Chow *et al.*
BMC Medical Research Methodology (2024) 24:170
https://doi.org/10.1186/s12874-024-02291-6

BMC Medical Research Methodology

RESEARCH

Open Access

Evaluating quasi-experimental approaches for estimating epidemiological efficacy of non-randomised field trials: applications in *Wolbachia* interventions for dengue

Jo Yi Chow[1*], Lin Geng[1], Somya Bansal[2], Borame Sue Lee Dickens[2], Lee Ching Ng[3], Ary Anthony Hoffmann[4] and Jue Tao Lim[1]

Abstract

Background *Wolbachia* symbiosis in *Aedes aegypti* is an emerging biocontrol measure against dengue. However, assessing its real-world efficacy is challenging due to the non-randomised, field-based nature of most intervention studies. This research re-evaluates the spatial–temporal impact of *Wolbachia* interventions on dengue incidence using a large battery of quasi-experimental methods and assesses each method's validity.

Methods A systematic search for *Wolbachia* intervention data was conducted via PUBMED. Efficacy was reassessed using commonly-used quasi-experimental approaches with extensive robustness checks, including geospatial placebo tests and a simulation study. Intervention efficacies across multiple study sites were computed using high-resolution aggregations to examine heterogeneities across sites and study periods. We further designed a stochastic simulation framework to assess the methods' ability to estimate intervention efficacies (IE).

Results *Wolbachia* interventions in Singapore, Malaysia, and Brazil significantly decreased dengue incidence, with reductions ranging from 48.17% to 69.19%. IEs varied with location and duration. Malaysia showed increasing efficacy over time, while Brazil exhibited initial success with subsequent decline, hinting at operational challenges. Singapore's strategy was highly effective despite partial saturation. Simulations identified Synthetic Control Methods (SCM) and its variant, count Synthetic Control Method (cSCM), as superior in precision, with the smallest percentage errors in efficacy estimation. These methods also demonstrated robustness in placebo tests.

Conclusions *Wolbachia* interventions exhibit consistent protective effects against dengue. SCM and cSCM provided the most precise and robust estimates of IEs, validated across simulated and real-world settings.

Keywords *Wolbachia*, Dengue, Non-randomised, Quasi-experiment, Synthetic control

*Correspondence:
Jo Yi Chow
joyi001@e.ntu.edu.sg
Full list of author information is available at the end of the article

Introduction

The burden of dengue increased globally by 30-fold over the past five decades [1], placing a heavy burden on healthcare systems. With limited effective therapeutics [2, 3] and vaccines [4, 5], vector-control remains the primary means to reduce dengue transmission [6, 7].

Emerging vector-control strategies, such as the Sterile Insect Technique (SIT) and *Wolbachia*-based methods, show promise in reducing mosquito populations and the transmission of dengue. SIT utilises radiation to sterilise male mosquitoes, limiting population growth. *Wolbachia*-based techniques employ two methods: the Incompatible Insect Technique (IIT), where *Wolbachia*-infected males render females infertile, and introgression (also referred to as "replacement"), which spreads *Wolbachia* within mosquito populations through infected females, reducing their vectorial capacity.

Despite the initial success of *Wolbachia*-infected Aedes aegypti mosquitoes in field trials across several countries like Malaysia [8, 9], Indonesia [10], Singapore [11, 12], Brazil [13, 14], Colombia [15], and Australia [16, 17], accurately gauging the intervention's epidemiological impact presents substantial challenges. These challenges stem from factors that can obscure causal inferences regarding intervention efficacy (IE), such as (1) intervention spillovers to control areas, (2) heterogeneities in baseline disease trajectories in spatial units prior to intervention, (3) staggered adoption of intervention due to resource limitations, (4) environmental heterogeneity, leading to different doses of interventions across target units despite uniform implementation of interventions, and (5) biological intricacies of the intervention, such as imperfect suppression of wildtype mosquitoes or incomplete introgression.

While gold-standard, cluster randomised controlled trials (cRCTs) have successfully evaluated the epidemiological efficacy of *Wolbachia* introgression [10], they face practical and ethical constraints in population-level infectious disease control [18–20]. Additionally, although cRCTs balance baseline characteristics and historical dengue risk in intervention and control arms, differences in disease pre-trends cannot be explicitly balanced even if randomisation equalises historical dengue risk. Spillover effects due to the close proximity of intervention sites to controls also cannot be easily mitigated even with geographical features, underestimating the epidemiological efficacy of interventions.

Given these constraints, cRCTs can be augmented with alternative, quasi-experimental methods to triangulate IEs, including Pre-post analysis [16], Bayesian time series analysis [8], Difference-in-Differences (DiD) [13, 14, 21] and Regression Discontinuity Design (RDD) [22]. Currently, there is no consensus on a universally superior method, and choices often depend on study context. Other methodologies, such as compartmental models [17] can account for complicated disease transmission processes, and the Synthetic Control Method (SCM) [12] has been employed to simultaneously accommodate differences in spatiotemporal characteristics and historical dengue risk. These approaches aim to counteract selection bias and confounding factors, helping in robust causal inference in the absence of traditional randomisation.

While existing literature acknowledges the challenges of evaluating *Wolbachia* interventions, there is yet to be an examination of the robustness of quasi-experimental methods in this application. To address this gap, our study conducted a thorough analysis of various quasi-experimental designs to determine the causal impact of *Wolbachia* on dengue incidence across diverse settings in Singapore, Malaysia, and Brazil. Our study provides a threefold contribution to the literature on quasi-experimental methodology. Firstly, we systematically searched for and re-evaluated all available open datasets pertaining to *Wolbachia* field trials using a large battery of quasi-experimental tools to re-estimate IEs by spatial units and event time. Secondly, we implemented a series of spatial and temporal placebo tests to verify the methodological robustness of each method. Finally, we stress-tested each method against a diverse range of hypothetical dengue epidemic scenarios through a simulation study to further underscore the practical relevance of these tools in public health planning.

