on 15 subjects (7 vs 8), so single-gene and model performance estimates carry wide Wilson intervals and remain hypothesis-generating only; with only three mRNA-vaccine trials in the registry, the platform contrast is descriptive by construction (Fisher exact two-sided p=1.0), and the Zika coefficient in the multivariable model was non-estimable owing to quasi-complete separation. The RTS,S pooled series encodes sample identity only in submitter column names and has no series-matrix file, and the NIAID Data Ecosystem API remained unreachable (HTTP 404) throughout the access window. Registry fields were near-complete but not perfect (enrolment missing 0.21%, start date 0.43%), and the platform taxonomy, although sensitivity-tested, remains keyword-based.

Conclusion

Using only genuinely public, downloaded and checksummed data, this project established that (i) no dataset permitting direct estimation of mRNA-vaccine efficacy against dengue, malaria, Zika or chikungunya exists as of 2026-11-07; (ii) the 469-trial registry landscape shows significant platform-phase structure in which the three mRNA vaccine trials remain early-phase; and (iii) existing tropical-disease vaccines across two further platforms converge on interferon-module activation, providing the immunological benchmark for future mRNA designs. All findings are exactly reproducible from the shipped pipeline, checksums and logs.

Data and code availability

Data availability. Every file used in this study was downloaded from the official NCBI GEO tree and the ClinicalTrials.gov API v2 during the study window (2026-11-07) and recorded with SHA-256 checksums in a download manifest; the registry pull is reproducible from the documented query pattern, and full numerical outputs for every analysis are shipped in 05_RESULTS (JSON + CSV). The NIAID Data Ecosystem API was unreachable (HTTP 404) throughout the access window, and no dataset used here depends on that source.

Code availability. Every number reported in this paper is regenerated by the accompanying pipeline from the raw files. Reproduction commands are shipped verbatim in 07_DOCUMENTATION/DATA_ACQUISITION.md, all 14 preprocessing operations are logged in 02_DATA/PROCESSED_DATA/preprocessing_log.csv, 29 validation checks passed with 0 failures, and all figures are provided in both PNG (300 dpi) and editable SVG. Analyses used Python 3.12.16 with pandas 2.2.3, numpy 2.1.3, scipy 1.14.1, statsmodels 0.14.5 and scikit-learn 1.5.2; the random seed was 42 throughout.

REFERENCES

[1] NCBI Gene Expression Omnibus (GEO). https://www.ncbi.nlm.nih.gov/geo/ [web resource]
[2] ClinicalTrials.gov API v2 (National Library of Medicine). https://clinicaltrials.gov/api/v2 [web resource]
[3] NIAID Data Ecosystem Discovery Portal. https://data.niaid.nih.gov/ [web resource]
[4] mygene.info gene annotation services. https://mygene.info/ [web resource]
[5] Kim EY, Che Y, Dean HJ, et al. Transcriptome-wide changes in gene expression, splicing, and lncRNAs in response to a live attenuated dengue virus vaccine. Cell Rep. 2022;38(6):110341. doi:10.1016/j.celrep.2022.110341. PMID: 35139383.
[6] Du Y, Thompson EG, Muller J, et al. The ratiometric transcript signature MX2/GPR183 is consistently associated with RTS,S-mediated protection against controlled human malaria infection. Front Immunol. 2020;11:669. doi:10.3389/fimmu.2020.00669. PMID: 32411130.
[7] NCBI GEO series GSE211626: Transcription profiling of PfSPZ Vaccine trial subjects (US, malaria naive) receiving a 3-dose or 5-dose regimen. https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE211626
[8] Pardi N, Hogan MJ, Porter FW, Weissman D. mRNA vaccines - a new era in vaccinology. Nat Rev Drug Discov. 2018;17(4):261-279.

[9] Richner JM, Himansu S, Dowd KA, et al. Vaccine-mediated protection against Zika virus-induced congenital disease. Cell. 2017;170(2):273-283.
[10] Essink B, et al. The safety and immunogenicity of two Zika virus mRNA vaccine candidates: a randomised, placebo-controlled, dose-ranging, multicentre, phase 1 trial (mRNA-1325 and mRNA-1893). Lancet Infect Dis. 2023. PMID: 36682364.
[11] August A, Shaw CA, Lee H, et al. A phase 1 trial of lipid-encapsulated mRNA encoding a monoclonal antibody with broad neutralizing activity against chikungunya virus. Nat Med. 2021. doi:10.1038/s41591-021-01573-6. PMID: 34887572.
[12] Hadinegoro SR, Arredondo-Garcia JL, Capeding MR, et al. Efficacy and long-term safety of a dengue vaccine in regions of endemic disease. N Engl J Med. 2015;373(13):1195-1206.
[13] Biswal S, Borja-Tabora C, Martinez Vargas L, et al. Efficacy of a tetravalent dengue vaccine in healthy children aged 4-16 years: a randomised, placebo-controlled, phase 3 trial. Lancet. 2020;395(10234):935-946.
[14] RTS,S Clinical Trials Partnership. Efficacy and safety of RTS,S/AS01 malaria vaccine with or without a booster dose in infants and children in Africa: final analysis of a phase 3, individually randomised, controlled trial. Lancet. 2015;386(9988):31-45.
[15] Datoo MS, Natama HM, Some A, et al. Efficacy of a low-dose candidate malaria vaccine, R21 in adjuvant Matrix-M, with seasonal administration to children in Burkina Faso: a randomised, controlled, phase 2b trial. Lancet. 2021;397(10279):1809-1818.
[16] Seder RA, Chang LJ, Enama ME, et al. Protection against malaria by intravenous immunization with a nonreplicating sporozoite vaccine. Science. 2013;341(6152):1359-1365.
[17] World Health Organization. Dengue vaccine: WHO position paper - September 2018. Wkly Epidemiol Rec. 2018;93(36):457-476.
[18] World Health Organization. WHO prequalifies a second dengue vaccine (Qdenga). 15 May 2024. https://www.who.int/news [web resource]
[19] U.S. FDA. FDA approves first chikungunya vaccine, Ixchiq. November 2023. https://www.fda.gov/news-events/press-announcements [web resource]
[20] Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc B. 1995;57(1):289-300.
[21] Schoggins JW, Rice CM. Interferon-stimulated genes and their antiviral effector functions. Curr Opin Virol. 2011;1(4):519-525.
[22] Edgar R, Domrachev M, Lash AE. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. Nucleic Acids Res. 2002;30(1):207-210.
[23] Barrett T, Wilhite SE, Ledoux P, et al. NCBI GEO: archive for functional genomics data sets - update. Nucleic Acids Res. 2013;41(D1):D991-D995.
[24] Xin J, Mark A, Afrasiabi C, et al. High-performance web services for querying gene and variant annotation. Genome Biol. 2016;17(1):91.
[25] Pedregosa F, Varoquaux G, Gramfort A, et al. Scikit-learn: machine learning in Python. J Mach Learn Res. 2011;12:2825-2830.
[26] Virtanen P, Gommers R, Oliphant TE, et al. SciPy 1.0: fundamental algorithms for scientific computing in Python. Nat Methods. 2020;17(3):261-272.
[27] Seabold S, Perktold J. statsmodels: econometric and statistical modeling with Python. Proc 9th Python in Science Conf. 2010:92-96.
[28] Hunter JD. Matplotlib: a 2D graphics environment. Comput Sci Eng. 2007;9(3):90-95.