Protocol

This trial protocol has been provided by the authors to give readers additional information about their work.

Protocol for: Utarini A, Indriani C, Ahmad RA, et al. Efficacy of wolbachia-infected mosquito deployments for the control of dengue. N Engl J Med 2021;384:2177-86. DOI: 10.1056/NEJMoa2030243

Protocol and Statistical Analysis Plan

This supplement contains the following items:

1. Original protocol (version 1), final protocol (version 5.1), and summary of changes
2. Original statistical analysis plan (version 1.3), final statistical analysis plan (version 1.6), and summary of changes

Title: **A non-blinded cluster randomised controlled trial to assess the efficacy of *Wolbachia*-infected mosquito deployments to reduce dengue incidence in Yogyakarta, Indonesia**

Short title: The impact of *Wolbachia* mosquito deployments on dengue in Yogyakarta

Version: 1.0

Trial registration identifier: this trial protocol will be registered with ANZCTR after obtaining ethics committee approval.

Protocol contributors:

Professor Cameron Simmons	Eliminate Dengue Program, Monash University, Australia
Dr Katherine Anders	Eliminate Dengue Program, Monash University, Australia
Ms Zoe Cutcher	Eliminate Dengue Program, Monash University, Australia
Dr Edwige Rances	Eliminate Dengue Program, Monash University, Australia
Professor Adi Utarini	Universitas Gadjah Mada, Yogyakarta, Indonesia Eliminate Dengue Project, Yogyakarta, Indonesia
Dr Riris Adono Ahmad	Universitas Gadjah Mada, Yogyakarta, Indonesia Eliminate Dengue Project, Yogyakarta, Indonesia
Dr Citra Indriani	Universitas Gadjah Mada, Yogyakarta, Indonesia Eliminate Dengue Project, Yogyakarta, Indonesia
Dr Eggi Arguni	Universitas Gadjah Mada, Yogyakarta, Indonesia Eliminate Dengue Project, Yogyakarta, Indonesia
Dr Warsito Tantowijoyo	Eliminate Dengue Project, Yogyakarta, Indonesia
Professor Nicholas Jewell	University of California, Berkeley, USA
Professor Christl Donnelly	Imperial College, London, UK
Professor Neil Ferguson	Imperial College, London, UK
Dr Peter Ryan	Eliminate Dengue Program, Monash University, Australia
Professor Scott O'Neill	Eliminate Dengue Program, Monash University, Australia

Trial sponsor:
Universitas Gadjah Mada, Yogyakarta, Indonesia

TABLE OF CONTENTS

List of figures:

List of tables:

Appendices:

1. Synopsis

1.1. Quick reference table

Primary registry and trial identification number	*Registration planned following ethical review*
Date of registration in primary registry	*Registration planned following ethical review*
Source of financial support	The Tahija Foundation, Jakarta, Indonesia
Sponsor	Universitas Gadjah Mada, Yogyakarta, Indonesia
Title	A non-blinded cluster randomised controlled trial to assess the protective efficacy of *Wolbachia* mosquito deployments for dengue control in Yogyakarta, Indonesia
Short title	CRT of *Wolbachia* against dengue
Study setting	Yogyakarta City, Indonesia
Health condition(s) studied	Dengue, Zika and chikungunya virus infection
Intervention	Intervention arm: Deployment of *Wolbachia*-infected *Aedes aegypti* mosquitoes, in addition to standard practice dengue control activities. Comparison arm: Standard practice dengue control activities.
Primary endpoint	Symptomatic, virologically-confirmed dengue virus (DENV) infection of any severity.
Secondary endpoints	Symptomatic, virologically-confirmed Zika virus (ZIKV) infection of any severity. Symptomatic, virologically-confirmed chikungunya virus (CHIKV) infection of any severity.
Study design	Study type: intervention study with test-negative design Allocation: cluster randomised Assignment: parallel 1:1 Masking: non-blinded Primary purpose: prevention
Study duration	36 months (12 months lead in during deployment and establishment; 24 months participant enrolment)
Target sample size	Allocation of the intervention will be randomised to 24 clusters (12 intervention and 12 untreated). All patients meeting the eligibility criteria will be invited to participate in the study. From baseline historical data we expect approximately 6000 participants per annum to be enrolled, among which 10-20% will be subsequently classified as virologically confirmed dengue. Power calculations estimate that approximately 1000 dengue cases and 4000 arbovirus-negative age and time-matched controls will be needed to detect a 50% or greater reduction in dengue incidence in

	Wolbachia-treated clusters compared to untreated clusters, with 80% power. These estimations are dependent upon assumptions regarding the expected distribution of cases and controls across clusters during the study period. As dengue distribution can fluctuate substantially from year to year, the required sample size will be re-calculated after twelve and twenty-four months of recruitment using the observed distribution of participants. Enrolment will continue for 24 months, or longer if required to achieve the minimum sample size.
Analysis	A Cox proportional hazards model with shared frailty will be used to estimate the relative hazard (incidence rate ratio) of dengue in *Wolbachia*-treated versus untreated clusters, accounting for time-matching of cases and controls and the non-independence of study participants resident in the same intervention cluster. The intention-to-treat analysis will consider *Wolbachia* exposure as binary depending on the allocation of the cluster of residence. The per-protocol analysis will consider *Wolbachia* exposure as a continuous weighted index based on *Wolbachia* prevalence in trapped mosquitoes in the cluster of residence and other clusters visited during the ten days prior to illness onset.

2. Introduction

2.1. Background

2.1.1. The burden of arboviral diseases transmitted by *Aedes aegypti*

The health and economic impacts of arboviral diseases transmitted by *Aedes aegypti* mosquitoes are escalating globally. The World Health Organisation (WHO) has stated that dengue is the most threatening and fastest spreading mosquito-borne disease, citing a 30-fold increase in global incidence during the past 50 years. A 2012 study suggested that almost 4 billion people in 128 countries are at risk of acquiring dengue [1]. In 2013, the estimated global burden of dengue was revised upward to 390 million infections per year [2], with almost 100 million infections manifesting some level of disease. The burden of dengue has a cost of ~$2.1 billion/year in the Americas [3] and almost $1 billion/year in Southeast Asia [4,5]. Clinically, dengue is a systemic viral illness of 3-7 days duration. Headache, fever, myalgia, anorexia and rash are common features. The defining pathophysiological feature of severe dengue is dysfunction of the vascular endothelium resulting in plasma leakage. When severe, plasma leakage can result in hypovolemic shock, a life threatening complication that requires urgent fluid resuscitation and other supportive care. Other features of severe dengue include leukopenia, thrombocytopenia and disturbed coagulation profiles that predispose to hemorrhagic tendencies, particularly at mucosal surfaces. Since the prognosis of dengue is difficult, many dengue cases are hospitalised for careful monitoring. As a consequence hospitals become overloaded with dengue cases and this places a significant economic impost on the health care system and to affected families. The only licensed medical specific intervention against dengue is the Dengvaxia vaccine [6]. Large phase III trials of Dengvaxia revealed both the burden of disease [7] and the vaccines complex efficacy profile, with highly variable efficacy across dengue virus (DENV) serotypes [8,9] and unanswered questions around long term efficacy and safety [10,11]. In April 2016, the WHO's Scientific Advisory Group of Experts (SAGE) gave qualified support to Dengvaxia, but with “guard rails” on where and how to use this complex intervention in endemic countries.

Another epidemic arbovirus, the chikungunya virus, came to global attention in 2004 when it caused epidemics on several Indian Ocean islands before spreading to southern Europe and South and South East Asia. Like dengue, chikungunya is a febrile systemic viral illness of 4-7 days duration. Debilitating polyarthralgia can be a long-lasting sequelae of chikungunya virus

infection [12]. In 2013, the chikungunya virus emerged again in the Caribbean and caused epidemics in Latin American countries that are ongoing [13]. There are no licensed vaccines or specific therapies for chikungunya.

Against a backdrop of endemic or epidemic dengue in over 100 countries, and recent explosive outbreaks of chikungunya, the Zika virus emerged in epidemic fashion in the Western Pacific in 2013 and in Latin America in 2015 [14]. As evidence accumulated that it causes congenital infections with severe outcomes including fetal death and severe microcephaly, it was declared a public health emergency of international concern (PHEIC) by the WHO[15]. Like chikungunya, there are no licensed vaccines or specific therapies for Zika.

There is a consensus that *Ae. aegypti* mosquitoes are the primary vectors of dengue, chikungunya and Zika. Hence the WHO has recommended well implemented vector control programmes against this species. The WHO also recommended the carefully planned pilot deployment, under operational conditions, of *Wolbachia*-based biocontrol accompanied by rigorous independent monitoring and evaluation [16].

2.1.2. Dengue in Indonesia

With a population of ~250 million, Indonesia is one of the largest dengue endemic countries in Asia. Correspondingly, the economic burden of dengue is estimated to be amongst the highest of countries in the region [4]. The first 58 dengue cases in Indonesia were reported from Jakarta and Surabaya in 1968 and thereafter dengue (or more specifically dengue hemorrhagic fever cases) was a notifiable disease [17]. Figure 1 shows the incidence of dengue hemorrhagic fever (DHF) since 1968 in Indonesia. Epidemic peaks have occurred at irregular intervals with a progressive increase in intensity, with large outbreaks evident in 1973, 1988, 1998, 2007, and 2010. Dengue remains predominantly a disease of children <15yrs of age in Indonesia, although there has been a trend towards increasing median age in the last decade [17]. In 2013, the five provinces with highest incidence of DHF were Bali (168.5/100,000 population), DKI Jakarta (104.0/100,000), DI Yogyakarta (96.0/100,000), East Kalimantan (92.7/100,000) and Sulawesi Tenggara (66.8/100,000).

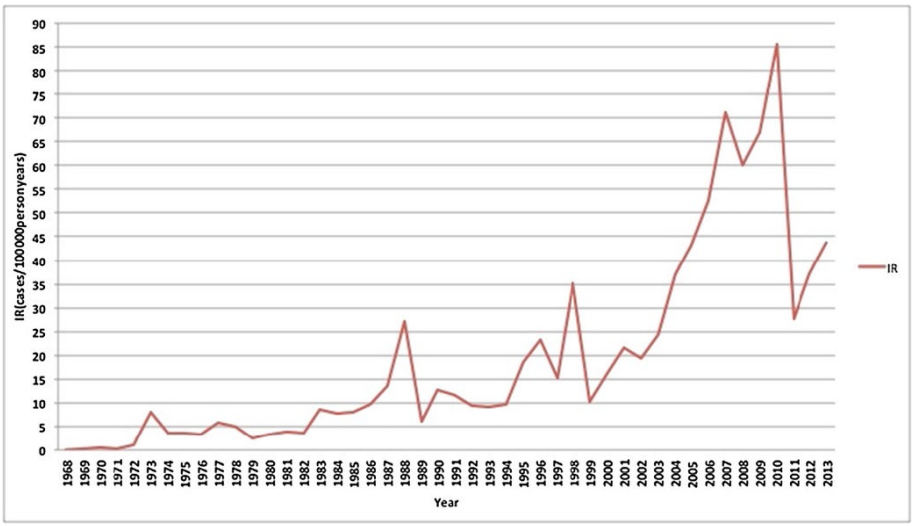


Figure 1. Trend in incidence rate of DHF cases in Indonesia from 1968 to 2013, measured in numbers of cases per 100,000 person years. Reproduced from [17]

2.1.3. Dengue in Yogyakarta

Dengue has been endemic in Yogyakarta for decades. Graham *et al* described high levels of anti-DENV antibody seroprevalence, indicating past infection, in Yogyakarta children in 1996 [18]. Between 2006 and 2014 the local public health surveillance system in Yogyakarta City received notification of 6,772 dengue hemorrhagic fever cases, including a large outbreak in 2010 (Figure 2). The large dengue epidemic in 2010 coincided with a national spike in disease incidence. These data reported to the surveillance system include only hospitalised cases that are classified as dengue hemorrhagic fever, so do not include the large ambulatory dengue patient population. The administrative area of Yogyakarta City, with a population in 2015 of 408,000 in an area of 32 km^2 [19], has generally had a higher dengue incidence than surrounding districts [20]. The seasonal distribution of dengue cases reported to Yogyakarta City health authorities between 2006 -2014 is shown in Figure 3.

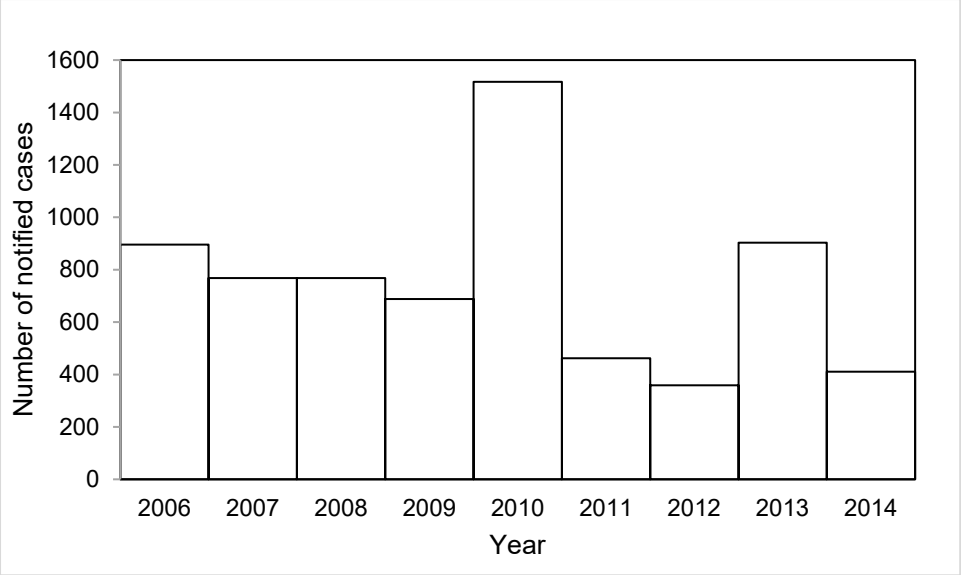


Figure 2. Dengue cases notified to the dengue surveillance system in Yogyakarta City (2006-2014)

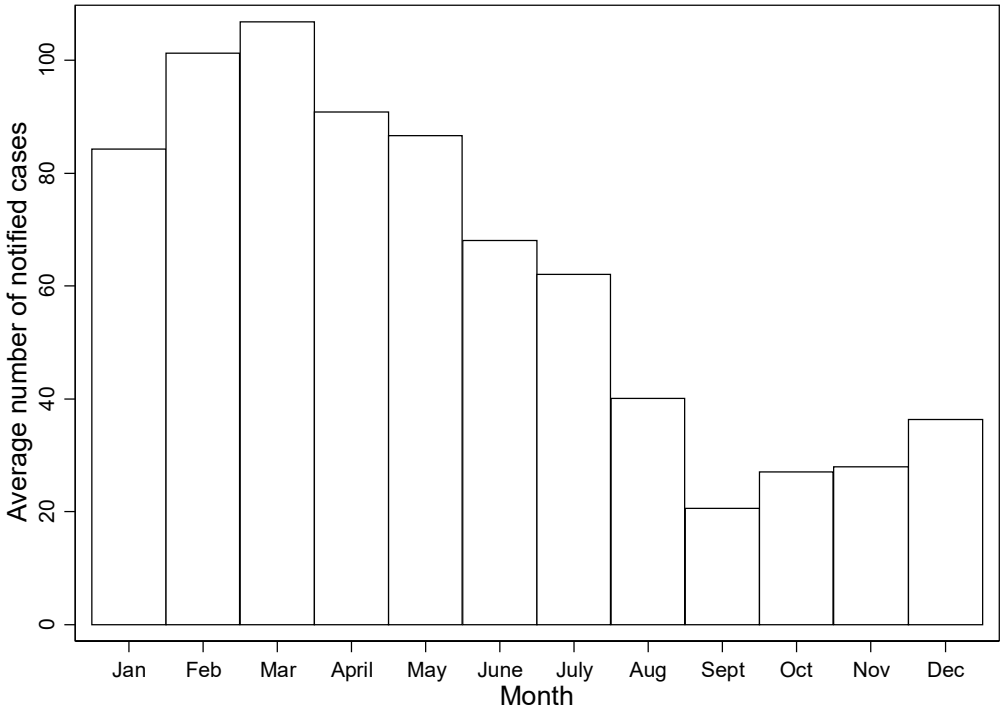


Figure 3. Average monthly dengue cases notified in Yogyakarta City (2006-2014)

2.1.4. Evidence for Zika and chikungunya transmission in Indonesia

Data on the prevalence of Zika and chikungunya (CHIK) in Yogyakarta are sparse and reflect limited availability of molecular diagnostics for these diseases in most clinical settings in Indonesia. A study from 1999 documented the presence of anti-CHIK antibodies in residents of Yogyakarta. Seroprevalence was ~30% in healthy volunteers [21], suggesting the transmission of CHIK, or of a closely-related, serologically cross-reactive alphavirus in Yogyakarta. Elsewhere in

Java independent studies have documented autochthonous CHIK transmission occurring between 2000-2011 [22–25]. Collectively these data confirm that CHIK transmission has occurred in Java, and likely continues to occur at a variable intensity.

The Zika virus has circulated in Asia for decades [26]. Zika viruses derived from Asia have been responsible for epidemics in the Western Pacific and Latin America and precipitated the WHO declaration of a public health emergency in early 2016. Although there is no data on Zika transmission in Yogyakarta, between December 2014 and April 2015 a Zika case was detected in a 27-year-old man in Jambi Province, central Sumatra, Indonesia [27]. The isolation and characterization of the Zika virus from this patient with no travel history confirms that the virus is circulating in Indonesia and that, by mimicking mild dengue infection, this infection is likely contributing to the large number of undiagnosed cases of acute febrile illness. The assumption is supported by confirmation of Zika infection among returned travellers following exposure in Jakarta in 2013 [28] and Bali in 2014 [29]. Since the clinical manifestations of CHIK and Zika virus infections are similar to that of uncomplicated dengue, and there is a paucity of specific diagnostics tests being performed for these pathogens, it is likely these two infections are underreported in the Indonesian archipelago. There are no licensed medical interventions for Zika or chikungunya.