By comparing these methods, we aim to identify the quasi-experimental techniques that have the necessary rigor for reliable evaluations of public health measures, a critical step to aid in the implementation of evidence-based strategies against dengue and similar diseases when randomisation is impractical.

Methods

Study cohorts and data collection

We sought to re-evaluate the efficacies of *Wolbachia* interventions to stem dengue transmission where reports were publicly available, using quasi-experimental tools. First, eligible study cohorts were identified through a systematic review conducted in accordance with the PRISMA guidelines (Supplementary Fig. 1). A search was performed on the PUBMED platform from December 2023 to January 2024 using specific terms: (1) "dengue" OR "breakbone fever," (2) "population" OR "community," and (3) "*Wolbachia*". The final search was conducted on January 4, 2024.

The inclusion criteria for article selection were studies that investigated the epidemiological impact of *Wolbachia* on dengue. On the other hand, the following exclusion

criteria were applied: (1) articles without open-sourced data, (2) articles that only reported aggregated dengue incidence across all control and intervention sites, (3) articles with no control arm or only a single control arm, (4) articles where the final reported metric was not dengue incidence, (5) review articles, and (6) articles not published in English. These criteria ensured that the included studies were accessible for re-evaluating IEs using selected quasi-experimental methods as described later.

The initial search yielded a total of 212 articles. Following screening and application of the specified inclusion and exclusion criteria, 2 articles met the requirements and were selected for data extraction. The third study setting, Singapore, was included in the analysis despite not having open-sourced data due to the team's pre-existing access to the relevant data from previous studies [12, 23]. For more detailed data descriptions, see Supplementary Methods.

Quasi-experimental methods for understanding intervention effects

We evaluate IEs of *Wolbachia* interventions to reduce dengue incidence rates using 6 commonly employed quasi-experimental methods, which include the (1) Pre-post design – where dengue incidence rates are compared based on mean differences by utilising a simple regression model; (2) regression discontinuity design (RDD) – where dengue incidence rates are compared in pre- and post-intervention time for intervention units while controlling for temporal trends; (3) 2×2 difference-in-differences (2×2 DiD), which computes the intervention efficacies by comparing first, the Pre-post intervention differences in dengue incidence for intervention units, and the differences in dengue incidence between intervention and control units over time; (4) difference-in-differences with multiple time periods (Panel DiD), which is an extension of 2×2 DiD that explicitly accounts for the intervention adoption time in each treated unit, allowing for estimation of treatment effects that vary spatially and temporally (Supplementary Table S2, S3); (5) the canonical synthetic control methods (SCM), which generates counterfactuals for the intervention site by minimizing the differences in the dengue incidence between a weighted set of donors (controls) versus the intervention unit in the pre-intervention period; and (6) count SCM (cSCM), which relaxes the bounding of synthetic control weights but still maintains the zero-lower bound for dengue incidence for the synthetic controls. (1) – (2) can be considered a single-group design as it includes only treated units in the model for evaluation, while (3) – (6) can be considered multiple-group designs which have both treated and control groups for evaluation of IEs.

For regression-based analyses (1) – (3) with the Malaysian and Singapore data, we utilised zero-inflated negative binomial models from the 'pscl' R package, as the data exhibited non-negative, continuous characteristics with a significant number of zero values. Meanwhile, for the Brazil data, we employed truncated normal regressions using the R package 'truncreg'. This decision was based on the absence of a population to use as an offset and the continuous, non-negative nature of the data. Supplementary Methods provides specific details on implementation of all quasi-experimental methods, computation of IEs and the counterfactuals.

Aggregates of intervention efficacy (IE)

As outcomes were both temporally and spatially explicit, we could re-aggregate IEs in meaningful ways, to examine potential spatial heterogeneity in IEs and considering time needed for *Wolbachia* interventions to take effect – either the time required for *Wolbachia* to introgress into local populations for the introgression approach, or suppression of wild-type mosquitoes to take place using the incompatible insect technique (IIT). In the preceding sections, $IE_{RDD,i}$, $IE_{DiD,i}$ and $IE_{SCM,i}$ can be considered as aggregations of IEs for a specific treated unit i. However, we can also compute.

(1) the overall effect of the intervention across all study sites (IE_I), as the difference between the cumulative sum of actual cases $y_{i,t}$ and cumulative sum of counterfactual cases $\widehat{y}_{i,t}$ across all sites. This is calculated from each site i's specific start time of intervention $t_{start(i)}$ to the end of the study t_{end}.

$$IE_I = \frac{\left(\sum_{i=1}^{I}\sum_{t=t_{start(i)}}^{t_{end}}\widehat{y}_{i,t}\right) - \left(\sum_{i=1}^{I}\sum_{t=t_{start(i)}}^{t_{end}} y_{i,t}\right)}{\left(\sum_{i=1}^{I}\sum_{t=t_{start(i)}}^{t_{end}}\widehat{y}_{i,t}\right)} \times 100 \quad (1)$$

(2) the IE by event-study time ($IE_{t_1:t_2}$), where $t_1 : t_2$ are the start and end periods for some pre-specified time in the post-intervention period. Intervention and counterfactuals can similarly be obtained in preceding sections.

$$IE_{t_1:t_2} = \left(\sum_{t=t_1:t_2} \frac{\text{counterfactual}_t - \text{intervention}_t}{\text{counterfactual}_t}\right) \times 100 \quad (2)$$

Robustness checks for quasi-experimental methods

Each method was stress-tested to validate efficacy estimates. Placebo interventions, set two years prior to actual interventions, discerned if observed effects were due to

pre-intervention dengue incidence trends or the *Wolbachia*, ensuring that each methods fits well to the pre-intervention data and that the measured effects were attributable to the intervention rather than to poorly predictive counterfactuals. For each dataset and method, we truncated all observations to be the pre-intervention period data, and recalculated placebo IEs after adjusting the placebo-treatment to two years prior to the actual intervention date. Post-adjustment placebo efficacies were then compared to actual intervention effects in the complete dataset.

To test for spatial validity, in-space placebo tests were conducted by reassigning the control group to be the treated units and measuring whether there were any large intervention effects for control sites, while excluding the actual treated units from the donor pool. Supplementary Methods provides further details on the execution of in-space placebos for each method.

Simulation study

We conducted a simulation study to examine whether quasi-experimental methods can recover IEs of *Wolbachia* interventions. We defined data generating processes using a custom-built ASEI-SEIR transmission dynamic model which incorporates host-vector components with waning immunity with stochasticity in the observation process for the *Wolbachia* introgression intervention (Supplementary Methods, Supplementary Table S1). In summary, the simulation was conducted 1000 times for various control pool sizes, ranging from 5 to 40 sites with increments of 5 per 1000 simulations.

Each simulation run created two intervention sites first with the absence, then the presence of *Wolbachia* releases using the same initial parameters (i.e. initial populations of susceptible humans S_h, dengue-infected humans I_h, and wild-type juvenile mosquitoes $A_{(n,w)}$). After the simulations were run, we assumed stochasticity using a Poisson distribution for the number of infected individuals in both intervention and control sites. The 'true' IE was then computed by:

$$IE_{t_1:t_2} = \left(\sum_{t=t_1:t_2} \frac{\text{no intervention}_t - \text{intervention}_t}{\text{no intervention}_t} \right) \times 100 \quad (3)$$

Each quasi-experimental tool was then fitted according to the observable number of infected individuals to generate counterfactuals for the case where the intervened group had no interventions. The *IE* of *Wolbachia* introgression was hence calculated using each method and its respective formulae. The percentage error between the 'true' *IE* and the calculated values was subsequently computed to determine the ability of each method to recover the 'true' *IE*.

Data analysis was performed using R software (V. 4.3.1) and RStudio (V.2023.06.1 + 524).

Results

Characteristics and outcomes of *Wolbachia* trials across study settings

Table 1 provides a summary of the essential characteristics pertaining to the *Wolbachia* trials conducted in Malaysia, Brazil, and Singapore. The Malaysia trial involved *w*AlbB *Wolbachia* introgression across six sites, with a pre-intervention period of 221 to 255 weeks and a post-intervention period of 78 to 88 weeks, achieving the highest intervention saturation among the studies at 74.62%. In Malaysia, saturation is calculated as the average percentage of wAlbB detected by qPCR in Ae. aegypti mosquitoes, starting 4 or 8 weeks post-release depending on the study site. In Brazil, a similar intervention with *w*Mel *Wolbachia* spanned 32 sites, with a pre-intervention duration of 42 to 51 months and a shorter post-intervention timeframe of 3 to 12 months, resulting in a 50.65% saturation rate. In Brazil, saturation is the average percentage of wMel detected by qPCR, starting 1 month post-release. Conversely, Singapore utilised a combination of IIT and SIT with high-fidelity sex sorting across four sites with saturation gradually expanding over the study period with the largest control group of 30 sites. Saturation in Singapore is calculated as the sum of (intervention area * intervention timepoints) / (total area * max timepoints). The study in Singapore featured the most extensive pre-intervention data with 247 to 335 weeks of observation time, and post-intervention data of 109 to 209 weeks.

Aggregated intervention efficacies by site and event time

Reductions in dengue incidence were attributed to *Wolbachia* interventions in all study settings, indicated by positive IEs for methods which passed all robustness checks. For Malaysia, the IE estimates from quasi-experimental methods ranged from 46.73% to 69.19% (Table 2), surpassing the 40.00% (95%CI 5.06–64.59) estimate from the original study [8]. Brazil estimates showed a high degree of efficacy (55.35% to 89.40%), broadly consistent with the original study [14]. However, only SCM and cSCM passed spatial robustness checks in this location. In Singapore, SCM and cSCM passed both placebo checks and these methods respectively produced positive IEs despite incomplete saturation of interventions in intervention sites over the study period, similar to the original study [12]. Estimated IEs across study settings revealed distinct patterns when both in-time and in-space checks were employed.

The efficacy of *Wolbachia* interventions can be mediated by their duration. As quasi-experimental tools can

Table 1 Summary of Wolbachia intervention approaches over the 3 study settings (Malaysia, Brazil, Singapore)

Study Setting	Selangor, Malaysia	Niterói, Brazil	Singapore, Singapore
Intervention type	Introgression	Introgression	IIT-SIT, High-fidelity sex sorting
Study dates	EW 1 2013 – EW 19 2019	EW 1 2007 – EW 27 2020	EW 1 2014 – EW 26 2022
Intervention sites	6	32	4
Control sites	19	19	30
Pre-intervention period[a]	221–255 weeks	42–51 months	247–335 weeks
Post-intervention period[a]	78–88 weeks	3–12 months	109–209 weeks
Intervention time[b]	94.5 weeks	9.4 months	156.75 weeks
At risk population[c]	179,305	373,000	5,151,316
Average intervention saturation over study duration[d]	74.62%	50.65%	40.28%

[a] given as the range of pre and post intervention lengths across all units in the specific study setting

[b] calculated as average across all intervention sites

[c] sum of population across all intervention and control sites

[d] Computation varies across locations. In Singapore, saturation is determined by the sum of (the area of intervention*total timepoints of intervention)/(total intervention area*max timepoints of intervention). In Malaysia and Brazil,, average saturation is determined by the average percentage of the wAlbB (Malaysia) or wMel (Brazil) Wolbachia strain detected by qPCR in Ae. aegypti mosquitoes across all sites, from the initiation of monitoring. Monitoring commences 4 or 8 weeks post-initial release in Malaysia (varying by site) and 1 month post-release in Brazil

Table 2 Aggregate intervention efficacy (IE) estimates (%) of Wolbachia releases on total dengue incidence rates across all intervention sites. The estimates are reported for each method and are accompanied by the corresponding confidence intervals (in parentheses). The IE estimate reported in the respective original papers, highlighted in blue, is provided for comparison. Numbers in parenthesis represent lower and upper bounds for 95% confidence intervals, estimated using the bootstrapping procedure [8, 12, 14]