2.1.5. Traditional vector control strategies to control dengue, chikungunya and Zika transmission

Vector control targeted against *Ae. aegypti* is the mainstay of the fight against dengue, chikungunya and Zika disease transmission. Integrated control strategies include (i) targeted residual spraying, (ii) space spraying, (iii) larval control and (iv) personal protection measures. The public health response to episodic dengue outbreaks in northern Australia relies upon active case finding and vector control to interrupt dengue transmission. However, successful broad-scale application of integrated vector control has been especially difficult to achieve in resource-limited endemic countries and impossible to sustain. Additionally, the evidence base to prioritise one intervention over another (e.g., larvicides and outdoor versus indoor insecticide space spraying), is weak as none have been robustly evaluated for impact on human infection and disease [30,31]. Some intervention trials have evaluated entomological impact [32], but reductions in mosquito populations do not correlate well with predictable reductions in dengue disease [33]. Collectively, the inability to rationally prioritise vector control interventions, coupled

with resource limitations in endemic settings, helps explain why contemporary vector-borne disease control programs have failed to stop regular epidemics and global dispersal of dengue, chikungunya and Zika.

2.1.6. The need for a strengthened evidence base for vector control interventions

A recent meta-analysis of entomological intervention trials demonstrated the remarkable paucity of reliable evidence for the effectiveness of any vector control method on dengue incidence [34]. Strikingly, none of the randomised controlled trials (RCTs) of vector control that were included in the meta-analysis investigated epidemiological impact (i.e. clinical disease endpoint) [34]. As examples, Andersson et al 2015 and Degener et al 2014 [35,36] recently reported cluster randomised trials of vector control for dengue but neither used an objectively measured clinical endpoint. The difficulty of making evidence-based policy in relation to vector control has resulted in calls for improved trial methods [37].

2.1.7. The Eliminate Dengue Program approach

The Eliminate Dengue Program is an international research collaboration that is delivering a paradigm shift in the control arboviral of diseases transmitted by *Ae. aegypti* mosquitoes. Our method utilises *Wolbachia*, obligate intracellular endosymbionts that are common in insect species [38–41] but were not present in *Ae. aegypti* mosquitoes until they were stably transinfected in the laboratory. In insects *Wolbachia* is maternally transmitted via the egg and manipulates insect reproduction to favour its own population dissemination via cytoplasmic incompatibility (CI). The result is that *Wolbachia* rapidly enter into naïve mosquito populations in a self-sustaining, durable manner. Multiple *Ae. aegypti*:*Wolbachia* combinations have been generated by the O'Neill laboratory where they form stable, maternally-transmitted infections that cause CI.[42–44] Strikingly, the presence of *Wolbachia* in *Ae. aegypti* mosquitoes renders them more resistant to disseminated arbovirus infection, including dengue, Zika, chikunguyna and Yellow fever viruses [45–47]. Thus the critical and signature effect of *Wolbachia* as a public health intervention is to severely reduce the vectorial capacity of mosquito populations to transmit arboviral infections between humans. For field implementation, the approach works by seeding wild mosquito populations with *Wolbachia* through controlled releases of relatively small numbers of *Wolbachia* infected mosquitoes (Figure 4). Over several months, and through the actions of CI, the prevalence of *Wolbachia* in the local mosquito population increases, until such time as the majority of mosquitoes in the area carry *Wolbachia*.

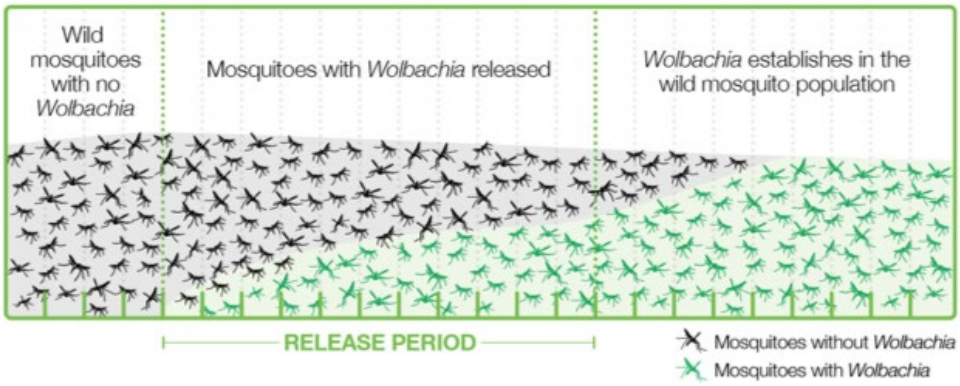


Figure 4. The *Wolbachia* biocontrol method. *Ae. aegypti* mosquitoes with *Wolbachia* (green) are released into the wild mosquito population (black).

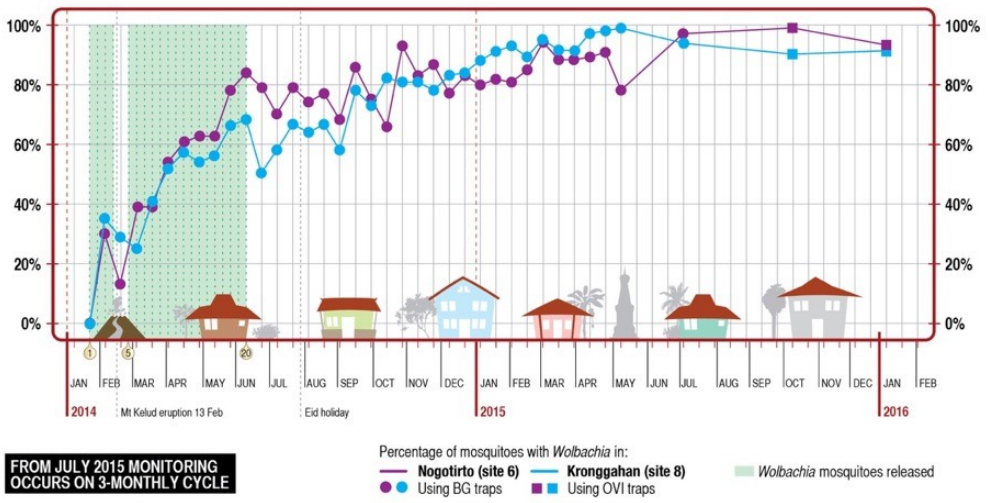
Over a series of releases, the percentage of *Wolbachia* mosquitoes increases. Once a threshold frequency of *Wolbachia* mosquitoes is reached, *Wolbachia* will continue to spread after releases have finished until the majority of mosquitoes carry *Wolbachia*. Laboratory vector competence studies show that *Wolbachia*-infected mosquitoes have a significantly reduced ability to transmit dengue, Zika and chikungunya viruses. Our Program has undertaken extensive vector competence assessments to determine the effect of *Wolbachia* (*w*Mel strain) on DENV infection and dissemination in *Ae. aegypti* mosquitoes [48–51]. Of note, we have elegantly demonstrated reduced vector competence in *Wolbachia*-infected mosquitoes obtained from the field using human dengue viremic blood and a novel read-out to measure infectious mosquito saliva [52]. *w*Mel viral interference effects were found to impact all four DENV serotypes, resulting in predicted reductions of 66-75% in the basic reproduction number R_0 for DENV-1-4 [52]. Reductions of this magnitude are predicted to result in local elimination of DENV transmission in most epidemiological circumstances [52].

The reduction in mosquito vector competence imparted by *Wolbachia,* together with the ability of *Wolbachia* (*w*Mel strain) to establish itself in *Ae. aegypti* populations, has led to regulatory and community acceptance of the technology in five countries; Australia (Cairns, Townsville), Indonesia (Yogyakarta), Vietnam (Nha Trang), Colombia (Medellin) and Brazil (Rio de Janeiro). Currently, approximately 160,000 people live under the protective umbrella of *w*Mel deployments. In 2016, in response to the emergence of Zika virus, the WHO endorsed pilot deployments of the *Wolbachia*-based biocontrol method to combat arboviral diseases [16].

2.1.8. Previous *Wolbachia* releases in Yogyakarta

Small-scale proof-of-concept field trials of *Wolbachia* (*w*Mel) deployment have been conducted in four small communities in districts adjacent to Yogyakarta City since 2014, with releases beginning in January 2014 in two sites in Sleman district and in November 2014 in two sites in Bantul district. In all sites, *Wolbachia* achieved a high prevalence in field-caught mosquitoes following the completion of releases, which has since been sustained (Figure 5). One year after *Wolbachia* establishment in Sleman (2015), *w*Mel-infected *Ae. aegypti* were collected from Nogotirto and Kronggahan field sites and injected with the four serotypes of dengue (isolated from East Timor) to look at dengue replication-blockage phenotype mediated by *w*Mel in wild-type *Ae. aegypti.* The data obtained show continued strong blockage of dengue 1-4 replication by *w*Mel.

A.



B.

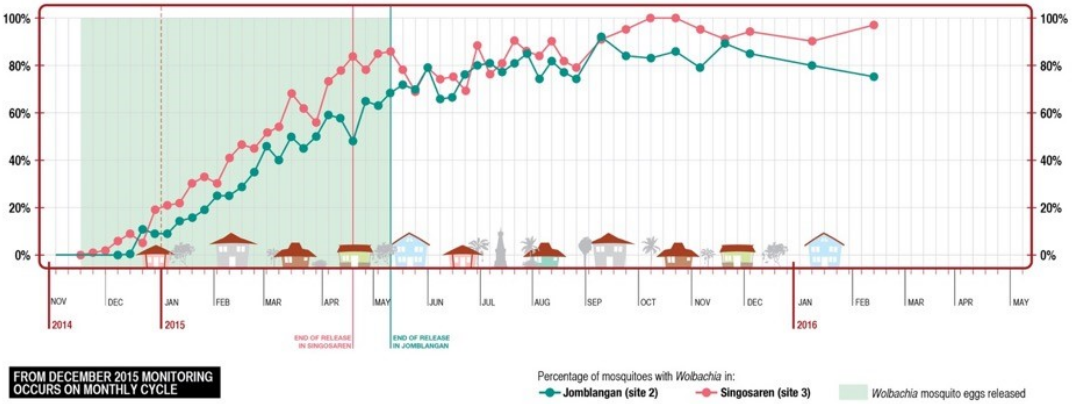


Figure 5. Establishment of *Wolbachi*a in A) Sleman and B) Bantul districts, Yogyakarta Province

2.2. Research question

Does large-scale deployment of *Wolbachia*-infected *Ae. aegypti* mosquitoes lead to a measurable reduction in dengue incidence in people living in release areas, compared to those living outside release areas?

2.3. Rationale

The successful introduction of the intracellular bacterium *Wolbachia* into *Ae. aegypti* mosquito populations is predicted to provide a long-term and sustainable approach to reducing dengue transmission. A critical next step, and the aim of the study outlined in this protocol, is to measure experimentally the efficacy of *Wolbachia* in reducing dengue virus transmission in the field. To this end, cluster randomised trials (CRTs) are the gold standard design to provide evidence on the efficacy of an intervention that has a community-wide impact [53]. The value of providing estimates of the epidemiological impact of *Wolbachia* from a randomised controlled trial are threefold. First, randomised controlled trials are the gold standard and preferred method for estimating the benefit or harm of health interventions. Second, results of CRTs are usually more influential in shaping policy and practice in medicine and public health than observational studies and this is likely to be true for the *Wolbachia* technology. Third, as noted previously, limitations in the design of most previous vector control trials for dengue mean that the scientific community expects that the *Wolbachia* technology be trialled using gold-standard methods wherever feasible [16,37].

3. Study objectives

3.1. Primary objective

To assess the efficacy of community-based deployments of *Wolbachia*-infected *Ae. aegypti* mosquitoes in reducing the incidence of symptomatic, virologically-confirmed dengue cases of any severity in Yogyakarta residents aged 3-30 years in release areas, relative to non-release areas.

3.2. Secondary objectives

- To measure the efficacy of the *Wolbachia* method in reducing the incidence of symptomatic virologically confirmed Zika virus and chikungunya virus infection in release areas, relative to non-release areas, and

- To quantify the level of human mobility within Yogyakarta City, and to estimate the degree to which this mobility reduces the power to measure the effect of a cluster-randomised intervention.

4. Study design

4.1. Type of study

This is a parallel two-arm non-blinded cluster randomised controlled trial which will be conducted in a single site in Yogyakarta City, Indonesia. The study site will be subdivided into twenty-four contiguous clusters, approximately 1km^2 in size (range 0.7km^2-1.65km^2), Figure 6. Clusters will be randomly allocated in a 1-to-1 ratio to receive *Wolbachia* deployments or no intervention, such that 12 clusters receive *Wolbachia* deployments and 12 receive no intervention.

There will be no buffer areas between clusters, but natural borders (roads, rivers, non-residential areas) will be used to define cluster boundaries as much as possible, to limit the spatial spread of *Wolbachia* from treated clusters into untreated areas, and of wild-type mosquitoes in *Wolbachia* treated clusters. Exclusion areas will be minimised, but any areas within the study site where releases are not possible for reasons of logistics, public acceptance or absence of mosquito populations (e.g. hospitals, public space, open parkland) will be pre-specified prior to randomisation and balanced between study arms. No attempt will be made to alter the routine dengue prevention and vector control activities conducted by public and private agencies throughout the study area (treated and untreated clusters). It is worth noting the capacity of the disease surveillance system to detect (and thus respond to) dengue will be enhanced across the city through increased availability of diagnostic kits, which have been supplied to primary care clinics and hospitals since March 2016 by the Eliminate Dengue Project Indonesia, to support efforts to enhance the surveillance of dengue across Yogyakarta.

The impact of *Wolbachia* deployments on dengue incidence will be assessed by comparing the exposure distribution (probability of living in a *Wolbachia*-treated area) among virologically-confirmed dengue cases presenting to a network of public primary clinics (Puskesmas), against the exposure distribution among patients with febrile illness of non-arboviral aetiology presenting to the same network of clinics in the same temporal windows. Dengue cases and arbovirus-negative controls will be sampled concurrently from within the population of patients presenting with febrile illness to the study clinic network, with case or control status classified retrospectively based on the results of laboratory diagnostic testing. The dataset for analysis will retain all enrolled cases and all controls that are matched to a case by calendar month of illness onset and age group (3-10; 11-20; 21-30 years), up to a maximum of 4 controls per case. Unmatched controls will be excluded from analysis.

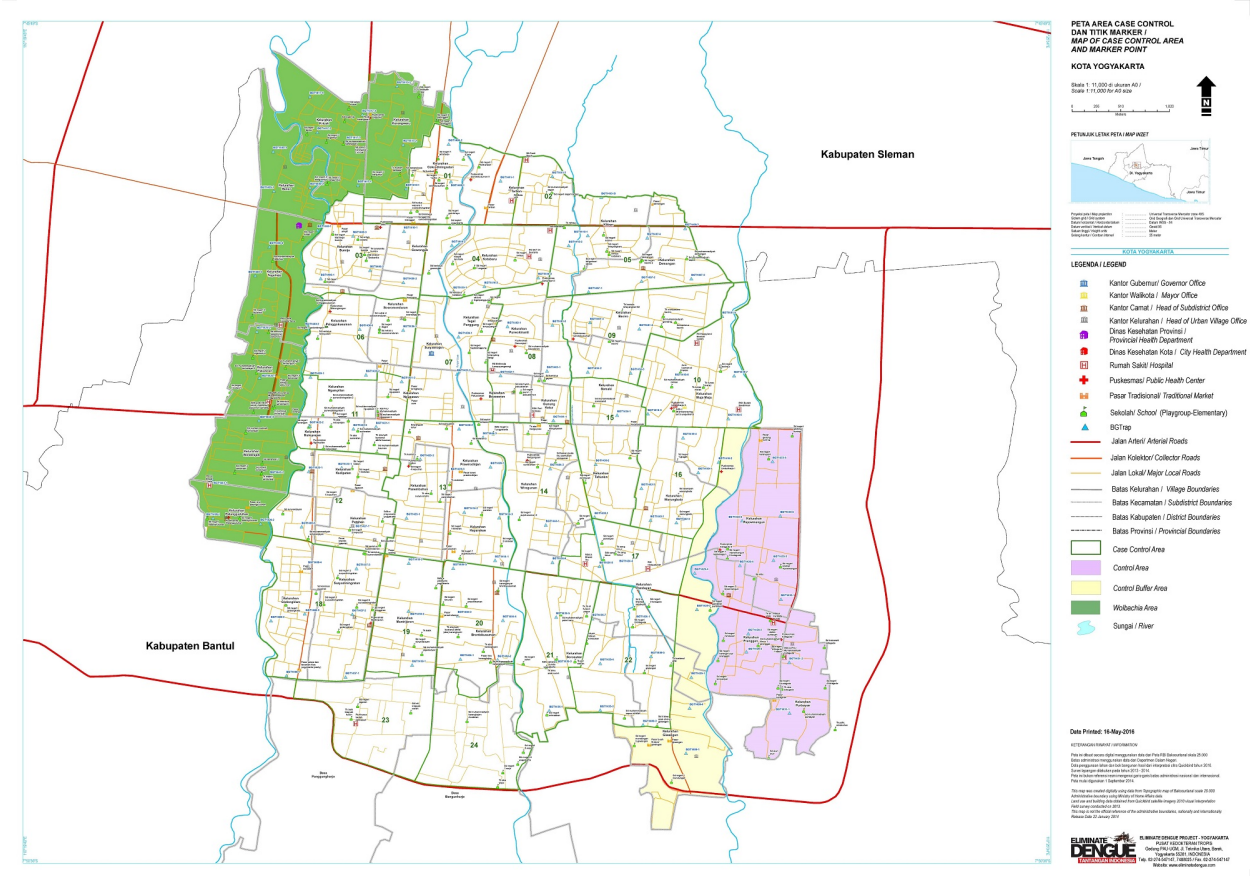


Figure 6. Map of study area, proposed cluster boundaries, and Puskesmas clinics. The study area is the non-shaded central area. Green lines outline proposed cluster boundaries. The green shaded area to the north-west represents a contiguous pilot deployment area, while the purple and yellow areas to the east of the city are non-release areas.