	Aggregate site IE (%)		
	Selangor	Niterói	Singapore
Pre-post	**58.43 (51.96-63.70)[ab]**	89.40 (79.10-93.50)	-7.27 (-23.03-6.05)[b]
RDD	**69.19 (37.25-85.90)[ab]**	60.76 (88.90-129.02)	10.08 (-98.34-64.57)[b]
2x2 DiD	**46.73 (43.14-49.78)[ab]**	87.41 (69.15-92.63)	42.86 (39.40-46.12)[b]
SCM	**64.26 (62.92-65.81)[ab]**	66.59 (60.18-70.83)[b]	**53.33 (51.66-54.89)[ab]**
cSCM	**66.38 (65.25-67.51)[ab]**	55.35 (52.16-65.95)[b]	**48.71 (47.67-53.18)[ab]**
Estimate from paper	40.00 (5.06-64.59)	69.40 (54.40-79.40)	56.88 (51.88-58.46)
Average saturation (%)over study period[c]	74.62%	50.65%	40.28%

Pre-Post analysis compares the dengue incidence rates before and after the intervention. RDD analyses the relationship between the intervention time (assignment variable) and dengue incidence (outcome) at a specific threshold (start of the intervention). The 2 × 2 Difference in Differences (DiD) approach compared changes over time between treatment and control groups. SCM and cSCM methods construct a counterfactual by using a weighted combination of control units, with the latter relaxing the convex hull assumption

[a] indicates that the estimate passed in-time placebo checks while [b] indicates that it passed in-space placebo check. Bolded figures represent significant IE estimates which also passed in-space and in-time placebo checks

[c] Computation varies across locations. In Singapore, the saturation is determined by averaging the product of the area covered by the intervention within each specific site and the number of effective weeks, divided by the total area of the Wolbachia intervention in the site multiplied by the number of weeks in a year. In Malaysia and Brazil, average saturation is determined by the average percentage of the wAlbB or wMel Wolbachia strain detected by qPCR in Ae. aegypti mosquitoes across all sites, from the initiation of monitoring. Monitoring commences 4 or 8 weeks post-initial release in Malaysia (varying by site) and 1 month post-release in Brazil

generate counterfactuals for each location and time post-intervention, we re-aggregated IEs by event time relative to intervention start (Table 3). Results indicated that *Wolbachia* releases were associated with reduced dengue incidence in all locations, but geographic variation in magnitude of efficacy was evident, likely due to local factors such as saturation levels, release strategies and *Wolbachia* strain employed (Table 3).

Malaysia's IE estimates were consistently high across methods (Table 3, Supplementary Figure S2a), demonstrating an upward trend over time, suggesting that the intervention is stable and increasingly effective as *Wolbachia* introgresses into the local *Aedes* population. Conversely, Brazil showed significant variability and a notable decline in efficacy in later months (Table 3, Supplementary Figure S2b). However, Brazil's IE estimates lack robustness, failing all temporal checks and passing few spatial checks. In Singapore, IIT-SIT showed reductions in dengue incidence ranging from 35.77–53.11%, using SCM and cSCM (Table 3, Supplementary Figure S2c), stabilising as intervention saturation increases.

This variation in IE reflects the complexity of factors affecting *Wolbachia* intervention outcomes, highlighting the need to account for local conditions in efficacy assessments.

Empirical validation by simulation study

Table 4 presents the results of the simulation study, displaying the average percentage errors across 1000 simulation runs along with their corresponding standard deviations (s.d.) across methods. Both the Pre-post and RDD methods exhibited underestimation of the true IE. However, the RDD method, with a percentage error of -34.98% (s.d. 23.848), demonstrated the greatest extent of underestimation and spread when compared to other methods.

2×2 DiD consistently underestimated the true IE across all control pool sizes, with a percentage error of approximately -2.5% (s.d. 3.39), suggesting a minor bias that remained consistent irrespective of the control pool size.

Meanwhile, SCM and cSCM showed notable improvements in performance as the control pool size increased. SCM's error decreased from 2.52% (s.d. 3.327) to 0.23% (s.d. 1.370), and its s.d. decreased as the number of controls increased, indicating greater accuracy and consistency with larger control pools. On the other hand, cSCM exhibited a transition from a small positive error of 1.62 (s.d. 2.465) to a small negative error of -0.27 (s.d. 0.572) as the control pool expanded, along with diminishing s.d., suggesting enhanced precision, but a tendency to slightly underestimate the IE with increasing control pool size. cSCM also exhibited the least variability in its estimates across varying control pool sizes, reflected by consistently narrow standard deviations, indicating a high level of precision in estimates. The results also suggest a requisite minimum number of control groups for accurate estimation in SCM and cSCM. With a baseline of 10 control groups, SCM's accuracy improved by a factor of four (2.52 to 0.60). Likewise, cSCM's accuracy saw a tenfold increase (1.62 to -0.15) with at least 5 control groups, but declined when more than 20 groups were used.

Discussion

Dengue necessitates robust vector control strategies. Yet, assessing the effectiveness of these interventions in field conditions is difficult, due to the challenge of undertaking randomised control studies and confounding in non-randomised studies. This study revisits the IEs of three *Wolbachia* field trials targeting dengue from a systematic search through a variety of commonly-used quasi-experimental methods, all underpinned by stringent validity checks. Despite incomplete intervention saturation over the study period, methods passing these checks confirm *Wolbachia*'s protective effect against dengue (Tables 2 and 3), aligning with existing literature [8–10, 12–17, 21].

Comparative analysis reveals how different statistical methods which complied with spatio-temporal placebo checks can yield diverse IE estimates. The Bayesian model from the Malaysia study yielded a conservative IE estimate of 40.00% (95%CI 5.06–64.59) [8] while other methods in this study reported higher efficacies 46.73 – 69.19%. However, a revised Bayesian analysis based on a longer timespan and additional study sites were concordant with our higher estimates of intervention efficacy (IE: 62.4% (95%CI: 50.00–71.00)) [9]. In Singapore, SCM and cSCM derived IEs (IE: 48.71–53.33%) were in close agreement with the original study's 56.88% (95%CI 51.88–58.46) [12], with the application of similar methodologies and data, confirming the intervention's effectiveness even with partial *Wolbachia* saturation. These consistent results, validated by robust placebo checks, highlight how the choice of analytical method can significantly impact intervention assessments.

The Brazil study's application of Interrupted Time Series Analysis (ITSA), operationally mirroring 2×2 DiD, resulted in a higher reported IE of 69.40 (95%CI: 54.40–79.40) [14] compared to 66.59% (95%CI: 60.18–70.83) and 55.35% (95%CI 52.16–65.95) from our use of SCM and cSCM respectively. Discrepancies in IE estimates may arise from different methodological approaches. Firstly, SCM and cSCM have theoretically enhanced accuracy over DiD through better balancing of pre-intervention historical outcomes compared to DiD, thus accounting for unobserved confounders.

Table 3 Intervention efficacy (IE) estimates of Wolbachia releases on total dengue incidence rates across all intervention sites aggregated by event time, with corresponding intervention saturation. For Malaysia and Brazil, saturation is determined by the average percentage of the wAlbB/wMel Wolbachia detected by qPCR in Ae. aegypti mosquitoes across all sites during monitoring phase relative to the start of intervention (introgression). In Singapore, saturation is defined as the area covered by the intervention (IIT-SIT)

			Method				
Approach	**Duration (months)**	**Saturation (%)**	**Pre-post**	**RDD**	**2×2 DiD**	**SCM**	**cSCM**
			Intervention Efficacy (IE)				
Introgression			**Selangor, Malaysia**				
	1 – 6	72.16	**65.18 (59.74–69.62)**	**58.65 (40.38–68.64)**	**53.86 (50.82–56.44)**	**44.36 (43.00–45.81)**	**51.47 (50.31–52.44)**
	7 – 12	76.97	**49.57 (41.69–55.99)**	**52.69 (3.51–71.95)**	**34.13 (29.54–38.00)**	**65.45 (63.98–67.23)**	**67.36 (66.22–68.82)**
	13 – 18	66.99	**54.60 (47.54–60.35)**	**72.37 (33.42–89.51)**	**42.85 (38.90–46.18)**	**74.43 (73.32–75.73)**	**74.84 (73.81–75.88)**
	19 – 24	80.84	**63.42 (57.98–67.85)**	**89.24 (71.83–96.31)**	**56.61 (53.78–59.01)**	**72.14 (70.45–73.59)**	**73.17 (71.82–74.04)**
Introgression			**Niterói, Brazil**				
	1 – 6	45.32	83.90 (68.57–90.08)	64.93 (-47.56–82.65)	83.88 (68.87–89.82)	74.59 (68.90–77.71)	63.71 (60.03–73.97)
	7 – 12	40.66	92.27 (84.62–95.27)	75.36 (90.26–176.78)	92.26 (84.73–95.16)	63.48 (59.36–68.36)	58.45 (54.25–64.44)
	13 – 18	56.74	95.81 (91.67–97.44)	79.28 (95.11–109.37)	95.81 (91.73–97.38)	26.29 (17.65–45.28)	14.55 (3.26–47.43)
	19 – 24	44.38	85.90 (71.96–91.38)	-64.44 (86.08–116.35)	77.03 (-46.31–88.50)	20.97 (0.30–40.79)	-6.98 (-17.46–35.67)
IIT-SIT			**Singapore**				
	1 – 6	18.73	26.97 (15.91–36.24)	-23.31 (-79.14–14.53)	64.34 (62.12–66.42)	**53.82 (51.45–55.78)**	**48.18 (46.01–53.79)**
	7 – 12	21.93	-22.78 (-41.35–7.21)	-4.19 (-78.67–42.33)	48.08 (44.73–51.20)	**35.77 (33.41–39.48)**	**35.28 (33.89–40.02)**
	13 – 18	27.75	34.57 (24.67–42.86)	70.53 (38.45–87.19)	72.33 (70.54–74.00)	**53.09 (50.89–54.35)**	**53.11 (51.37–55.18)**
	19 – 24	53.77	-125.96 (-160.13–97.30)	52.58 (-29.71–84.37)	4.45 (-1.72–10.20)	**51.61 (50.21–53.83)**	**45.67 (44.65–52.93)**

Numbers in parenthesis represent 95% confidence intervals estimated by bootstrapping. Bolded figures represent significant IE where 95% confidence intervals do not contain 0, and were estimated using methods which also passed in-space and in-time placebo checks

Additionally, due to absence of population data at the neighbourhood level, the granularity of data differs; our study's neighbourhood-level analysis with min–max normalisation to adjust for population sizes and reporting disparities contrasts with the original study's zone-level approach, in which the authors could directly account for population. Therefore, interpreting IE differences between studies requires concomitant consideration of each analytical approach's unique assumptions and constraints.

Aggregating IEs by the period post-intervention (Table 3, Supplementary Figure S2), our analysis facilitates equitable comparison of intervention impacts across sites, currently unexamined in existing literature, illuminating the dose–response relationship of *Wolbachia* interventions. The IEs in Malaysia displayed an upward trend with longer intervention periods, indicating that sustained *Wolbachia* presence may lead to more significant reductions in dengue incidence. This observation indicates the potential long-term benefits of sustained *Wolbachia* releases until *Wolbachia* frequencies reach a high and stable level without further releases, and is also consistent with the recent Bayesian analysis completed on additional Malaysian data [9]. In Singapore, IIT-SIT first showed a moderate efficacy, which may be due to wildtype vector declines as *Wolbachia*-infected males began reducing viable mosquito offspring. The dip in mid-term efficacy likely reflected interepidemic dengue transmission or ecological adjustments to the new intervention–emphasising the need for persistent releases to maintain pressure on the mosquito population. The recovery and eventual stabilisation of IE in the later months signifies that the consistent release of *Wolbachia*-infected males leads to a progressive and sustained reduction in the mosquito population despite incomplete saturation across the study period. In Brazil, a decline in IE over time pointed to challenges in maintaining the intervention's effectiveness. Varied factors affect the *wMel* strain's homogenous spread within mosquito populations, including persistent wild-type eggs, urban infrastructure barriers, and disruptions like the COVID-19 pandemic impeding mosquito dispersal [13, 14]. The Brazil dataset, inclusive of both suspected and confirmed dengue cases, confronts limitations in the specificity and sensitivity of case reporting, wherein potential misclassified or unreported cases could skew the assessment of *Wolbachia*'s efficacy. Declines in maternal transmission of *Wolbachia* at higher temperatures may also account for decreased efficacy, especially in Brazil's dense urban environment with its variable microclimates [24, 25]. The need for adaptive management of *Wolbachia* interventions, as emphasised in public health discussions [26], is critical to respond to these environmental and biological challenges to maintain the success of the program.