The distribution of *Wolbachia* exposure in the sampled arbovirus negative controls will reflect the distribution of *Wolbachia* exposure in the underlying source population that gave rise to cases, as long as a core assumption is met that the relative propensity to seek healthcare for febrile illness at a Puskesmas in intervention versus untreated arms is the same for dengue cases as other febrile illness controls. This should be upheld if dengue cases and other febrile illness controls are clinically indistinguishable until laboratory diagnosis. The concurrent sampling of controls and cases means that the odds of *Wolbachia*-exposure among sampled dengue cases relative to febrile controls (odds ratio), is an unbiased estimate of the relative incidence of medically-attended dengue in *Wolbachia*-treated versus untreated clusters (incidence rate ratio), from which protective efficacy can be estimated directly.

The null hypothesis is that the relative incidence of virologically-confirmed dengue in *Wolbachia*-treated and untreated areas is one. If *Wolbachia* has a protective effect against DENV transmission, we would expect the incidence rate ratio for virologically-confirmed dengue in *Wolbachia*-treated areas compared to untreated areas to be below one.

4.2. Justification of study design

Cluster randomised trials are the gold standard design to provide evidence on the efficacy of an intervention that has a community-wide impact[54]. This design involves the random allocation of intervention or no intervention to communities ('clusters') in the study area. Traditionally, an endpoint of infection and/or disease incidence would be measured in a prospective cohort of participants in *Wolbachia* treated and untreated clusters, and the measure of effect is the risk or rate ratio between intervention and untreated arms. Protective efficacy is estimated by one minus the risk/rate ratio. Given that DENV infection and/or disease in children living in an endemic area might be conservatively estimated at 5-10% and 1-5% per annum respectively[55], this would necessitate a large cohort of several thousand children, followed for several years, to detect a few hundred DENV infections each year.

An alternative method for measuring efficacy in a cluster randomised trial of *Wolbachia,* with a clinical endpoint, is to use concurrent sampling of dengue cases and non-dengue controls from within the whole study population to derive the estimate of the incidence rate ratio. In this method of concurrent sampling of controls – also called incidence density or risk-set sampling – the control group represents a sample of the underlying dynamic population from

which the cases arose, with an exposure distribution that is temporally matched to the timing of case onset [56]. This takes account of the time-varying nature of the exposure distribution (*Wolbachia* prevalence in local *Ae. aegypti* mosquitoes) and the seasonality of dengue. By recruiting participants from within the population of patients presenting to clinics with febrile illness – with dengue test-positive patients classified as cases and test-negative patients classified as controls – the controls are necessarily drawn from the same source population as the cases, thus avoiding the common pitfalls of control selection in traditional case-control studies that can introduce selection bias [57]. In this situation, the odds ratio is an unbiased estimate of the rate ratio in the source population over the period of participant enrolment (the 'risk' period), without the need for any rare disease assumption [56,58]. This approach to measuring the efficacy endpoint in a CRT has the advantage of being more efficient, cost effective, and logistically simpler to achieve than a large prospective cohort of children.

This study design has a precedent in the 'test negative design' (TND) used for evaluating the effectiveness of seasonal influenza vaccination. In that design, patients seeking health care for an acute respiratory illness (ARI) are recruited into the study and tested for influenza. Influenza vaccine effectiveness is then estimated as one minus the ratio of the odds of vaccination in subjects testing positive for influenza to the odds of vaccination in subjects testing negative [59]. Several authors have explored the statistical rationale and underlying assumptions of this design, and have demonstrated that the odds ratio for vaccination in influenza cases vs test-negative controls is directly equivalent to the relative risk of influenza in vaccinated vs unvaccinated individuals *if* test-negative controls are allowed to include participants who may test positive for influenza at any other time during the study period (i.e. risk-set sampling) [57] and if the distribution of non-influenza ARI is not associated with the intervention status [59]. The design outlined in this protocol extends the TND approach by including concurrent sampling of test-positive cases and test-negative controls, such that the odds ratio will approximate the rate ratio, rather than the risk ratio.

The design allows for differences in health care seeking behavior between vaccinated and unvaccinated individuals (i.e. *Wolbachia*-exposed and unexposed populations in our study) – e.g. due to spatial variation in the preferences for attending government vs private clinics - as long as the relative propensity between exposed and unexposed populations to seek care (and be enrolled) at a participating clinic is the same for test-positive and test-negative

patients[60]. This should be the case if test-positive and test-negative patients are clinically indistinguishable, and only classified after enrollment on the basis of subsequent laboratory diagnostic testing. The internal validity of the TND depends primarily upon the avoidance of selection bias in the sampling of cases and controls, and the extent to which controls can be assumed to be representative of the source population that gave rise to cases. A core assumption of the TND, translated to our study design, is that among persons who would seek care for febrile illness (at a participating clinic), the incidence of arbovirus-negative febrile illness does not differ, on average, between *Wolbachia*-treated and untreated areas [57,59,60]. If this is upheld, then the sampled controls will represent an unbiased estimate of the exposure distribution (i.e. residence in a *Wolbachia*-treated area) in the source population.

Our design introduces some advantages over the TND for influenza vaccine effectiveness, principally that the allocation of *Wolbachia* deployments is randomised. This helps to ensure that the intervention and non-intervention arms are balanced with respect to confounding variables, so that the study arms have similar baseline dengue risk and any measured difference in dengue incidence during the study period can be attributed to the effects of *Wolbachia*. This also means that, whereas the TND as applied to influenza vaccination is an observational study and can estimate only vaccine *effectiveness* under field conditions from the proportionate reduction of risk, ours is an experimental design and the proportionate reduction of risk gives an estimate of protective *efficacy* of *Wolbachia* [57]. A further difference is that *Wolbachia* is a community-level intervention, unlike influenza vaccination which is delivered to the individual; this introduces additional complexity into the analysis approach. However the methodological foundations of the TND, and many of the assumptions on which the statistical inference is based, translate well to the study design and analysis approach detailed in this protocol.

4.3. Number of participants

The study area as a whole has a population of approximately 350,000, of which approximately half will be resident in areas randomised to receive *Wolbachia* deployments and half in untreated areas.

The study population for measurement of the efficacy endpoint is the population of patients resident in the study area, presenting to the network of participating health clinics with

febrile illness, and meeting the eligibility criteria as described in section 7. Based on two years of historic data collated from the network of health clinics (Puskesmas) in the study area, it is estimated that approximately 6000 patients per year present to these clinics with febrile illness (range 200-1500 per clinic per annum).

We will enroll all participants presenting to any of the participating clinics who meet the eligibility criteria as described in section 7. Following laboratory testing and classification of participants' diagnostic status, matched sets of cases and controls will be formed based on month of illness onset and age, as described in 11.2.1. Enrolment will continue for two years, or longer if required to attain the minimum sample size for intention-to-treat analysis, as described in 11.1.

4.4. Expected duration of study

Wolbachia deployments are planned to commence in February 2017, and continue for approximately seven to eight months. The clinic-based sampling of febrile patients is expected to commence in pilot phase in September 2017, with active enrolment in all clinics by December 2017. The dataset to be included in the primary 'intention-to-treat' analysis will include only participants enrolled after *Wolbachia* is considered established in treated clusters, defined as when 80% of treated clusters reach a *Wolbachia* prevalence in trapped mosquitoes of ≥80% in two sequential screening events, or three months after completion of releases in the last cluster, whichever occurs first. The study timeline is depicted in Figure 7.

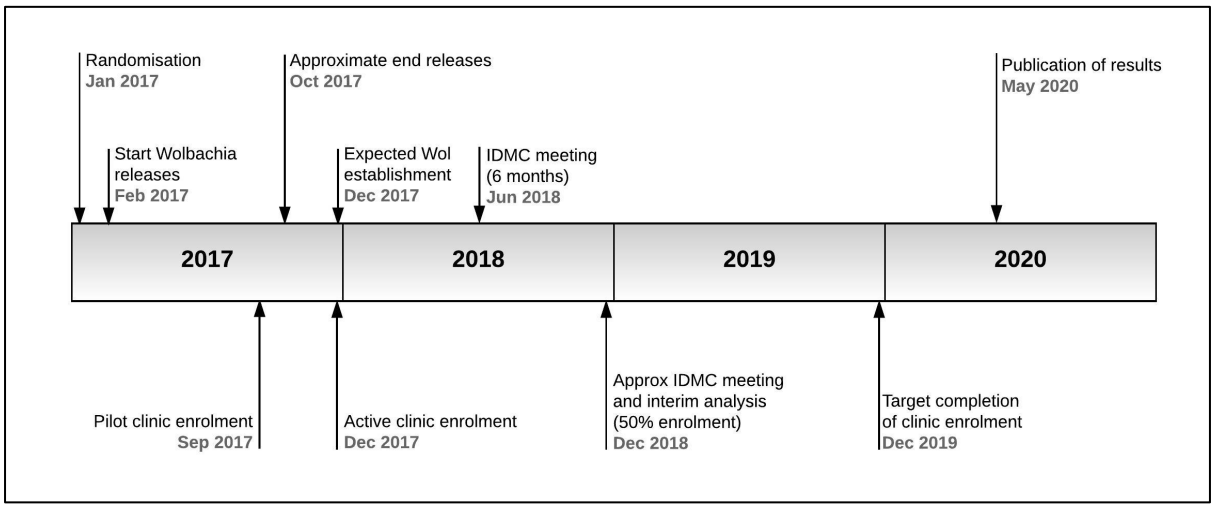


Figure 7. Study time line. Wol: *Wolbachia;* IDMC: Independent Data Monitoring Committee

4.5. Primary and secondary outcome measures

4.5.1. Primary outcome: dengue

The primary outcome measure will be virologically confirmed dengue virus infection in patients reporting febrile illness. Participants will be classified as dengue cases if plasma samples collected within 1-4 days after onset of fever test positive for dengue virus NS1 antigen (BioRad Platelia NS1 ELISA) and/or dengue virus nucleic acid by RT-qPCR.

4.5.2. Secondary outcomes: chikungunya and Zika

Secondary outcome measures include chikungunya and Zika virus infection in patients reporting febrile illness. Participants will be classified as virologically-confirmed chikungunya cases if chikungunya nucleic acid is detected in plasma samples by RT-qPCR. Participants will be classified as virologically-confirmed Zika virus cases if Zika virus nucleic acid is detected in plasma samples by RT-qPCR. A reduction in chikungunya and/or Zika virus infections association with *Wolbachia* deployment will be measured in the same way as for dengue, as described in section 11.3.1.

5. Study setting

The study will be conducted in Yogyakarta City and Bantul District, both located in the province of Yogyakarta Special Region, Indonesia. Yogyakarta City is 32 km^2 in size and had a population of 408,000 in 2015. The study site is 26 km^2 in size, including 24 km^2 within Yogyakarta City, and extending into 2km^2 of the adjacent administrative area, Bantul District, to the south of Yogyakarta City (see Figure 6). The study site is a continuous urban area, with a total population of approximately 350,000 and an average population density of 13,460 persons per km^2. The annual dengue incidence rate in Yogyakarta ranged between 83-390 cases per 100,000 population during the years 2006-2014. Even though dengue cases are reported all year, the high dengue season usually begins in December, peaks in March, and tapers off between June and July.

The study site will be subdivided into twenty-four contiguous clusters, each approximately 1km^2 in size (range 0.7km^2-1.65km^2). Among the 24 clusters, 12 will be randomised to receive *Wolbachia* deployments and 12 will be untreated.

Participant enrolment to measure the efficacy endpoint will be conducted at a network of health clinics throughout the study site. During the 6-12 months prior to commencement of *Wolbachia* deployments, data will be collected from a network of local government health clinics (Puskesmas) within the study site (Figure 6) on the number of patients meeting the clinical criteria for study participation, and their residential address, to determine the expected number and residential catchment area of potential participants. This will provide an indication of the feasibility of achieving the minimum target sample size, as described in 11.1.

6. Study intervention

6.1. Randomisation method

Constrained randomisation will be used to prevent a chance imbalance in the baseline characteristics or spatial distribution of treated and untreated clusters.

Covariate constrained (also referred to as "restricted") randomisation ensures balance and minimises loss of statistical power without the need for large numbers of strata, and is an excellent method to achieve balance when the number of clusters is small [54,61,62].

The approach to randomisation follows the method outlined by Hayes and Moulton [54]. First, all potential allocations of twelve intervention and twelve non-intervention clusters will be identified amongst the twenty-four clusters. Next, each allocation will be assessed against pre-defined balance criteria. All potential allocations that satisfy these balancing criteria will be retained, and non-balanced possibilities rejected. Finally, a single allocation pattern will be randomly selected from within the restricted list of balanced possibilities.

Traditionally, balancing variables include those that may be potentially confounding covariates; may impact sample size; or are relevant for logistics [54]. To avoid bias the TND approach also requires there is no association between the probability of other febrile illness (OFI) and the intervention [60]. To account for this given the propensity for spatial clustering of OFIs, randomisation will also balance on numbers of patients presenting to Puskesmas with undifferentiated fever during the preceding two years.

Sample size includes both the number of clusters and the total number of sampled individuals within each treatment arm [62], and is included because precision and power is maximised when sample sizes in treatment arms are similar. Balancing covariates are listed in Table 1.

Table 1. Balancing covariates

Reason for balance	Balancing covariates
Potential confounders	1. Age - % under 15 years 2. Average dengue incidence rate over most recent 3 years 3. Education - % completed high school, as a proxy for socioeconomic status
Potential sources of bias	4. Incidence of other febrile illness presenting to Puskesmas in 2014-15
Sample size	5. Number of clusters 6. Cluster population
Logistics	7. Total cluster area (km^2) 8. Total size of non-release area in cluster 9. Three spatial strata (to minimise number of contiguous areas)

6.2. *Wolbachia* deployment strategy

Wolbachia-infected *Ae. aegypti* will be deployed by setting Mosquito Release Containers (MRCs) in residential and non-residential properties throughout the intervention clusters. An MRC is a small plastic tub containing approximately 80 *Ae. aegypti* eggs, Tetramin food and water. Adult mosquitoes emerge from small holes in the side of the MRC, approximately 7-12 days after the MRC is deployed.

6.2.1. Density and duration of release

Clusters have been mapped and overlaid with 50-meter grids. One property in each grid square in the clusters allocated to *Wolbachia* deployments will host an MRC. Entomology field staff will set up the MRC in an outdoor area, by filling the container with water and adding the eggs and the Tetramin food to the container. Each MRC will be serviced every two weeks by adding a new batch of eggs, fresh food and water, at which point the previous release event will be classified as OK or Fail. Criteria for a fail include the container tipping over or missing, or presence of predators in the container.

As a quality assurance, a subset of MRCs will be checked and the number of emerged adult mosquitoes will be recorded based on the number of pupal skins and dead adults per container. The average emergence rate in each cluster will be used as a multiplier together with the number of eggs per MRC to estimate the number of adult mosquitoes released per cluster. Based on previous field work, the expected emergence rate is approximately 60%, and so a 50-meter grid network of MRCs (400 per km^2) with 80 eggs per MRC is expected to equate to 19,200 adults released per km^2, in each release week.

Eight fortnightly releases per cluster are planned (total duration of 16 weeks).
If *Wolbachia* prevalence in trapped mosquitoes is ≤60% in a specific cluster at the end of the release, the release will continue in that particular cluster until the threshold of 60% is reached.

Once *Wolbachia* prevalence has reached ≥60% in a cluster and releases have stopped, there will be no remediation with additional releases if *Wolbachia* prevalence drops below 60% in the future.

6.2.2. Timeline for completion of releases

Wolbachia will be deployed through rolling releases across treatment clusters within a 6-9 month period, with the aim of achieving *Wolbachia* establishment throughout treatment clusters within twelve months (from the start of the release). Deployments will continue in a cluster until the cluster-level *Wolbachia* prevalence in trapped *Ae. aegypti* reaches a pre-defined threshold of 60%. Deployment will then stop in that cluster and monitoring of *Wolbachia* prevalence in trapped mosquitoes will continue throughout the study period.

6.2.3. Handling individual and community-level refusal to release

Permission will be sought from community leaders (heads of Kelurahan administrative areas) prior to randomisation. If permission is not granted for a given Kelurahan or part of a Kelurahan, that area will be excluded when drawing cluster boundaries, and therefore from the study area for randomisation. Residents of these excluded areas who present to a study clinic will not be enrolled into the study, as they will not meet the inclusion criterion of residence within the study area.

Where individual householders refuse to host a mosquito release container, field staff will not release at the individual's home. Another release location within the same 50m^2 grid will be sought instead. Individuals from this household will still be eligible to participate in the clinic-based efficacy study.

6.3. *Wolbachia* monitoring strategy

6.3.1. Trapping method and density: during and post-deployment

A network of BG-Sentinel adult mosquito traps (BioGents) will be established across the study site prior to the commencement of releases, with a minimum of ten traps per cluster (intervention and untreated clusters) evenly spaced throughout residential areas, at a density of approximately one trap per 250-400 m. Additional traps (with a maximum density of 16 traps/km^2) will be set in intervention clusters during deployments to guide the release strategy, then removed when *Wolbachia* prevalence is stable at ≥80% for two consecutive screening events. BG traps will be serviced weekly, with trapped mosquitoes screened for *Wolbachia* at weekly, fortnightly or monthly intervals throughout the duration of the trial, depending on the stage of release and establishment. Mosquitoes will be bio-banked in the intervening weeks when screening is not done.