Our study introduces a novel empirical validation of quasi-experimental methods using the ASEI-SEIR model to simulate dengue dynamics with *Wolbachia* interventions, addressing the previously unexplored efficacy of such methods in recovering IEs in vector-borne disease transmission and interventions. Consistently minimal errors affirm SCM and cSCM's exceptional empirical precision and accuracy in recovering IEs, outperforming other methods and demonstrating their reliability across different control configurations (Table 4). Conversely, RDD showed a notable negative bias, and the Pre-post and 2×2 DiD methods, despite relatively modest errors, did not match SCM and cSCM's performance. Although parallel trends were validated in the Brazil study's DiD analysis, our simulations point out that 2×2 DiD's sensitivity to unaccounted factors can lead to erroneous estimations of IE, even with historically consistent parallel trends. This illustrates the advantage of SCM and cSCM in adjusting for time-varying confounding factors, offering a reduction in bias within quasi-experimental settings, given an adequate pre-intervention duration [27].

In public health research, the choice between cRCTs and quasi-experimental designs hinges on the study's resources, objectives, and constraints. cRCTs may offer high validity but face logistical and ethical issues in outbreak situations. Our study illustrates that quasi-experimental methods can provide comparable validity with thorough checks, yet require cautious interpretation due to potential biases from non-randomisation. In the simple one-group Pre-post design, researchers must adjust for confounding factors, including disease incidence fluctuations (regression towards the mean) and behavioural changes (maturation), to attribute effects to the intervention, given that this method cannot infer causality on its own due to the lack of a control group [28]. For SCM and cSCM to produce reliable counterfactuals, certain conditions should be met. While pre-intervention data directly before the intervention reflects the most recent trends, a longer pre-intervention period better captures historical changes and strengthens counterfactuals [27, 29]. While a larger donor pool offers more control options for weight selection, it is essential to select well-matched controls to minimise over-fitting and interpolation biases [30]. A well-matched control unit is characterised by its ability to closely mimics the outcome trajectory of the treated unit over an extended pre-intervention period [30]. Valid causal inference with 2×2 DiD requires confirmed parallel trends [31]. RDD's effectiveness is bandwidth-dependent, requiring the balance between breadth of data against the risk of bias [32, 33]. Lastly, traditional quasi-experimental methods

Table 4 Percentage errors of intervention efficacies for each method compared to the true intervention efficacy (IE). The values in parentheses represent the standard deviations, reflecting the variability observed across 1,000 bootstrapping iterations. These errors were computed using the formula (true IE–- IE from method) / (true IE). Here, 'true IE' represents the IE determined by simulating the dengue epidemic cycle with the same initial parameters and comparing scenarios with and without the intervention. For the intervention scenario, the initial Wolbachia-infected mosquito population is set to a random multiplier (1.2 to 2 times) of the wild-type mosquito population. In the non-intervention scenario, the initial Wolbachia mosquito population is set to zero

Control pool size	Intervention efficacy (IE) percentage error (%)				
	Pre-post	RDD	2×2 DiD	SCM	cSCM
5	-1.85 (3.36)	-34.98 (23.848)	-2.63 (3.393)	2.52 (3.327)	1.62 (2.465)
10			-2.53 (3.389)	1.24 (2.116)	0.37 (1.250)
15			-2.63 (3.388)	0.60 (1.571)	-0.15 (0.825)
20			-2.59 (3.388)	0.52 (1.530)	-0.18 (0.736)
25			-2.58 (3.389)	0.36 (1.474)	-0.22 (0.679)
30			-2.53 (3.388)	0.40 (1.497)	-0.21 (0.651)
35			-2.52 (3.386)	0.30 (1.402)	-0.26 (0.595)
40			-2.52 (3.386)	0.23 (1.370)	-0.27 (0.572)

The analysis did not include Pre-post and RDD because the efficacy estimates derived from these methods are not affected by the size of the control group. This independence arises because neither method relies on a control or comparison group for efficacy calculations. Specifically, RDD estimates IE by comparing the incidence of dengue immediately before and after the interventions' initiation at a specific threshold, while Pre-post assesses efficacy through within-unit comparisons of outcomes before and after the intervention

approximate treatment effects as binary step functions, lacking the ability to account for the increasing magnitude of treatment effects over time. This approach fails to capture the gradual and cumulative impacts of interventions like Wolbachia, where small levels of saturation can significantly impact over time. Our time-aggregated analysis (Table 3) demonstrates how variations in treatment saturation correlate with changes in intervention efficacy providing a more accurate understanding of the intervention's incremental effects. These findings highlight that when applied rigorously, quasi-experimental methods extend the toolkit for reliable causal inferences in public health interventions where cRCTs are impractical.