6.3.2. Laboratory methods for mosquito ID and screening

Trapped mosquitoes will be identified using microscopy, based on morphological criteria that allow differentiation of adult *Ae. aegypti* from other mosquito species present in Yogyakarta. *Ae. aegypti* collected from a single BG trap will be preserved together, but separated by female and male, in tubes containing 80% Ethanol.

After identification, samples will be sent to the diagnostic laboratory and individual mosquitoes (male and female) will be homogenised in a buffer solution to extract DNA and screened using quantitative PCR assay to detect the presence of *Wolbachia* and to confirm the species as *Ae. aegypti*. For each tube tested, corresponding to male or female *Ae. aegypti* from a single BG trap, the data recorded will include the number tested, the number positive by *Ae. aegypti* PCR, and the number positive by *Wolbachia* PCR. The *Wolbachia* prevalence in trapped *Ae. aegypti* will be reported aggregated to the cluster level. A minimum of 100 *Ae. aegypti* sampled from across 10 traps are required per cluster per screening event, in order to

have adequate precision around the point estimate (maximum uncertainty +/-10%). In the event that fewer than 100 *Ae. aegypti* are trapped within one week in a cluster during the post-establishment phase of the study, the mosquitoes trapped the following week can be added to the screened samples instead of bio-banked, in order to have a sufficiently large denominator for the *Wolbachia* prevalence estimate.

6.3.3. Definition of establishment

Establishment is defined by ≥80% *Wolbachia* prevalence in trapped *Ae. aegypti* (aggregated across all traps in the cluster), for two consecutive screening events. For the purposes of measuring the efficacy endpoint in the primary intention-to-treat analysis, *Wolbachia* will be considered established throughout intervention clusters when 80% of intervention clusters have had two consecutive screening events with ≥80% *Wolbachia* prevalence, or three months after completing releases in the last cluster, whichever occurs first.

7. Selection and enrolment of participants

A lag period between the *Wolbachia* releases and the start of clinical surveillance is planned to provide sufficient time for *Wolbachia* to establish in the wild *Ae. aegypti* population. Study processes for enrolling patients presenting with febrile illness will be established a network of primary care clinics (Puskesmas) throughout the study area. The clinic-based enrolment will operate in a pilot capacity from approximately September 2017, with a staged implementation across study clinics. The pilot period will be considered complete when study processes have been successfully implemented in all clinics. Recruitment will be continuous, with the dataset for the secondary 'per-protocol' analysis including all participants enrolled following the completion of the pilot period. The dataset for the primary 'intention-to-treat' analysis will include only participants enrolled after *Wolbachia* is considered established in the treated clusters (defined as ≥80% of treated clusters with *Wolbachia* prevalence ≥80%, or three months after completion of releases in the last cluster, whichever occurs first). Participants will be enrolled from within the population of patients (aged between 3-30 years old) presenting with undifferentiated fever of 1-4 days duration. All patients meeting the below inclusion criteria and providing written informed consent will be eligible for enrolment. Recruitment will continue for 24 months, at which point the required minimum sample size for intention-to-treat analysis is expected to have been achieved. In the event that the

minimum sample size has not yet been achieved, a protocol amendment to extend the study period will be submitted.

7.1. Recruitment procedures

All eligible participants meeting study inclusion criteria will be invited to enroll continuously throughout the study period. Recruitment will occur during normal clinic hours. Recruitment rates in each clinic and across the study site as a whole will be monitored monthly, including review of the screening logs to identify the proportion of eligible participants who did not consent to participate. The field coordinator will make regular visits to low-enrolling clinics to identify clinic-based, patient-based or other causes for low recruitment, and put measures in place to address these.

7.1.1 Screening log book

All patients presenting with febrile illness will be screened against the study inclusion criteria by trained staff. All eligible febrile individuals will be recorded in a screening log and invited to participate. Participation status (consent/decline) will be recorded against each participant in the log.

7.2. Informed consent procedures

Written informed consent will be sought from participants (or their guardian where the participant is a minor) by trained local staff, after explaining the study objectives, processes, data and sample collection and the participant has had an opportunity to ask questions. A verbal explanation of the written Explanatory Statement will be provided to all participants in the local language. In addition, participants aged between 13 and 17 years will be invited to sign an assent form indicating they understand the research and agree to participate. English translations of the Explanatory Statement, participant consent form and assent form are included as appendices; the documents used for participant recruitment will be in Bahasa Indonesia and all recruitment procedures will be conducted in the local language.

7.3. Inclusion criteria

Participants must meet the following inclusion criteria:

i) Fever (either self-reported or objectively measured, e.g. (tympanic membrane temperature ≥38^{o}C)) with a date of onset between 1-4 days prior to the day of presentation.

ii) Aged between 3-30 years old.

iii) Resided in the study area every night for the 10 days preceding illness onset.

7.4. Exclusion criteria

Participants will not be eligible for inclusion if any of the following are identified:

i) Localising features suggestive of an alternative diagnosis e.g. severe diarrhea, otitis, pneumonia

ii) Prior enrollment in the study within the previous 4 weeks.

An individual presenting to the clinic on repeat occasions for different febrile episodes will be eligible for enrollment during each different episode. However an individual may only be enrolled once during a single illness episode, which we define as illness occurring within 4 weeks of a previous febrile episode.

8. Data and sample collection procedures

8.1. Data to be collected

A unique identifier will be assigned to each participant at enrollment. Basic demographic details, eligibility against the inclusion criteria and illness onset date will be recorded in a standardised case report form. Table 2 summarises the data and samples to be collected from each participant. Data and samples are collected at a single time point at enrolment, with no longitudinal follow up of participants except for a phone call to establish their status at 14 days post-enrolment.

Table 2: Summary of data and samples collected

Data/sample type	Purpose

Demographic data (e.g. name, date of birth, address)	To uniquely identify participants; describe demographic characteristics of study population; ascertain study eligibility
Illness history data (e.g. symptoms, date of onset)	To ascertain study eligibility
Travel history in past 10 days (e.g. home and other places visited, including durations and geolocations)	To determine proportion of time spent in *Wolbachia*-treated and untreated clusters, for per-protocol analysis
3 ml venous blood sample	For DENV, chikungunya virus (CHIKV), Zika virus (ZIKV) diagnostic testing, in order to classify case/control status

8.1.1. Travel history

A brief travel history interview will be conducted at enrolment to determine the main places visited by each participant within the 10 days prior to illness onset, i.e. the incubation period for dengue. Thus, travel data are collected retrospectively. These data will be used to determine the proportion of time spent in *Wolbachia*-treated and untreated areas, for the per-protocol analysis. Because laboratory diagnostics are done retrospectively, interviewers will be blinded to the case/control status of the participant at the time of collection, which will avoid interviewer bias during collection of travel histories. However, potential bias in reporting of travel history between participants living in *Wolbachia*-treated areas and those living in untreated areas cannot be excluded, and will be minimised by the use of trained interviewers and standardised interview methods for eliciting travel histories.

8.1.2. Geolocation of participants residence and visited locations

The address of participants' residence and other locations visited during the 10 days prior to illness onset will be recorded during the travel history interview as above. The coordinates of locations visited will be verified by geo-locating on a map, and these geolocations retained for per-protocol analysis.

8.2. Data handling and record keeping

8.2.1. Roles and responsibilities of clinic staff and study staff

A study nurse will be stationed at each Puskesmas to ensure consistency in screening, recruitment and consent, data collection, sampling and transfer of specimens to laboratory. Clinical management and diagnostic testing will be provided by health center staff in accordance with standard of care. The field trial coordinator will oversee study processes in all participating Puskesmas clinics, to ensure adherence to the study protocol and standard operating procedures with respect to inclusion and exclusion criteria, informed consent procedures, case report form completion and the handling of samples and data.

8.3. Clinical sampling procedures

A single 3 ml venous blood will be collected from all consenting participants, on the day of enrolment. Blood samples from all participants will be transferred to the project laboratory on the day of collection and batch-tested within one month to determine case or control status (see Figure 8).

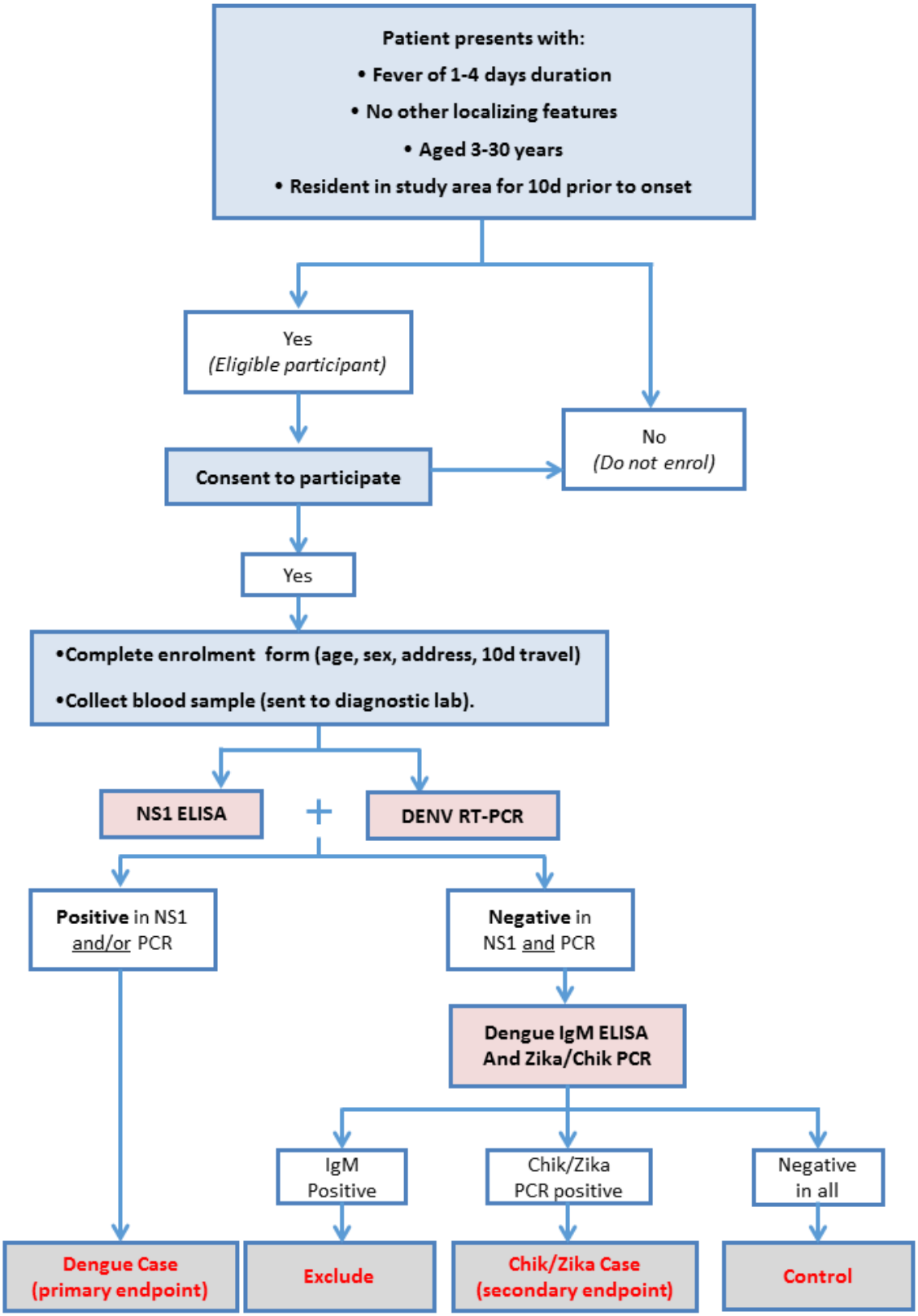


Figure 8. Flowchart of data and sample collection procedures and diagnostic algorithm

9. Laboratory assessments

9.1. Diagnostic testing for dengue, chikungunya and Zika

RT-qPCR is the gold standard method of diagnosing arboviral infections in the first few days of illness. We will use an internally controlled triplex RT-qPCR assay to detect DENV, CHIK and Zika viruses in plasma samples from all enrolled participants. Dengue NS1 Platelia ELISA (BioRad) and IgM serology (Panbio, Australia) will be performed according to the manufacturers' instructions. All research diagnostic investigations will be performed by the Eliminate Dengue Project diagnostics reference laboratory at the Universitas Gadjah Mada. External quality assurance panels will be used to monitor the performance of the molecular diagnostic tests.

9.2. Batch testing procedures

Diagnostic specimens will be tested in batch fashion in such a way as to maximise the throughput and minimise the cost of testing.

9.3. Sample handling and storage procedures

Clinical specimens will be collected and transferred to the reference laboratory according to standard operating procedures. All diagnostic specimens will be processed and stored on the same day as sample receipt and plasma stored at minus 80^{o}C.

9.4. Reporting of results

Diagnostic test results will not be reported back to individual participants since the testing will be performed in a research laboratory, not a certified diagnostic laboratory, and the batch processing of samples will mean that results are not available in time to inform clinical management. Participants will be managed according to standard clinical practice by the treating clinicians.

Given the potential risk of congenital Zika virus syndrome in a developing fetus exposed to Zika virus, we will report back to the primary care clinic a line listing of participants with positive results in Zika virus PCR at least once per month, so that standard procedures for

follow up of patients at risk of Zika virus infection can be followed at the clinician's discretion. The remaining blood specimen will be made available for forwarding to a diagnostic laboratory if the clinician requests it.

9.5. Case/control classification algorithm

Dengue cases are defined as patients with virologically-confirmed DENV infection, meeting the clinical criteria for enrolment and also with a positive result in NS1 ELISA and/or DENV RT-qPCR.

Controls are patients meeting the clinical criteria for enrolment, but with negative test results for CHIK RT-qPCR, Zika RT-qPCR, DENV NS1 ELISA, DENV RT-qPCR and DENV IgM ELISA (see Figure 8).

For the secondary endpoints, Zika or chikungunya cases are defined as patients with virologically confirmed Zika or chikungunya infections, meeting the clinical criteria for enrolment and also with a positive result in Zika RT-qPCR or CHIK RT-qPCR, respectively, and controls are defined as above.

10. Monitoring of unintended adverse effects of *Wolbachia* releases

Given the well-established safety profile of *Wolbachia*-infected *Ae. aegypti*, we do not anticipate any adverse effects associated with *Wolbachia* deployment during this trial. In order to demonstrate that the deployment is not associated with any excess of a severe adverse outcome, we will follow up all enrolled participants (test-positive cases and test-negative controls) by telephone within 7-14 days post-enrolment to ascertain their health status, recorded categorically as recovered/died, and whether or not they were ever hospitalised during this illness. The proportion of participants in each arm that were hospitalised or died will be reviewed by the Independent Data Monitoring Committee each time they meet (see Section 13.4), and at any other time at the request of the Trial Steering Committee or other agencies.

11. Statistical methods

11.1. Sample size estimation

It is estimated that approximately 1000 cases plus four times as many age- and time-matched controls will be sufficient to detect a 60% reduction in dengue incidence with 80% power. The estimate relies on several assumptions, outlined below. Sample size requirements will be re-estimated using observed data after 50% of the target recruitment is completed, to account for possible violations to these assumptions.

There are no published formulae to estimate sample size for the proposed study design, ie. a cluster randomised trial where the effect measure is the ratio of odds between study arms. Thus, we used the anticipated odds ratio to determine the proportion of enrolled participants resident in each cluster that are expected to be cases. We then estimated sample size requirements using a formula for comparison of proportions in cluster randomised trials (formula 7.6 in [54]), after solving the formula to give the harmonic mean number of participants per cluster required to detect the desired effect size, given a fixed power (80%) and number of clusters (n=24) (equation 1 and Table 3, below). We then estimated the total number of cases required by multiplying this number across the 24 clusters, with inflation of the harmonic mean to account for the heterogeneity in cluster size, using equation 2. The null hypothesis is that the proportion of total enrolled participants that are cases is the same in treated and untreated study arms. The alternative hypothesis is that the proportion of enrolled participants that are cases is lower in the *Wolbachia* treated arm than the untreated arm.

Equation (1): Harmonic mean number of participants per cluster, *H:*

$$H = \frac{{\pi_0 - \pi_0}^2 + {\pi_1 - \pi_1}^2}{(c-1)(\pi_0 - \pi_1)^2 / \left(z_{\alpha/2} + z_\beta\right)^2 - {k_0}^2 . {\pi_0}^2 + k^2 . {\pi_1}^2}$$

Equation (2): Total number of cases required to detect the desired effect, N:

$$N = \frac{2iHc}{r+1}$$

Table 3. Sample size parameters and assumed values

Parameter	Definition	Assumed value
π_0	Average proportion of participants that are cases, in non-intervention clusters	0.2459
π_1	Average proportion of participants that are cases, in intervention clusters	0.1703
k_0	Coefficient of variation for π_0	0.3841
k_1	Coefficient of variation for π_1	0.4563
c	Number of clusters per arm	12
$z_{\alpha/2}$	Standard normal distribution value corresponding to an upper tail probability of $\alpha/2$, and the significance difference obtained by a two-tailed significance test is P< α	1.96
z_β	Standard normal distribution value corresponding to an upper tail probability of β, where the sample size provides a power of 100(1-β)% to detect the difference in proportions	0.84
i	Inflation factor to convert harmonic mean to arithmetic mean, derived through simulations (*i*=arithmetic mean/harmonic mean)	1.30
r	Number of controls per case	4

Parameters (Table 3) were estimated using available historical data. Dengue data for the years 2003-2014 were sourced from the Yogyakarta dengue surveillance system. Calculations were based on dengue distribution from 2010-2011 as this period showed mid-range dengue heterogeneity. Data for other febrile illness during 2014-2015 were sourced from individual Puskesmas using ICD10 codes for non-localising fever (fever of unknown origin R50; Typhus A75.9; and acute infection due to bacteria at an unspecified site A49). Clusters were randomly allocated to intervention or untreated arms to estimate the likely values of π_0, π_1, k_0, $k_1$and *i*.