Our study has several strengths. (1) We systematically reviewed publicly available data on *Wolbachia* field trials to minimise selection bias of included studies for re-evaluation of IEs. (2) We employed a comprehensive suite of quasi-experimental methods, accompanied by extensive robustness checks, to confidently demonstrate *Wolbachia*'s consistent protective effect against dengue. (3) We used a novel data generating process using a custom-built transmission dynamic model to validate the efficacy of these quasi-experimental methods in estimating true IEs. However, there are limitations. The study did not account for spillover effects, possibly underestimating Wolbachia's efficacy by diluting the observed differential in disease incidence between intervention and control sites. Subsequent research could use spatial analysis to account for indirect protection effects due to the migration of *Wolbachia*-infected mosquitoes into control areas. Moreover, data constraints precluded the inclusion of relevant covariates and demographic subgroup analyses. Future studies with richer datasets could perform more granular analyses using spatial interpolation to better understand *Wolbachia*'s differential impact across population strata and control for confounders. Lastly, our study focused on a single-serotype, single-season simulation to clarify IE recovery in simplified dengue dynamics. Multi-season, multi-serotype expansions are recommended for future work to triangulate long-term efficacy while encompassing the interplay of cross-immunity to antibody-dependent enhancement.

Conclusion

In conclusion, our study rigorously re-evaluated the impact of *Wolbachia*-based vector control interventions on dengue incidence across three locations using quasi-experimental methods. The findings revealed a consistent protective effect against dengue, particularly in analyses supported by robust internal validity checks. SCM and cSCM stand out for their precision and validity in estimating IEs in both real-world and simulated settings, displaying promise for guiding public health strategies where randomised controlled trials are not always feasible.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12874-024-02291-6.

Supplementary Material 1.

Supplementary Material 2. Supplementary Figure S2. Aggregated intervention efficacies for (a) Malaysia, (b) Brazil, and (c) Singapore based on event time. Each efficacy assessment method is represented by a specific color: red for Pre-post, yellow for RDD (Regression Discontinuity Design), green for 2x2 DiD (Difference-in-Differences), blue for SCM (Synthetic Control Method), and purple for cSCM (count Synthetic Control Method). The intervention efficacy (IE) reported in the original paper is depicted on the right side of each diagram, represented by the colour black. The dot represents the IE point estimate for each method, while the line indicates the corresponding confidence intervals. Extreme confidence interval values and point estimates are denoted by arrows, with their respective values displayed next to the arrow.

Acknowledgements
Not applicable.

Authors' contributions
J.T.L. and B.L.D. conceived the study. J.Y.C. and G.L. developed the underlying code for the study's analysis. J.Y.C., J.T.L., S.B., and G.L. implemented the study. A.A.H. and L.C.N. provided data resources. J.Y.C., J.T.L., and S.B. analysed the data. J.Y.C. and J.T.L. drafted the manuscript with contributions in the interpretation of the findings from A.A.H. and L.C.N. All authors have read and reviewed the manuscript.

Funding
This research is hosted by CNRS@CREATE and supported by the National Research Foundation, Prime Minister's Office, Singapore, under its Campus for Research Excellence and Technological Enterprise (CREATE) program, and is funded by the Lee Kong Chian School of Medicine—Ministry of Education Start-Up Grant.

Availability of data and materials
The data underlying this article is available in Mendeley Data at https://doi.org/https://doi.org/10.17632/v8vn35zj3g.1 for Malaysia and Figshare at https://doi.org/https://doi.org/10.6084/m9.figshare.13662203.v3 for Brazil. Data for Singapore are the property of the Ministry of Health, Singapore, and were shared under the Infectious Disease Act. Permission to access Singapore data should be obtained from the Ministry of Health, Singapore. The code required to reproduce this study's findings can be found at github.com/joyichow/wolbachia-quasiexperiments.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Author details
[1]Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore 308232, Singapore. [2]Saw Swee Hock School of Public Health, National University of Singapore, Singapore 117549, Singapore. [3]National Environment Agency, 11, Biopolis Way, #06-05/08, Helios Block, Singapore 138667, Singapore. [4]Pest and Environmental Research Group, Bio21 Institute, University of Melbourne, Melbourne, Vic 3010, Australia.