11.2. Analysis plan for primary endpoint

11.2.1. Matching criteria for controls and cases

Matched sets of cases and controls will be defined by matching enrolled confirmed dengue cases to arbovirus-negative controls who had illness onset in the same calendar month and are within the same age group (3-10; 11-20; 21-31 years). In the unlikely event that a minimum of four age-matched controls cannot be found for a case within the same calendar month, the window for matching can be extended until four matched controls are identified,

for that case only. For any calendar month where only controls are enrolled, and no cases, those controls will be excluded from analysis unless matched to a case using the criteria above. Each control will be included in the matched set for only a single case.

11.2.2. Intention-to-treat analysis

The intention-to-treat (primary) analysis will consider *Wolbachia* exposure as a binary classification based on residence in a cluster allocated to *Wolbachia* deployment or not. Residence will be defined as the primary place of residence during the 10 days prior to illness onset. The intention-to-treat analysis will be performed on data acquired during the case surveillance period, i.e. the 24-month period commencing when *Wolbachia* is deemed to have been established throughout intervention clusters, defined as ≥80% of treated clusters with *Wolbachia* prevalence ≥80% in two consecutive reads, or three months after completion of releases in the last cluster, whichever occurs first.

The association between *Wolbachia* deployment and incidence of dengue will be quantified in the time- and age-matched data sets using a Cox proportional hazards model with a shared frailty for cluster, which will provide an estimate of the incidence rate ratio (IRR, the relative hazard). This model will also provide a confidence interval for the IRR that accounts for correlation within clusters, and a p-value for the null hypothesis that the IRR equals one. The latter inference depends on the model being a reasonable fit to the trial data, and as an additional check we will perform a permutation of intervention allocation to quantify the probability that the trial result arose through chance alone, free from the assumptions implicit in the Cox model. The analysis dataset will be re-analysed against each potential balanced allocation identified during the restricted randomisation, and the effect estimates for each permutation compared against those obtained in the trial.

11.2.3. Per-protocol analysis

The per-protocol analysis will consider *Wolbachia* exposure as a quantitative index based on measured *Wolbachia* prevalence in local *Ae. aegypti* mosquitoes in the locations visited by the participant during the 10 days prior to illness onset. The per-protocol analysis therefore allows for *Wolbachia* exposure to vary in a location over time, and also accounts for human mobility, in terms of the exposure-time that individuals spend outside their cluster of

residence as reported in the travel history interview at enrolment. The per-protocol analysis will include all participants enrolled from the commencement of the main phase of clinic-based sampling (i.e. excluding the pilot phase, but including participants enrolled before *Wolbachia* was established in treated clusters).

Participants will be asked about their mobility during the ten days prior to illness onset using a structured interview administered at enrolment. This will record the duration of time spent at home, work or school, and up to three other most-visited locations during daylight hours (5am – 9pm) in the ten-day period. The geographic coordinates of those locations will be derived by geo-locating them on a digital map, with the assistance of the respondent. A weighted '*Wolbachia* exposure index' (WEI) will be calculated from the most recent cluster-level estimate of *Wolbachia* prevalence in trapped *Ae. aegypti* (n=10 traps per cluster) at each of the locations visited, multiplied by the proportion of time spent at each location, to give a value on a continuous scale from 0 to 1. The process of calculating WEI will be conducted blinded to participants' case/control status, by partitioning the travel history data from the laboratory diagnostic data, to remove any possibility of observer bias.

Cases and controls will be classified by strata of their WEI (e.g. 0-0.2; 0.2-0.4; 0.4-0.6; 0.6-0.8; 0.8-1). This acknowledges that the WEI is not a highly precise measure, and serves to reduce error in exposure classification. Another Cox proportional hazard model with shared frailty for cluster will be fitted, including the strata as categorical variables to calculate stratum-specific IRRs (relative to the baseline 0-0.2 stratum). This will allow examination of a 'dose response' relationship. An additional benefit of transforming WEI to a categorical variable is that it avoids any assumption of linearity in the dose response relationship.

11.3. Analysis of secondary objectives

11.3.1. Impact of *Wolbachia* deployment on Zika and chikungunya

There exists no baseline data on the prevalence of Zika or chikungunya infection among febrile patients presenting to primary health care clinics in Yogyakarta City, from which to estimate the expected number of cases, therefore these secondary analyses are exploratory only and not subject to any formal sample size or power calculations. Blood samples from enrolled participants will be tested by Zika and chikungunya PCR for the purpose of defining

arbovirus-negative controls for the primary analysis, as described above. These results will permit estimation of the prevalence of virologically confirmed Zika virus and chikungunya virus infection among the study population of ambulatory febrile patients presenting to primary health care.

If virologically confirmed Zika or chikungunya cases are detected, a secondary analysis will estimate the efficacy of *Wolbachia* deployments in reducing the incidence of symptomatic virologically confirmed Zika virus and chikungunya virus infection. The same enrolled patient population will be used to analyse all three arbovirus endpoints (dengue, Zika and chikungunya), and the same intention-to-treat and per-protocol analyses will be used as described for the primary (dengue) endpoint above. For Zika and chikungunya, the cases will be defined as enrolled participants who test positive by Zika or chikungunya PCR, respectively, and the controls will be those who test negative to all three arboviruses. Cases and controls will be matched by month of illness onset and age group, as described above (11.2.1). Statistical methods will be as described above (11.2.2 and 11.2.3).

11.3.2. Impact of *Wolbachia* deployment on notified dengue cases

The existing system for routine notification of dengue cases in Yogyakarta City is based on hospital-reporting of cases diagnosed clinically as Dengue Hemorrhagic Fever, which historically have not been accompanied by supportive laboratory testing. Since March 2016, hospitals have been encouraged to record a serological testing result, where available, on the report form, and also to report cases diagnosed clinically as Dengue Fever where there is a confirmatory NS1-positive test result. A separate reporting system, established in March 2016, collates data on the number of NS1 rapid tests performed – and number positive – in primary health clinics (Puskesmas) across the city. Both of these reporting systems include address information for notified cases.

We will collate data from these two reporting systems on a monthly basis from 2016-2020, aggregated by Kelurahan of residence, to monitor trends in reported dengue incidence across the City and by Kelurahan, before, during and after *Wolbachia* deployment.

11.3.3. Human mobility in Yogyakarta and implications for measuring efficacy of *Wolbachia* deployment

The level and distribution of human mobility among the study population is critical to success of this study design, as it determines the degree to which the per-protocol analysis can retain comparison groups with different levels of *Wolbachia* exposure after taking into account participant's crude movement patterns. The data captured through the travel history interview will be analysed to quantify the geographical extent and duration of participants' travel outside the home, and to estimate the degree to which this mobility reduces the power to measure the effect of the cluster-randomised *Wolbachia* intervention by making participants in the study arms more similar in terms of their true exposure distributions. An age-stratified analysis will describe the proportion of participant's time (5am – 9pm) spent at home versus away from home, and will estimate the distribution of participant's time as a function of increasing distance from home. This information can inform the design of future trials of cluster-randomised household-based interventions, by estimating the optimal size of the clusters needed to account for the majority of daily movements.

12. Data management

12.1. Data collection and coding

Field data on *Wolbachia* deployment and monitoring will be captured through standardised electronic data capture forms deployed on mobile devices. When connected to the internet, the devices will sync with a web-based Core Data Repository and all new data will be uploaded.

Data collected from participants in the clinical study will be similarly captured through standardised electronic data capture forms and digital mapping interfaces, deployed either on mobile devices or through web-based applications on desktop or laptop computers. Laboratory diagnostic results will be captured directly from laboratory assay output.

Validity controls will be applied at the point of data capture into electronic forms, by predefining value ranges, specifying categorical option lists, and minimising the use of free text fields. The use of carefully designed electronic forms will facilitate the coding of participant responses at the point of data collection.

12.2. Data storage and security

Field data on *Wolbachia* release and monitoring will be stored in the Core Data Repository, a custom designed relational database hosted on an Australian web-based server.

Clinical study data will be uploaded initially to a database hosted on an Indonesian server, to comply with Indonesian requirements for storage of identifiable data. The cluster of residence will be derived from the geo-coordinates of the home address for each participant, and similarly the cluster of visited locations recorded in the travel history interview will be derived. A dataset including all variables except identifiers (name, address, geocoordinates) will be extracted each week from the Indonesia-based database and imported into the Core Data Repository.

In order to maintain blinding of research staff and data managers, measures will be put in place to ensure the datasets identifying participant's exposure status (cluster of residence and clusters visited during 10 days prior to illness) will remain unlinked and partitioned from the dataset that classifies their case/control status until the final analysis. In the event that the Independent Data Monitoring Committee requires data to be unblinded following the interim analysis, a single member of the Eliminate Dengue Program, Monash University data management unit will be responsible for linking the participant dataset to the exposure status.

Role based, tiered access permissions will be used to control access to the Core Data Repository and associated data capture applications. User logs will document the activities of all users. Security of the web-hosted Core Data Repository will be assured by the security processes of the cloud service (Amazon Web Services), namely: automated backups and database snapshots, high-level availability and 24/7 incident response and detection. The overall Core Data architecture has been subject to a security audit by Monash University's IT operations, eSolutions.

12.3. Data quality assurance

Quality control in the form of logic and consistency checks will be applied at several stages of data capture and management: i) at the point of data capture into an electronic form; ii) at the point of upload into the web-based database; and iii) during routine monitoring processes

by internal and external data monitors. An audit trail will be preserved within the database to capture the history of any changes made to data records after their initial capture.

12.4. Study record retention

All data relating to the trial, including field entomology and epidemiological data, will be retained indefinitely, and for a minimum of 5 years after study completion, in accordance with ICH-GCP requirements.

13. Ethical considerations and trial governance

13.1. Summary of governance structure

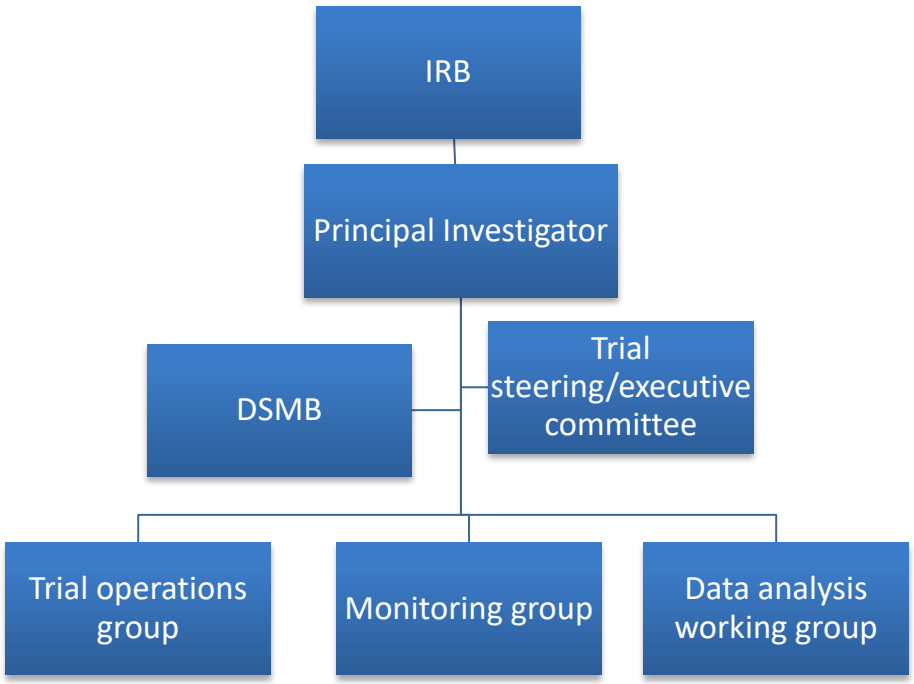


Figure 9: Trial governance structure

The Principal Investigator (PI) from Universitas Gadjah Mada, Yogyakarta, supported by the Chief Investigator from Monash University, will be responsible for ensuring the study is performed in compliance with the approved protocol and the principles of Good Clinical Practice.

The Trial Steering Committee (TSC), chaired by the PIs, will include one or more co-investigators and one or more members who are independent of the investigators and sponsors. The TSC will provide overall supervision of the trial, including monitoring of recruitment progress, and will consider and act upon (as appropriate) any recommendation from the DSMB with regards to early stopping of the trial.

The Trial Operations Group will, under the delegation of the PI, be responsible for day-to-day coordination of the trial processes.

The Monitoring Group will be independent of the investigators, and will conduct periodic monitoring of study processes including data collection and storage, sample collection and chain of custody, and laboratory processes.

The Data Analysis Working Group will be chaired by the trial statistician, and will be responsible for developing the statistical methods for randomization, data cleaning and validation, and preparing and implementing the statistical analyses.

13.2. Ethical review

This protocol and the informed consent document and any subsequent modifications will be reviewed and approved by the institutional review boards (IRBs) of Universitas Gadjah Mada, Yogyakarta, and Monash University, Melbourne, prior to the commencement of the trial. A letter of protocol approval by the ethical review boards will be obtained prior to the commencement of the trial.

If any substantive changes to study processes are required after commencement of the study, a protocol amendment request will be submitted to both review boards.

13.3. Modifications to the protocol

This study will be conducted in compliance with the current approved version of the protocol. Any change to the protocol document or informed consent form that affects the scientific intent, study design, participant safety, or may affect a participant's willingness to participate in the study is considered an amendment, and therefore will be written as a protocol amendment and submitted to the ethical review boards for approval prior to becoming effective.

13.4. Independent Data Monitoring Committee

An Independent Data Monitoring Committee (IDMC) will be constituted from local and international experts in accordance with standard practice for randomised clinical trials.

The IDMC will meet at study initiation, six months following the commencement of clinic-based enrolment, and at 50% enrolment of the target participants sample size, as well as any other time at the request of the TSC or other agencies. Their primary role is to safeguard the interests of the trial participants, to assess the safety and efficacy of the intervention during the trial, and to monitor the overall conduct of the trial.

The IDMC will provide recommendations about stopping or continuing the trial, and may also make recommendations relating to trial procedures, and data management and quality control. Any proposed major modifications to the study protocol will be reviewed by the IDMC, and approval for a protocol amendment will be sought from the relevant IRBs, prior to their implementation. Detailed responsibilities and terms of reference will be set out in an IDMC charter, and agreed to by all IDMC members, prior to study commencement.

13.5. Interim analyses and stopping rules

An interim analysis of the primary endpoint (intention-to-treat analysis only, as described in 11.2.2) will be conducted when enrolment reaches 50% of the planned participant enrolment target. The study statistician and other members of the data analysis working group will prepare the dataset and analysis code for interim analysis, retaining blinding of the exposure status of dengue test-positive and test-negative study participants. These data and code will be provided to the IDMC independent statistician who will generate the tables and distribute the interim report among IDMC members.

The IDMC may recommend modification or termination of the study if analyses of data from the first 50% of the planned participant enrolment target indicate beyond reasonable doubt that exposure to *Wolbachia* confers a reduced risk of dengue in the intention-to-treat analysis. The Haybittle-Peto boundary [63], requiring $p<0.001$ at interim analysis to consider stopping for efficacy, will be used as a guidance. The IDMC may also recommend termination if preliminary data clearly suggest that *Wolbachia* is associated with an excess of dengue (or Zika or chikungunya) cases. A less conservative $p<0.01$ in direction of harm will be used as a guidance. Termination or modification may also be recommended for any other operational reason (e.g. participant enrolment rates), perceived safety concern, or external factor.

The final decision to terminate or modify the study rests with the TSC.

13.6. Confidentiality

Confidentiality of participant information will be strictly maintained at all times by the participating investigators, research staff, and the sponsoring institution. This confidentiality is extended to cover testing of biological samples in addition to the clinical, demographic and geospatial information relating to participating subjects. All laboratory specimens, reports, data collection forms and log books, and geo-located records will be identified by a coded ID number only to maintain participant confidentiality. All records that contain names or other personal identifiers, such as informed consent forms, will be stored separately from study records identified by ID numbers. All local databases will be secured with password-protected access systems. No information concerning the study or the data will be released to any unauthorised third party, without prior written approval of the sponsoring institution. Clinical or personal information will not be released without written permission of the subject, except as necessary for monitoring by an ethical review board or regulatory agencies. Reporting of study results will not be done in any way that permits identification of individual participants, or the location of their homes or other visited locations.

13.7. Participant reimbursement

A small gift will be provided to participants after completion of study processes, to acknowledge their contribution. The value of this gift will not exceed $1 USD per participant. Participants will not be paid for their participation, nor will the study team be liable for payment of any medical costs.

14. Dissemination and publications policy

14.1. Dissemination of trial results

The scientific integrity of the trial requires that only the results of final analyses will be disseminated publicly; there will be no dissemination of any interim analysis, unless the results lead to early stoppage of the trial. Dissemination of trial results, including any publications arising, will be subject to the prior approval of the Trial Steering Committee. Final trial results will be disseminated to community leaders, healthcare professionals, the public and other relevant stakeholders, as well as being submitted for publication in a scientific journal.

14.2. Publication plan

The trial findings will be submitted for peer review and publication in an appropriate open access journal. Every attempt will be made to reduce to a minimum the interval between the completion of data collection and the release of study results. After finalising recruitment, we expect to take no more than four months to prepare the final results paper for submission.

14.3. Authorship eligibility guidelines

Named protocol contributors will be included as authors on the primary report of trial findings, assuming that they have fulfilled international criteria for authorship at the time of manuscript submission. Authors will be expected to have made a substantive contribution to the design, conduct, interpretation and reporting of the trial.

14.4. Data sharing statement

A summary of the trial protocol will be published in an open access journal prior to study commencement, and the full trial protocol will be made publicly available within one year of the conclusion of data collection. The trial will be registered on an appropriate clinical trials database prior to study commencement.

15. References

1. Brady, O. J. *et al.* Refining the global spatial limits of dengue virus transmission by evidence-based consensus. *PLoS Negl. Trop. Dis.* **6,** e1760 (2012).
2. Bhatt, S. *et al.* The global distribution and burden of dengue. *Nature* **496,** 504–7 (2013).
3. Shepard, D. S., Coudeville, L., Halasa, Y. a, Zambrano, B. & Dayan, G. H. Economic impact of dengue illness in the Americas. *Am. J. Trop. Med. Hyg.* **84,** 200–7 (2011).
4. Shepard, D. S., Undurraga, E. a & Halasa, Y. a. Economic and disease burden of dengue in southeast Asia. *PLoS Negl. Trop. Dis.* **7,** e2055 (2013).
5. Shepard, D. S. *et al.* Cost-effectiveness of a pediatric dengue vaccine. *Vaccine* **22,** 1275–80 (2004).
6. Dengue Vaccine Initiative. Dengue vaccine candidates in clinical development. (2016). Available at: http://www.denguevaccines.org/vaccine-development. (Accessed: 13th June 2016)
7. L'Azou, M. *et al.* Symptomatic Dengue in Children in 10 Asian and Latin American Countries. *N. Engl. J. Med.* **374,** 1155–66 (2016).
8. Capeding, M. R. *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. *Lancet* **6736,** 1–8 (2014).
9. Villar, L. *et al.* Efficacy of a Tetravalent Dengue Vaccine in Children in Latin America. *N. Engl. J.*

Med. **372,** 113–123 (2015).

10. Guy, B., Lang, J., Saville, M. & Jackson, N. Vaccination Against Dengue: Challenges and Current Developments. *Annu. Rev. Med.* **67,** 387–404 (2016).
11. Hadinegoro, S. R. *et al.* Efficacy and Long-Term Safety of a Dengue Vaccine in Regions of Endemic Disease. *N. Engl. J. Med.* **373,** 1195–206 (2015).
12. Schilte, C. *et al.* Chikungunya virus-associated long-term arthralgia: a 36-month prospective longitudinal study. *PLoS Negl. Trop. Dis.* **7,** e2137 (2013).
13. Rolph, M. S., Foo, S. S. & Mahalingam, S. Emergent chikungunya virus and arthritis in the Americas. *Lancet. Infect. Dis.* **15,** 1007–8 (2015).
14. Weaver, S. C. *et al.* Zika virus: History, emergence, biology, and prospects for control. *Antiviral Res.* **130,** 69–80 (2016).
15. World Health Organization. *Outcome of the Emergency Committee regarding clusters of microcephaly and Guillain-Barre syndrome*. (2016).
16. World Health Organization. Mosquito (vector) control emergency response and preparedness for Zika virus. (2016). Available at: http://www.who.int/neglected_diseases/news/mosquito_vector_control_response/en/. (Accessed: 18th March 2016)
17. Karyanti, M. R. *et al.* The changing incidence of dengue haemorrhagic fever in Indonesia: a 45-year registry-based analysis. *BMC Infect. Dis.* **14,** 412 (2014).
18. Graham, R. R. *et al.* A prospective seroepidemiologic study on dengue in children four to nine years of age in Yogyakarta, Indonesia I. studies in 1995-1996. *Am. J. Trop. Med. Hyg.* **61,** 412–9 (1999).
19. BPS - Statistics Indonesia. Badan Pusat Statistik. (2016). Available at: www.bps.go.id.
20. Ramadona, A. L. *et al.* Prediction of Dengue Outbreaks Based on Disease Surveillance and Meteorological Data. *PLoS One* **11,** e0152688 (2016).
21. Porter, K. R. *et al.* A serological study of Chikungunya virus transmission in Yogyakarta, Indonesia: evidence for the first outbreak since 1982. *Southeast Asian J. Trop. Med. Public Health* **35,** 408–15 (2004).
22. Kosasih, H. *et al.* Evidence for Endemic Chikungunya Virus Infections in Bandung, Indonesia. *PLoS Negl. Trop. Dis.* **7,** e2483 (2013).
23. Riswari, S. F. *et al.* Study of viremic profile in febrile specimens of chikungunya in Bandung, Indonesia. *J. Clin. Virol.* **74,** 61–5 (2016).
24. Laras, K. *et al.* Tracking the re-emergence of epidemic chikungunya virus in Indonesia. *Trans. R. Soc. Trop. Med. Hyg.* **99,** 128–41 (2005).
25. Mulyatno, K. C., Susilowati, H., Yamanaka, A., Soegijanto, S. & Konishi, E. Primary isolation and phylogenetic studies of Chikungunya virus from Surabaya, Indonesia. *Jpn. J. Infect. Dis.* **65,** 92–4 (2012).
26. Haddow, A. D. *et al.* Genetic characterization of Zika virus strains: geographic expansion of the Asian lineage. *PLoS Negl. Trop. Dis.* **6,** e1477 (2012).
27. Perkasa, A. *et al.* Isolation of Zika Virus from Febrile Patient, Indonesia. *Emerg. Infect. Dis.* **22,** 924–5 (2016).
28. Kwong, J. C., Druce, J. D. & Leder, K. Zika virus infection acquired during brief travel to Indonesia. *Am. J. Trop. Med. Hyg.* **89,** 516–7 (2013).

29. Leung, G. H. Y., Baird, R. W., Druce, J. & Anstey, N. M. Zika virus infection in Australia following a monkey bike in Indonesia. *Southeast Asian J. Trop. Med. Public Health* **46,** 460–4 (2015).

30. Esu, E., Lenhart, A., Smith, L. & Horstick, O. Effectiveness of peridomestic space spraying with insecticide on dengue transmission; systematic review. *Trop. Med. Int. Health* **15,** 619–31 (2010).

31. Pilger, D., De Maesschalck, M., Horstick, O. & San Martín, J. L. Dengue outbreak response: documented effective interventions and evidence gaps. *TropIKA* **1,** (2010).

32. Erlanger, T. E., Keiser, J. & Utzinger, J. Effect of dengue vector control interventions on entomological parameters in developing countries: a systematic review and meta-analysis. *Med. Vet. Entomol.* **22,** 203–21 (2008).

33. Bowman, L. R., Runge-Ranzinger, S. & McCall, P. J. Assessing the relationship between vector indices and dengue transmission: a systematic review of the evidence. *PLoS Negl. Trop. Dis.* **8,** e2848 (2014).

34. Bowman, L. R., Donegan, S. & McCall, P. J. Is Dengue Vector Control Deficient in Effectiveness or Evidence?: Systematic Review and Meta-analysis. *PLoS Negl. Trop. Dis.* **10,** e0004551 (2016).

35. Andersson, N. *et al.* Evidence based community mobilization for dengue prevention in Nicaragua and Mexico (Camino Verde , the Green Way): cluster randomized controlled trial. *BMJ* **351,** h3267 (2015).

36. Degener, C. M. *et al.* Evaluation of the effectiveness of mass trapping with BG-sentinel traps for dengue vector control: a cluster randomized controlled trial in Manaus, Brazil. *J. Med. Entomol.* **51,** 408–20 (2014).

37. Wilson, A. L. *et al.* Evidence-based vector control? Improving the quality of vector control trials. *Trends Parasitol.* **31,** 380–90 (2015).

38. O'Neill, S. L. *et al.* In vitro cultivation of Wolbachia pipientis in an Aedes albopictus cell line. *Insect Mol. Biol.* **6,** 33–9 (1997).

39. Stouthamer, R., Breeuwer, J. A. & Hurst, G. D. Wolbachia pipientis: microbial manipulator of arthropod reproduction. *Annu. Rev. Microbiol.* **53,** 71–102 (1999).

40. Rousset, F., Vautrin, D. & Solignac, M. Molecular identification of Wolbachia, the agent of cytoplasmic incompatibility in Drosophila simulans, and variability in relation with host mitochondrial types. *Proc. Biol. Sci.* **247,** 163–8 (1992).

41. Hilgenboecker, K., Hammerstein, P., Schlattmann, P., Telschow, A. & Werren, J. H. How many species are infected with Wolbachia?--A statistical analysis of current data. *FEMS Microbiol. Lett.* **281,** 215–20 (2008).

42. McMeniman, C. J. *et al.* Stable introduction of a life-shortening Wolbachia infection into the mosquito Aedes aegypti. *Science* **323,** 141–4 (2009).

43. Joubert, D. A. *et al.* Establishment of a Wolbachia Superinfection in Aedes aegypti Mosquitoes as a Potential Approach for Future Resistance Management. *PLoS Pathog.* **12,** e1005434 (2016).

44. Walker, T. *et al.* The wMel Wolbachia strain blocks dengue and invades caged Aedes aegypti populations. *Nature* **476,** 450–3 (2011).

45. Dutra, H. L. C. *et al.* Wolbachia Blocks Currently Circulating Zika Virus Isolates in Brazilian Aedes aegypti Mosquitoes. *Cell Host Microbe* **19,** 771–4 (2016).

46. Johnson, K. N. The Impact of Wolbachia on Virus Infection in Mosquitoes. *Viruses* **7,** 5705–17 (2015).

47. Rainey, S. M., Shah, P., Kohl, A. & Dietrich, I. Understanding the Wolbachia-mediated inhibition

of arboviruses in mosquitoes: progress and challenges. *J. Gen. Virol.* **95,** 517–30 (2014).

48. Amuzu, H. E., Simmons, C. P. & McGraw, E. A. Effect of repeat human blood feeding on Wolbachia density and dengue virus infection in Aedes aegypti. *Parasit. Vectors* **8,** 246 (2015).

49. Ye, Y. H. *et al.* Wolbachia Reduces the Transmission Potential of Dengue-Infected Aedes aegypti. *PLoS Negl. Trop. Dis.* **9,** e0003894 (2015).

50. Frentiu, F. D. *et al.* Limited dengue virus replication in field-collected Aedes aegypti mosquitoes infected with Wolbachia. *PLoS Negl. Trop. Dis.* **8,** e2688 (2014).

51. Moreira, L. A. *et al.* A Wolbachia symbiont in Aedes aegypti limits infection with dengue, Chikungunya, and Plasmodium. *Cell* **139,** 1268–78 (2009).

52. Ferguson, N. M. *et al.* Modeling the impact on virus transmission of Wolbachia-mediated blocking of dengue virus infection of Aedes aegypti. *Sci. Transl. Med.* **7,** 279ra37–279ra37 (2015).

53. Wolbers, M., Kleinschmidt, I., Simmons, C. P. & Donnelly, C. a. Considerations in the design of clinical trials to test novel entomological approaches to dengue control. *PLoS Negl. Trop. Dis.* **6,** e1937 (2012).

54. Hayes, R. & Moulton, L. *Cluster randomized trials*. (Chapman & Hall/CRC, 2009).

55. Endy, T. P., Yoon, I.-K. & Mammen, M. P. Prospective cohort studies of dengue viral transmission and severity of disease. *Curr. Top. Microbiol. Immunol.* **338,** 1–13 (2010).

56. Vandenbroucke, J. P. & Pearce, N. Case-control studies: Basic concepts. *Int. J. Epidemiol.* **41,** 1480–1489 (2012).

57. Serres, G. De, Skowronski, D. M., Wu, X. W. & Ambrose, C. S. The test-negative design : validity , accuracy and precision of vaccine efficacy estimates compared to the gold standard of randomised placebo-controlled clinical trials. *Eurosurveillance* **18,** pii=20585 (2013).

58. Greenland, S. & Thomas, D. C. On the need for the rare disease assumption in case-control studies. *Am. J. Epidemiol.* **116,** 547–553 (1982).

59. Jackson, M. L. & Nelson, J. C. The test-negative design for estimating influenza vaccine effectiveness. *Vaccine* **31,** 2165–2168 (2013).

60. Haber, M. *et al.* A Probability Model for Evaluating the Bias and Precision of Influenza Vaccine Effectiveness Estimates from Case-Control Studies. *Epidemiol. Infect.* **143,** 1417–1426 (2015).

61. Smith, P., Morrow, R. & Ross, D. *Field trials of health interventions*. (Oxford University Press, 2015).

62. Ivers, N. M. *et al.* Allocation techniques for balance at baseline in cluster randomized trials: a methodological review. *Trials* **13,** 120 (2012).

63. Haybittle, J. L. Repeated assessment of results in clinical trials of cancer treatment. *Br. J. Radiol.* **44,** 793–7 (1971).

Title: **Applying *Wolbachia* to Eliminate Dengue (AWED): A non-blinded cluster randomised controlled trial to assess the efficacy of *Wolbachia*-infected mosquito deployments to reduce dengue incidence in Yogyakarta, Indonesia**

Short title: The AWED trial: Applying *Wolbachia* to Eliminate Dengue

Version: 5.1

Trial registration identifier: Clinicaltrials.gov identifier NCT03055585.

Protocol contributors:

Professor Cameron Simmons	World Mosquito Program*, Monash University, Australia
Dr Katherine Anders	World Mosquito Program, Monash University, Australia
Dr Edwige Rances	World Mosquito Program, Monash University, Australia
Dr Stephanie Tanamas	World Mosquito Program, Monash University, Australia
Professor Adi Utarini	Universitas Gadjah Mada, Yogyakarta, Indonesia Eliminate Dengue Project, Yogyakarta, Indonesia
Dr Riris Adono Ahmad	Universitas Gadjah Mada, Yogyakarta, Indonesia Eliminate Dengue Project, Yogyakarta, Indonesia
Dr Citra Indriani	Universitas Gadjah Mada, Yogyakarta, Indonesia Eliminate Dengue Project, Yogyakarta, Indonesia
Dr Eggi Arguni	Universitas Gadjah Mada, Yogyakarta, Indonesia Eliminate Dengue Project, Yogyakarta, Indonesia
Dr Warsito Tantowijoyo	Eliminate Dengue Project, Yogyakarta, Indonesia
Professor Nicholas P. Jewell	University of California, Berkeley, USA
Professor Christl Donnelly	Imperial College, London, UK
Professor Neil Ferguson	Imperial College, London, UK
Dr Peter Ryan	World Mosquito Program, Monash University, Australia
Professor Scott O'Neill	World Mosquito Program, Monash University, Australia

* Formerly Eliminate Dengue Program

Trial sponsor:
Universitas Gadjah Mada, Yogyakarta, Indonesia

TABLE OF CONTENTS

List of figures:

List of tables:

1. Synopsis

1.1. Quick reference table

Primary registry and trial identification number	Clinicaltrials.gov identifier: NCT03055585
Date of registration in primary registry	14 February 2017
Source of financial support	The Tahija Foundation, Jakarta, Indonesia
Sponsor	Universitas Gadjah Mada, Yogyakarta, Indonesia
Title	Applying *Wolbachia* to Eliminate Dengue (AWED): A non-blinded cluster randomised controlled trial to assess the protective efficacy of *Wolbachia* mosquito deployments for dengue control in Yogyakarta, Indonesia
Short title	CRT of *Wolbachia* against dengue
Study setting	Yogyakarta City, Indonesia
Health condition(s) studied	Dengue, Zika and chikungunya virus infection
Intervention	Intervention arm: Deployment of *Wolbachia*-infected *Aedes aegypti* mosquitoes, in addition to standard practice dengue control activities. Comparison arm: Standard practice dengue control activities.
Primary endpoint	Symptomatic, virologically-confirmed dengue virus (DENV) infection of any severity.
Secondary endpoints	Serotype-specific symptomatic, virologically-confirmed DENV infection of any severity. Symptomatic, virologically-confirmed Zika virus (ZIKV) infection of any severity. Symptomatic, virologically-confirmed chikungunya virus (CHIKV) infection of any severity.
Study design	Study type: intervention study with test-negative design Allocation: cluster randomised Assignment: parallel 1:1 Masking: non-blinded Primary purpose: prevention
Study duration	48 months (12 months lead in during deployment and establishment; up to 36 months participant enrolment)
Target sample size	Allocation of the intervention will be randomised to 24 clusters (12 intervention and 12 untreated). All patients meeting the eligibility criteria will be invited to participate in the study. From baseline historical data we expect approximately 5000 participants per annum to be enrolled, among which 10-20% will be subsequently classified as virologically confirmed dengue.

	Revised power calculations estimate that a minimum of 400 dengue cases and 1600 arbovirus-negative controls will be needed to detect a 50% or greater reduction in dengue incidence in *Wolbachia*-treated clusters compared to untreated clusters, with 80% power. These estimations are dependent upon assumptions regarding the expected distribution of cases and controls across clusters during the study period. Enrolment will continue for up to 36 months.
Analysis	Permutation tests and standard regression models (including the Cox proportional hazards model with shared frailty) will be used to estimate the relative risk and relative hazard (incidence rate ratio) of dengue in *Wolbachia*-treated versus untreated clusters, accounting for the non-independence of study participants resident in the same intervention cluster. The intention-to-treat analysis will consider *Wolbachia* exposure as binary depending on the allocation of the cluster of residence. The per-protocol analysis will consider *Wolbachia* exposure as a continuous weighted index based on *Wolbachia* prevalence in trapped mosquitoes in the cluster of residence, either with or without weighting for time spent in other clusters visited during the ten days prior to illness onset, and will also allow for time-matching of cases and controls.

2. Introduction

2.1. Background

2.1.1. The burden of arboviral diseases transmitted by Aedes aegypti

The health and economic impacts of arboviral diseases transmitted by *Aedes aegypti* mosquitoes are escalating globally. The World Health Organisation (WHO) has stated that dengue is the most threatening and fastest spreading mosquito-borne disease, citing a 30-fold increase in global incidence during the past 50 years. A 2012 study suggested that almost 4 billion people in 128 countries are at risk of acquiring dengue [1]. In 2013, the estimated global burden of dengue was revised upward to 390 million infections per year [2], with almost 100 million infections manifesting some level of disease. The burden of dengue has a cost of ~$2.1 billion/year in the Americas [3] and almost $1 billion/year in Southeast Asia [4,5]. Clinically, dengue is a systemic viral illness of 3-7 days duration. Headache, fever, myalgia, anorexia and rash are common features. The defining pathophysiological feature of severe dengue is dysfunction of the vascular endothelium resulting in plasma leakage. When severe, plasma leakage can result

in hypovolemic shock, a life threatening complication that requires urgent fluid resuscitation and other supportive care. Other features of severe dengue include leukopenia, thrombocytopenia and disturbed coagulation profiles that predispose to hemorrhagic tendencies, particularly at mucosal surfaces. Since the prognosis of dengue is difficult, many dengue cases are hospitalised for careful monitoring. As a consequence hospitals become overloaded with dengue cases and this places a significant economic impost on the health care system and to affected families. The only licensed medical specific intervention against dengue is the Dengvaxia vaccine [6]. Large phase III trials of Dengvaxia revealed both the burden of disease [7] and the vaccines complex efficacy profile, with highly variable efficacy across dengue virus (DENV) serotypes [8,9] and unanswered questions around long term efficacy and safety [10,11]. In April 2016, the WHO's Scientific Advisory Group of Experts (SAGE) gave qualified support to Dengvaxia, but with "guard rails" on where and how to use this complex intervention in endemic countries.