Received: 8 April 2024 Accepted: 18 July 2024
Published online: 06 August 2024

References

1. Wang GH, Gamez S, Raban RR, Marshall JM, Alphey L, Li M, et al. Combating mosquito-borne diseases using genetic control technologies. Nat Commun. 2021;12(1):4388.
2. Obi J, Gutiérrez-Barbosa H, Chua J, Deredge D. Current trends and limitations in dengue antiviral research. TropicalMed. 2021;6(4):180.
3. Troost B, Smit JM. Recent advances in antiviral drug development towards dengue virus. Curr Opin Virol. 2020;43:9–21.
4. Capeding MR, Tran NH, Hadinegoro SRS, Ismail HIHM, Chotpitayasunondh T, Chua MN, et al. Clinical efficacy and safety of a novel tetravalent dengue vaccine in healthy children in Asia: a phase 3, randomised, observer-masked, placebo-controlled trial. The Lancet. 2014;384(9951):1358–65.
5. Torresi J, Ebert G, Pellegrini M. Vaccines licensed and in clinical trials for the prevention of dengue. Hum Vaccin Immunother. 2017;13(5):1059–72.
6. Andersson N, Nava-Aguilera E, Arosteguí J, Morales-Perez A, Suazo-Laguna H, Legorreta-Soberanis J, et al. Evidence based community mobilization for dengue prevention in Nicaragua and Mexico (Camino Verde, the Green Way): cluster randomized controlled trial. BMJ. 2015;8(351):h3267.
7. Vanlerberghe V, Toledo ME, Rodríguez M, Gomez D, Baly A, Benitez JR, et al. Community involvement in dengue vector control: cluster randomised trial. BMJ. 2009;9(338):b1959.
8. Nazni WA, Hoffmann AA, NoorAfizah A, Cheong YL, Mancini MV, Golding N, et al. Establishment of Wolbachia strain wAlbB in Malaysian populations of aedes aegypti for dengue control. Curr Biol. 2019;29(24):4241-4248.e5.
9. Hoffmann AA, Ahmad NW, Wan MK, Cheong YL, Ahmad NA, Golding N, et al. Operationalized releases of wAlbB Wolbachia in Aedes aegypti lead to sharp decreases in dengue incidence dependent on Wolbachia frequency. medRxiv; 2023. p. 2023.11.08.23298240. Available from: https://www.medrxiv.org/content/10.1101/2023.11.08.23298240v1. Cited 2024 Jan 31.
10. Utarini A, Indriani C, Ahmad RA, Tantowijoyo W, Arguni E, Ansari MR, et al. Efficacy of Wolbachia-infected mosquito deployments for the control of dengue. N Engl J Med. 2021;384(23):2177–86.
11. National Environmental Agency. Wolbachia-Aedes Mosquito Suppression Strategy. n.d. Available from: https://www.nea.gov.sg/corporate-functions/resources/research/wolbachia-aedes-mosquito-suppression-strategy. Cited 2023 Dec 18.
12. Lim JT, Bansal S, Chong CS, et al. Efficacy of Wolbachia-mediated sterility to reduce the incidence of dengue: a synthetic control study in Singapore. The Lancet Microbe. Elsevier; 2024;0(0). Available from: https://www.thelancet.com/journals/lanmic/article/PIIS2666-5247(23)00397-X/fulltext. Cited 2024 Mar 14.
13. Ribeiro Dos Santos G, Durovni B, Saraceni V, Souza Riback TI, Pinto SB, Anders KL, et al. Estimating the effect of the wMel release programme on the incidence of dengue and chikungunya in Rio de Janeiro, Brazil: a spatiotemporal modelling study. Lancet Infect Dis. 2022;22(11):1587–95.
14. Pinto SB, Riback TIS, Sylvestre G, Costa G, Peixoto J, Dias FBS, et al. Effectiveness of Wolbachia-infected mosquito deployments in reducing the incidence of dengue and other Aedes-borne diseases in Niterói, Brazil: A quasi-experimental study. Donald C, editor. PLoS Negl Trop Dis. 2021;15(7):e0009556.
15. Velez ID, Tanamas SK, Arbelaez MP, Kutcher SC, Duque SL, Uribe A, et al. Reduced dengue incidence following city-wide wMel Wolbachia mosquito releases throughout three Colombian cities: Interrupted time series analysis and a prospective case-control study. Viennet E, editor. PLoS Negl Trop Dis. 2023;17(11):e0011713.
16. Ryan PA, Turley AP, Wilson G, Hurst TP, Retzki K, Brown-Kenyon J, et al. Establishment of wMel Wolbachia in Aedes aegypti mosquitoes and reduction of local dengue transmission in Cairns and surrounding locations in northern Queensland, Australia. Gates Open Res. 2020;8(3):1547.
17. Ogunlade ST, Adekunle AI, Meehan MT, McBryde ES. Quantifying the impact of Wolbachia releases on dengue infection in Townsville, Australia. Sci Rep. 2023;13(1):14932.
18. Adebamowo C, Bah-Sow O, Binka F, Bruzzone R, Caplan A, Delfraissy JF, et al. Randomised controlled trials for Ebola: practical and ethical issues. The Lancet. 2014;384(9952):1423–4.
19. Resnik DB. Ethical issues in field trials of genetically modified disease-resistant mosquitoes. Dev World Bioeth. 2014;14(1):37–46.

20. Nichol AA, Antierens A. Ethics of emerging infectious disease outbreak responses: Using Ebola virus disease as a case study of limited resource allocation. PLoS ONE. 2021;16(2):e0246320.
21. Quintero J, Ronderos Pulido N, Logan J, Ant T, Bruce J, Carrasquilla G. Effectiveness of an intervention for Aedes aegypti control scaled-up under an inter-sectoral approach in a Colombian city hyper-endemic for dengue virus. PLoS ONE. 2020;15(4):e0230486.
22. Lim JT, Dickens BSL, Chew LZX, Choo ELW, Koo JR, Aik J, et al. Impact of sars-cov-2 interventions on dengue transmission. PLoS Negl Trop Dis. 2020;14(10):e0008719.
23. Lim JT, Mailepessov D, Chong CS, Dickens B, Lai YL, Ng Y, et al. Efficacy of Wolbachia-mediated sterility for control of dengue: emulation of a cluster randomized target trial. medRxiv; 2023. p. 2023.11.29.23299172. Available from: https://www.medrxiv.org/content/10.1101/2023.11.29.23299172v1. Cited 2024 Jan 30.
24. Mancini MV, Ant TH, Herd CS, Martinez J, Murdochy SM, Gingell DD, et al. High temperature cycles result in maternal transmission and dengue infection differences between wolbachia strains in Aedes aegypti. mBio. 2021;12(6):e00250-21.
25. Ross PA, Wiwatanaratanabutr I, Axford JK, White VL, Endersby-Harshman NM, Hoffmann AA. Wolbachia Infections in Aedes aegypti differ markedly in their response to cyclical heat stress. PLoS Pathog. 2017;13(1):e1006006.
26. Ooi EE, Wilder-Smith A. Externalities modulate the effectiveness of the Wolbachia release programme. Lancet Infect Dis. 2022;22(11):1518–9.
27. Abadie A, Diamond A, Hainmueller J. Synthetic control methods for comparative case studies: estimating the effect of california's tobacco control program. J Am Stat Assoc. 2010;105(490):493–505.
28. Knapp TR. Why is the one-group pretest-posttest design still used? Clin Nurs Res. 2016;25(5):467–72.
29. Bonander C. A (Flexible) synthetic control method for count data and other nonnegative outcomes. Epidemiology. 2021;32(5):653–60.
30. Abadie A, Vives-i-Bastida J. Synthetic Controls in Action. arXiv; 2022. Available from: http://arxiv.org/abs/2203.06279. Cited 2024 Jan 30.
31. Semiparametric AA, Estimators D-i-D. Rev Econ Stud. 2005;72(1):1–19.
32. Hausman C, Rapson DS. Regression Discontinuity in Time: Considerations for Empirical Applications. National Bureau of Economic Research; 2017. (Working Paper Series). Available from: https://www.nber.org/papers/w23602. Cited 2024 Jan 30.
33. Imbens G, Kalyanaraman K. Optimal Bandwidth choice for the regression discontinuity estimator. Rev Econ Stud. 2012;79(3):933–59.

Publisher's Note