Another epidemic arbovirus, the chikungunya virus, came to global attention in 2004 when it caused epidemics on several Indian Ocean islands before spreading to southern Europe and South and South East Asia. Like dengue, chikungunya is a febrile systemic viral illness of 4-7 days duration. Debilitating polyarthralgia can be a long-lasting sequelae of chikungunya virus infection [12]. In 2013, the chikungunya virus emerged again in the Caribbean and caused epidemics in Latin American countries that are ongoing [13]. There are no licensed vaccines or specific therapies for chikungunya.

Against a backdrop of endemic or epidemic dengue in over 100 countries, and recent explosive outbreaks of chikungunya, the Zika virus emerged in epidemic fashion in the Western Pacific in 2013 and in Latin America in 2015 [14]. As evidence accumulated that it causes congenital infections with severe outcomes including fetal death and severe microcephaly, it was declared a public health emergency of international concern (PHEIC) by the WHO[15]. Like chikungunya, there are no licensed vaccines or specific therapies for Zika.

There is a consensus that *Ae. aegypti* mosquitoes are the primary vectors of dengue, chikungunya and Zika. Hence the WHO has recommended well implemented vector control programmes against this species. The WHO also recommended the carefully planned pilot

deployment, under operational conditions, of *Wolbachia*-based biocontrol accompanied by rigorous independent monitoring and evaluation [16].

2.1.2. Dengue in Indonesia

With a population of ~250 million, Indonesia is one of the largest dengue endemic countries in Asia. Correspondingly, the economic burden of dengue is estimated to be amongst the highest of countries in the region [4]. The first 58 dengue cases in Indonesia were reported from Jakarta and Surabaya in 1968 and thereafter dengue (or more specifically dengue hemorrhagic fever cases) was a notifiable disease [17]. Figure 1 shows the incidence of dengue hemorrhagic fever (DHF) since 1968 in Indonesia. Epidemic peaks have occurred at irregular intervals with a progressive increase in intensity, with large outbreaks evident in 1973, 1988, 1998, 2007, and 2010. Dengue remains predominantly a disease of children <15 years of age in Indonesia, although there has been a trend towards increasing median age in the last decade [17]. In 2013, the five provinces with highest incidence of DHF were Bali (168.5/100,000 population), DKI Jakarta (104.0/100,000), DI Yogyakarta (96.0/100,000), East Kalimantan (92.7/100,000) and Sulawesi Tenggara (66.8/100,000).

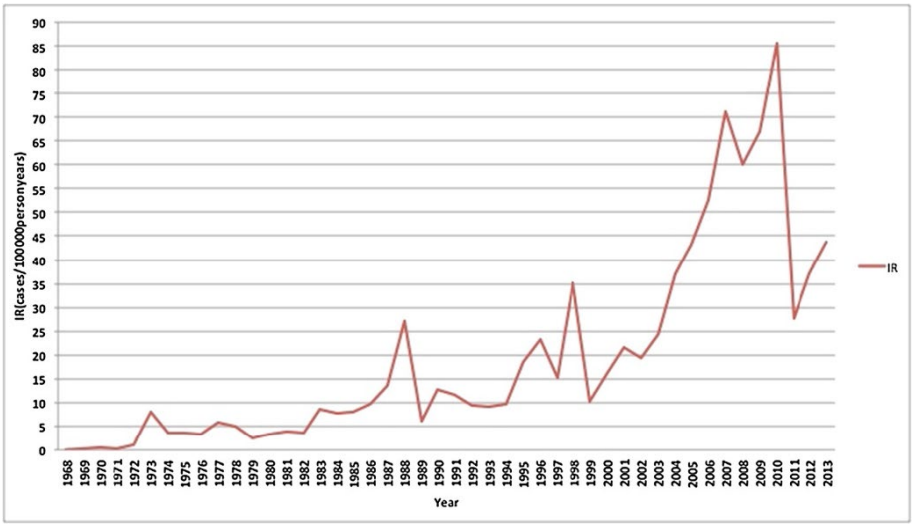


Figure 1. Trend in incidence rate of DHF cases in Indonesia from 1968 to 2013, measured in numbers of cases per 100,000 person years. Reproduced from [17]

2.1.3. Dengue in Yogyakarta

Dengue has been endemic in Yogyakarta for decades. Graham *et al* described high levels of anti-DENV antibody seroprevalence, indicating past infection, in Yogyakarta children in 1996 [18]. Between 2006 and 2014 the local public health surveillance system in Yogyakarta City received

notification of 6,772 dengue hemorrhagic fever cases, including a large outbreak in 2010 (Figure 2). The large dengue epidemic in 2010 coincided with a national spike in disease incidence. These data reported to the surveillance system include only hospitalised cases that are classified as dengue hemorrhagic fever, so do not include the large ambulatory dengue patient population. The administrative area of Yogyakarta City, with a population in 2015 of 408,000 in an area of 32 km^2 [19], has generally had a higher dengue incidence than surrounding districts [20]. The seasonal distribution of dengue cases reported to Yogyakarta City health authorities between 2006 -2014 is shown in Figure 3.

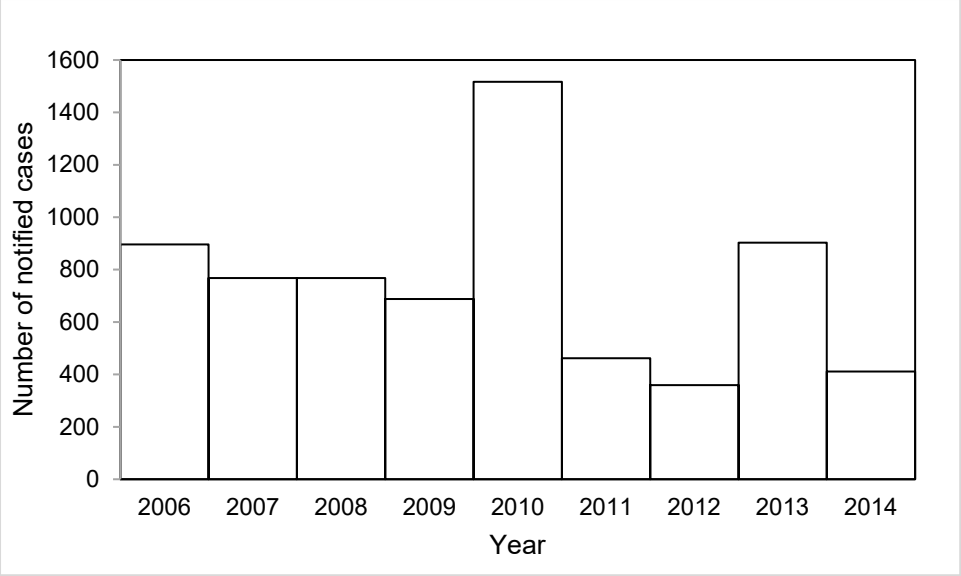


Figure 2. Dengue cases notified to the dengue surveillance system in Yogyakarta City (2006-2014)

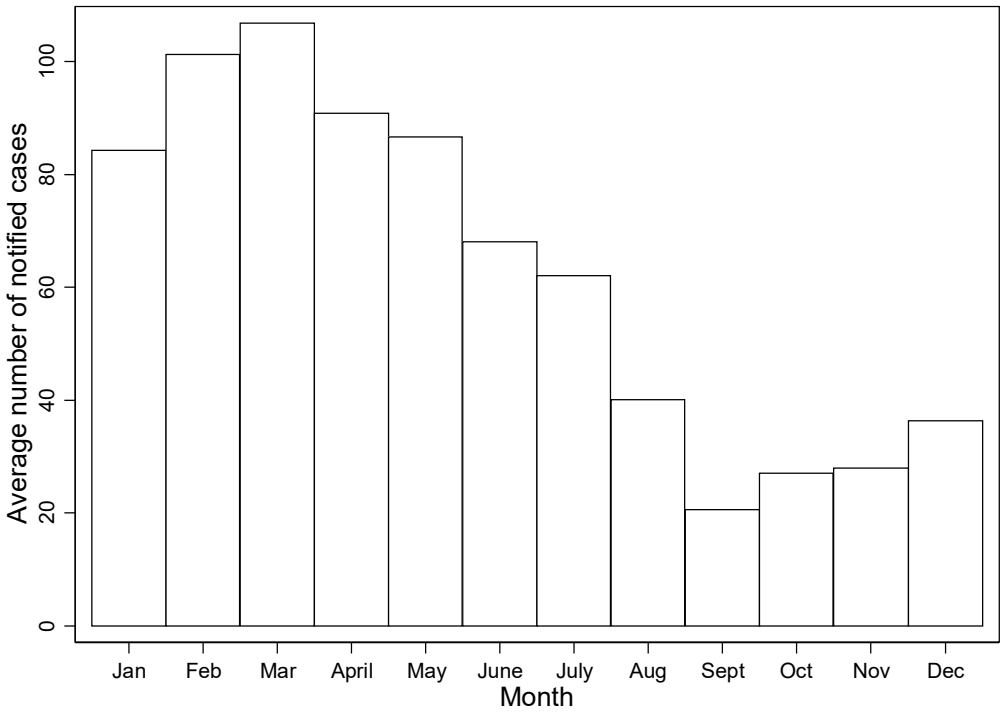


Figure 3. Average monthly dengue cases notified in Yogyakarta City (2006-2014)

2.1.4. Evidence for Zika and chikungunya transmission in Indonesia

Data on the prevalence of Zika and chikungunya (CHIK) in Yogyakarta are sparse and reflect limited availability of molecular diagnostics for these diseases in most clinical settings in Indonesia. A study from 1999 documented the presence of anti-CHIK antibodies in residents of Yogyakarta. Seroprevalence was ~30% in healthy volunteers [21], suggesting the transmission of CHIK, or of a closely-related, serologically cross-reactive alphavirus in Yogyakarta. Elsewhere in Java, independent studies have documented autochthonous CHIK transmission occurring between 2000-2011 [22–25]. Collectively these data confirm that CHIK transmission has occurred in Java, and likely continues to occur at a variable intensity.

The Zika virus has circulated in Asia for decades [26]. Zika viruses derived from Asia have been responsible for epidemics in the Western Pacific and Latin America and precipitated the WHO declaration of a public health emergency in early 2016. Although there is no data on Zika transmission in Yogyakarta, between December 2014 and April 2015 a Zika case was detected in a 27-year-old man in Jambi Province, central Sumatra, Indonesia [27]. The isolation and characterization of the Zika virus from this patient with no travel history confirms that the virus is circulating in Indonesia and that, by mimicking mild dengue infection, this infection is likely

contributing to the large number of undiagnosed cases of acute febrile illness. The assumption is supported by confirmation of Zika infection among returned travellers following exposure in Jakarta in 2013 [28] and Bali in 2014 [29]. Since the clinical manifestations of CHIK and Zika virus infections are similar to that of uncomplicated dengue, and there is a paucity of specific diagnostics tests being performed for these pathogens, it is likely these two infections are underreported in the Indonesian archipelago. There are no licensed medical interventions for Zika or chikungunya.

2.1.5. Traditional vector control strategies to control dengue, chikungunya and Zika transmission

Vector control targeted against *Ae. aegypti* is the mainstay of the fight against dengue, chikungunya and Zika disease transmission. Integrated control strategies include (i) targeted residual spraying, (ii) space spraying, (iii) larval control and, (iv) personal protection measures. The public health response to episodic dengue outbreaks in northern Australia relies upon active case finding and vector control to interrupt dengue transmission. However, successful broad-scale application of integrated vector control has been especially difficult to achieve in resource-limited endemic countries and impossible to sustain. Additionally, the evidence base to prioritise one intervention over another (e.g., larvicides and outdoor versus indoor insecticide space spraying), is weak as none have been robustly evaluated for impact on human infection and disease [30,31]. Some intervention trials have evaluated entomological impact [32], but reductions in mosquito populations do not correlate well with predictable reductions in dengue disease [33]. Collectively, the inability to rationally prioritise vector control interventions, coupled with resource limitations in endemic settings, helps explain why contemporary vector-borne disease control programs have failed to stop regular epidemics and global dispersal of dengue, chikungunya and Zika.

2.1.6. The need for a strengthened evidence base for vector control interventions

A recent meta-analysis of entomological intervention trials demonstrated the remarkable paucity of reliable evidence for the effectiveness of any vector control method on dengue incidence [34]. Strikingly, none of the randomised controlled trials (RCTs) of vector control that were included in the meta-analysis investigated epidemiological impact (i.e. clinical disease endpoint) [34]. As examples, Andersson et al 2015 and Degener et al 2014 [35,36] recently reported cluster randomised trials of vector control for dengue but neither used an objectively measured

clinical endpoint. The difficulty of making evidence-based policy in relation to vector control has resulted in calls for improved trial methods [37].

2.1.7. The World Mosquito Program approach

The World Mosquito Program (formerly Eliminate Dengue Program) is an international research collaboration that is delivering a paradigm shift in the control of arboviral diseases transmitted by *Ae. aegypti* mosquitoes. Our method utilises *Wolbachia*, obligate intracellular endosymbionts that are common in insect species [38–41] but were not present in *Ae. aegypti* mosquitoes until they were stably transinfected in the laboratory. In insects *Wolbachia* is maternally transmitted via the egg and manipulates insect reproduction to favour its own population dissemination via cytoplasmic incompatibility (CI). The result is that *Wolbachia* rapidly enter into naïve mosquito populations in a self-sustaining, durable manner. Multiple *Ae. aegypti*:*Wolbachia* combinations have been generated by the O'Neill laboratory where they form stable, maternally-transmitted infections that cause CI.[42–44] Strikingly, the presence of *Wolbachia* in *Ae. aegypti* mosquitoes renders them more resistant to disseminated arbovirus infection, including dengue, Zika, chikunguyna and Yellow fever viruses [45–47]. Thus the critical and signature effect of *Wolbachia* as a public health intervention is to severely reduce the vectorial capacity of mosquito populations to transmit arboviral infections between humans. For field implementation, the approach works by seeding wild mosquito populations with *Wolbachia* through controlled releases of relatively small numbers of *Wolbachia* infected mosquitoes (Figure 4). Over several months, and through the actions of CI, the prevalence of *Wolbachia* in the local mosquito population increases, until such time as the majority of mosquitoes in the area carry *Wolbachia*.

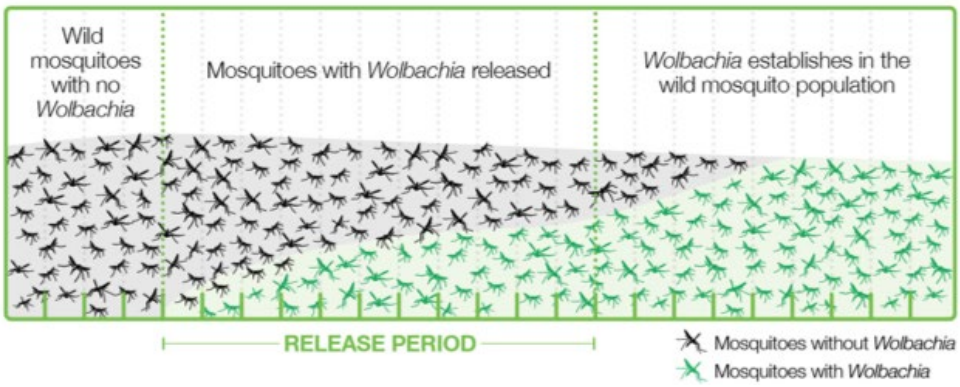


Figure 4. The *Wolbachia* biocontrol method. *Ae. aegypti* mosquitoes with *Wolbachia* (green) are released into the wild mosquito population (black).

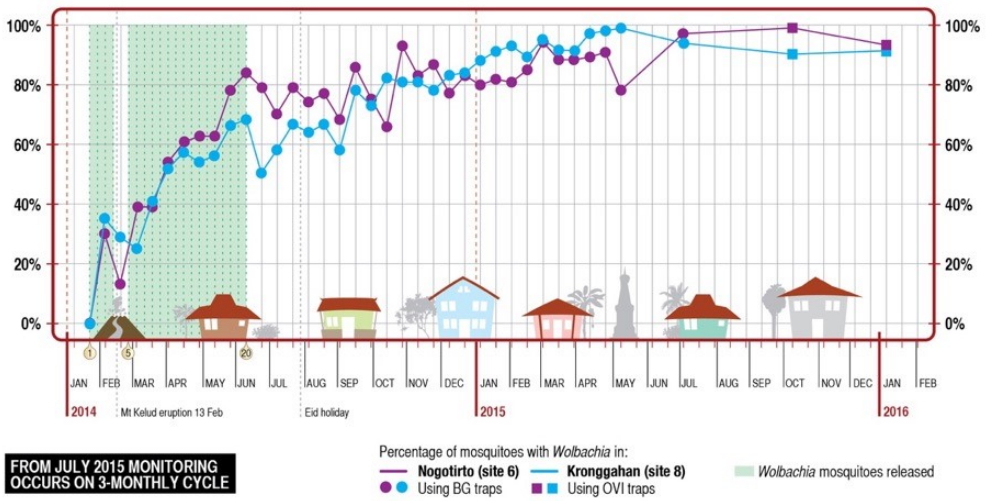
Over a series of releases, the percentage of *Wolbachia* mosquitoes increases. Once a threshold frequency of *Wolbachia* mosquitoes is reached, *Wolbachia* will continue to spread after releases have finished until the majority of mosquitoes carry *Wolbachia*. Laboratory vector competence studies show that *Wolbachia*-infected mosquitoes have a significantly reduced ability to transmit dengue, Zika and chikungunya viruses. Our Program has undertaken extensive vector competence assessments to determine the effect of *Wolbachia* (*w*Mel strain) on DENV infection and dissemination in *Ae. aegypti* mosquitoes [48–51]. Of note, we have elegantly demonstrated reduced vector competence in *Wolbachia*-infected mosquitoes obtained from the field using human dengue viremic blood and a novel read-out to measure infectious mosquito saliva [52]. *w*Mel viral interference effects were found to impact all four DENV serotypes, resulting in predicted reductions of 66-75% in the basic reproduction number R_0 for DENV-1-4 [52]. Reductions of this magnitude are predicted to result in local elimination of DENV transmission in most epidemiological circumstances [52].

The reduction in mosquito vector competence imparted by *Wolbachia,* together with the ability of *Wolbachia* (*w*Mel strain) to establish itself in *Ae. aegypti* populations, has led to regulatory and community acceptance of the technology in five countries; Australia (Cairns, Townsville), Indonesia (Yogyakarta), Vietnam (Nha Trang), Colombia (Medellin) and Brazil (Rio de Janeiro). Currently, approximately 160,000 people live under the protective umbrella of *w*Mel deployments. In 2016, in response to the emergence of Zika virus, the WHO endorsed pilot deployments of the *Wolbachia*-based biocontrol method to combat arboviral diseases [16].

2.1.8. Previous *Wolbachia* releases in Yogyakarta

Small-scale proof-of-concept field trials of *Wolbachia* (*w*Mel) deployment have been conducted in four small communities in districts adjacent to Yogyakarta City since 2014, with releases beginning in January 2014 in two sites in Sleman district and in November 2014 in two sites in Bantul district. In all sites, *Wolbachia* achieved a high prevalence in field-caught mosquitoes following the completion of releases, which has since been sustained (Figure 5). One year after *Wolbachia* establishment in Sleman (2015), *w*Mel-infected *Ae. aegypti* were collected from Nogotirto and Kronggahan field sites and injected with the four serotypes of dengue (isolated from East Timor) to look at dengue replication-blockage phenotype mediated by *w*Mel in wild-type *Ae. aegypti.* The data obtained show continued strong blockage of dengue 1-4 replication by *w*Mel.

A.



B.

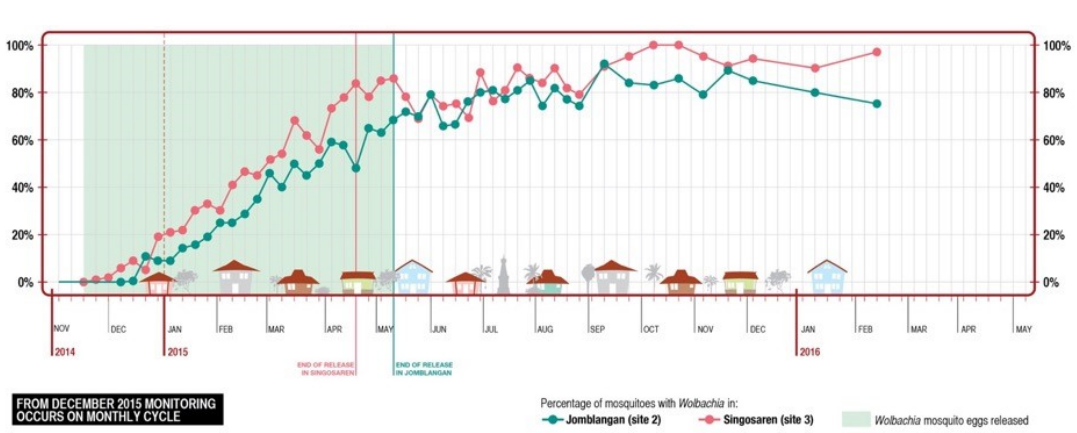


Figure 5. Establishment of *Wolbachi*a in A) Sleman and B) Bantul districts, Yogyakarta Province

2.2. Research question

Does large-scale deployment of *Wolbachia*-infected *Ae. aegypti* mosquitoes lead to a measurable reduction in dengue incidence in people living in release areas, compared to those living outside release areas?

2.3. Rationale

The successful introduction of the intracellular bacterium *Wolbachia* into *Ae. aegypti* mosquito populations is predicted to provide a long-term and sustainable approach to reducing dengue transmission. A critical next step, and the aim of the study outlined in this protocol, is to measure experimentally the efficacy of *Wolbachia* in reducing dengue virus transmission in the field. To this end, cluster randomised trials (CRTs) are the gold standard design to provide

evidence on the efficacy of an intervention that has a community-wide impact [53]. The value of providing estimates of the epidemiological impact of *Wolbachia* from a randomised controlled trial are threefold. First, randomised controlled trials are the gold standard and preferred method for estimating the benefit or harm of health interventions. Second, results of CRTs are usually more influential in shaping policy and practice in medicine and public health than observational studies and this is likely to be true for the *Wolbachia* technology. Third, as noted previously, limitations in the design of most previous vector control trials for dengue mean that the scientific community expects that the *Wolbachia* technology be trialled using gold-standard methods wherever feasible [16,37].

3. Study objectives

3.1. Primary objective

To assess the efficacy of community-based deployments of *Wolbachia*-infected *Ae. aegypti* mosquitoes in reducing the incidence of symptomatic, virologically-confirmed dengue cases of any severity in Yogyakarta residents aged 3-45 years in release areas, relative to non-release areas.

3.2. Secondary objectives

- To measure the efficacy of the *Wolbachia* method against each of the four DENV serotypes.
- To measure the efficacy of the *Wolbachia* method in reducing the incidence of symptomatic virologically-confirmed Zika virus and chikungunya virus infection in release areas, relative to non-release areas, and
- To quantify the impact of *Wolbachia* deployments on notifications of dengue haemorrhagic fever (DHF) cases to the Yogyakarta district health office
- To quantify the level of human mobility within Yogyakarta City, and estimate the proportion of residents' exposure time that they spend outside the treatment arm to which they were randomised.
- To determine whether community-based deployment of *Wolbachia*-infected *Ae. aegypti* mosquitoes reduces the abundance of wild-type *A. aegypti* adults, or

alternatively, alters the abundance of adults from *Aedes* species other than *A. aegypti* (e.g. *Ae. albopictus*)

4. Study design

4.1. Type of study

This is a parallel two-arm non-blinded cluster randomised controlled trial which will be conducted in a single site in Yogyakarta City, Indonesia. The study site will be subdivided into twenty-four contiguous clusters, approximately 1km^2 in size (range 0.7km^2-1.65km^2), Figure 6. Clusters will be randomly allocated in a 1-to-1 ratio to receive *Wolbachia* deployments or no intervention, such that 12 clusters receive *Wolbachia* deployments and 12 receive no intervention.

There will be no buffer areas between clusters, but natural borders (roads, rivers, non-residential areas) will be used to define cluster boundaries as much as possible, to limit the spatial spread of *Wolbachia* from treated clusters into untreated areas, and of wild-type mosquitoes in *Wolbachia* treated clusters. Exclusion areas will be minimised, but any areas within the study site where releases are not possible for reasons of logistics, public acceptance or absence of mosquito populations (e.g. hospitals, public space, open parkland) will be pre-specified prior to randomisation and balanced between study arms. No attempt will be made to alter the routine dengue prevention and vector control activities conducted by public and private agencies throughout the study area (treated and untreated clusters). It is worth noting the capacity of the disease surveillance system to detect (and thus respond to) dengue will be enhanced across the city through increased availability of diagnostic kits, which have been supplied to primary care clinics and hospitals since March 2016 by the Eliminate Dengue Project Indonesia, to support efforts to enhance the surveillance of dengue across Yogyakarta.

The impact of *Wolbachia* deployments on dengue incidence will be assessed by comparing the exposure distribution (probability of living in a *Wolbachia*-treated area) among virologically-confirmed dengue cases presenting to a network of public primary clinics (Puskesmas), against the exposure distribution among patients with febrile illness of non-arboviral aetiology presenting to the same network of clinics in the same temporal windows.

Dengue cases and arbovirus-negative controls will be sampled concurrently from within the population of patients presenting with febrile illness to the study clinic network, with case or control status classified retrospectively based on the results of laboratory diagnostic testing. The dataset for analysis will retain all enrolled cases and all controls that are matched to a case by calendar month of illness onset. Unmatched controls will not be used for the primary analysis.

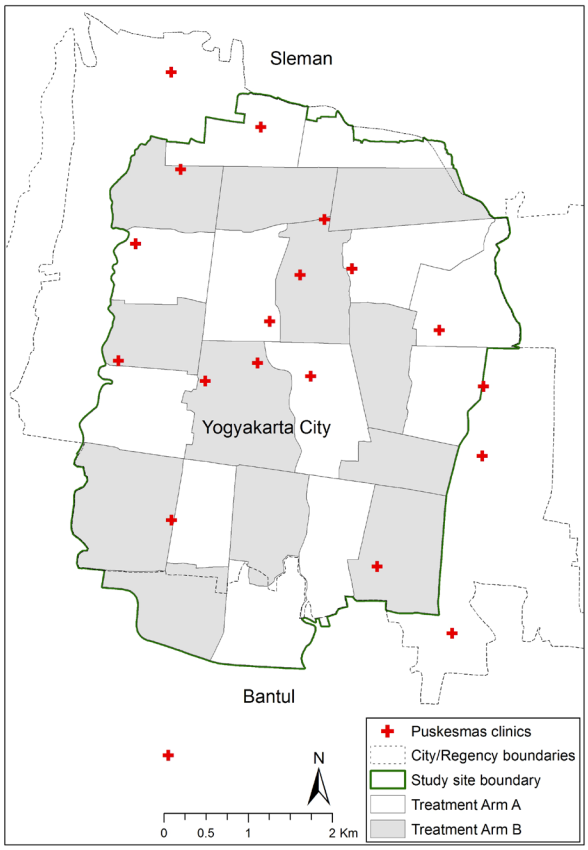


Figure 6. Map of study area, cluster boundaries, and Puskesmas clinics. The study area is outlined in green. The 12 clusters in each treatment arm are shown in grey and white. The location of the puskesmas clinics at which trial recruitment is conducted are shown by red crosses

The distribution of *Wolbachia* exposure in the sampled arbovirus negative controls will reflect the distribution of *Wolbachia* exposure in the underlying source population that gave rise to cases, as long as a core assumption is met that the relative propensity to seek healthcare for febrile illness at a Puskesmas in intervention versus untreated arms is the same for dengue cases as other febrile illness controls. This should be upheld if dengue cases and other febrile illness controls are clinically indistinguishable until laboratory diagnosis. The concurrent

sampling of controls and cases means that the odds of *Wolbachia*-exposure among sampled dengue cases relative to febrile controls (odds ratio), is an unbiased estimate of the relative incidence of medically-attended dengue in *Wolbachia*-treated versus untreated clusters (relative risk or incidence rate ratio), from which protective efficacy can be estimated directly.

The null hypothesis is that the relative risk of virologically-confirmed dengue in *Wolbachia*-treated and untreated areas is one. If *Wolbachia* has a protective effect against DENV transmission, we would expect the relative risk or incidence rate ratio for virologically-confirmed dengue in *Wolbachia*-treated areas compared to untreated areas to be below one.

4.2. Justification of study design

Cluster randomised trials are the gold standard design to provide evidence on the efficacy of an intervention that has a community-wide impact[54]. This design involves the random allocation of intervention or no intervention to communities ('clusters') in the study area. Traditionally, an endpoint of infection and/or disease incidence would be measured in a prospective cohort of participants in *Wolbachia* treated and untreated clusters, and the measure of effect is the risk or rate ratio between intervention and untreated arms. Protective efficacy is estimated by one minus the risk/rate ratio. Given that DENV infection and/or disease in children living in an endemic area might be conservatively estimated at 5-10% and 1-5% per annum respectively[55], this would necessitate a large cohort of several thousand children, followed for several years, to detect a few hundred DENV infections each year.

An alternative method for measuring efficacy in a cluster randomised trial of *Wolbachia,* with a clinical endpoint, is to use concurrent sampling of dengue cases and non-dengue controls from within the whole study population to derive the estimate of the relative risk and/or incidence rate ratio. In this method of concurrent sampling of controls – also called incidence density or risk-set sampling – the control group represents a sample of the underlying dynamic population from which the cases arose, with an exposure distribution that is temporally matched to the timing of case onset [56]. This takes account of the time-varying nature of the exposure distribution (*Wolbachia* prevalence in local *Ae. aegypti* mosquitoes) and the seasonality of dengue. By recruiting participants from within the population of patients presenting to clinics with febrile illness – with dengue test-positive patients classified as cases and test-negative patients classified as controls – the controls are necessarily drawn from the

same source population as the cases, thus avoiding the common pitfalls of control selection in traditional case-control studies that can introduce selection bias [57]. In this situation, the odds ratio is an unbiased estimate of the rate ratio in the source population over the period of participant enrolment (the 'risk' period), without the need for any rare disease assumption [56,58]. This approach to measuring the efficacy endpoint in a CRT has the advantage of being more efficient, cost effective, and logistically simpler to achieve than a large prospective cohort of children.

This study design has a precedent in the 'test negative design' (TND) used for evaluating the effectiveness of seasonal influenza vaccination. In that design, patients seeking health care for an acute respiratory illness (ARI) are recruited into the study and tested for influenza. Influenza vaccine effectiveness is then estimated as one minus the ratio of the odds of vaccination in subjects testing positive for influenza to the odds of vaccination in subjects testing negative [59]. Several authors have explored the statistical rationale and underlying assumptions of this design, and have demonstrated that the odds ratio for vaccination in influenza cases vs test-negative controls is directly equivalent to the relative risk of influenza in vaccinated vs unvaccinated individuals *if* test-negative controls are allowed to include participants who may test positive for influenza at any other time during the study period (i.e. risk-set sampling) [57] and if the distribution of non-influenza ARI is not associated with the intervention status [59]. The design outlined in this protocol extends the TND approach by including concurrent sampling of test-positive cases and test-negative controls, such that the odds ratio will approximate the rate ratio, rather than the risk ratio.

The test-negative design allows for differences in health care seeking behavior between vaccinated and unvaccinated individuals (i.e. *Wolbachia*-exposed and unexposed populations in our study) – e.g. due to spatial variation in the preferences for attending government vs private clinics - as long as the relative propensity between exposed and unexposed populations to seek care (and be enrolled) at a participating clinic is the same for test-positive and test-negative patients[60]. This should be the case if test-positive and test-negative patients are clinically indistinguishable, and only classified after enrollment on the basis of subsequent laboratory diagnostic testing. The internal validity of the TND depends primarily upon the avoidance of selection bias in the sampling of cases and controls, and the extent to which controls can be assumed to be representative of the source population that gave rise to cases.

A core assumption of the TND, translated to our study design, is that among persons who would seek care for febrile illness (at a participating clinic), the incidence of arbovirus-negative febrile illness does not differ, on average, between *Wolbachia*-treated and untreated areas [57,59,60]. If this is upheld, then the sampled controls will represent an unbiased estimate of the exposure distribution (i.e. residence in a *Wolbachia*-treated area) in the source population.

Our design introduces some advantages over the TND for influenza vaccine effectiveness, principally that the allocation of *Wolbachia* deployments is randomised. This helps to ensure that the intervention and non-intervention arms are balanced with respect to confounding variables, so that the study arms have similar baseline dengue risk and any measured difference in dengue incidence during the study period can be attributed to the effects of *Wolbachia*. This also means that, whereas the TND as applied to influenza vaccination is an observational study and can estimate only vaccine *effectiveness* under field conditions from the proportionate reduction of risk, ours is an experimental design and the proportionate reduction of risk gives an estimate of protective *efficacy* of *Wolbachia* [57]. A further difference is that *Wolbachia* is a community-level intervention, unlike influenza vaccination which is delivered to the individual; this introduces additional complexity into the analysis approach. However, the methodological foundations of the TND, and many of the assumptions on which the statistical inference is based, translate well to the study design and analysis approach detailed in this protocol.

4.3. Number of participants

The study area as a whole has a population of approximately 350,000, of which approximately half will be resident in areas randomised to receive *Wolbachia* deployments and half in untreated areas.

The study population for measurement of the efficacy endpoint is the population of patients resident in the study area, presenting to the network of participating health clinics with febrile illness, and meeting the eligibility criteria as described in section 7. Based on two years of historic data collated from the network of health clinics (Puskesmas) in the study area, it is estimated that at least 5000 patients per year present to these clinics with febrile illness (range 200-1500 per clinic per annum).

We will enroll all participants presenting to any of the participating clinics who meet the eligibility criteria as described in section 7. Following laboratory testing and classification of participants' diagnostic status, all cases and those controls enrolled within the same calendar month as any case will be retained in the dataset for analysis. Enrolment will continue for up to three years to attain a sufficient sample size for intention-to-treat analysis, as described in 11.1. Recruitment will continue for up to 36 months, unless early termination is recommended by the independent data monitoring committee (IDMC; see section 13.5).

4.4. Expected duration of study

Wolbachia deployments will commence in March 2017 and will continue for approximately seven to eight months. The clinic-based sampling of febrile patients is expected to commence in pilot phase in September 2017, with active enrolment in all clinics by December 2017. The dataset to be included in the primary 'intention-to-treat' analysis will include only participants enrolled after *Wolbachia* is considered established in treated clusters, defined as one month after completion of releases in the last cluster. The study timeline is depicted in Figure 7.

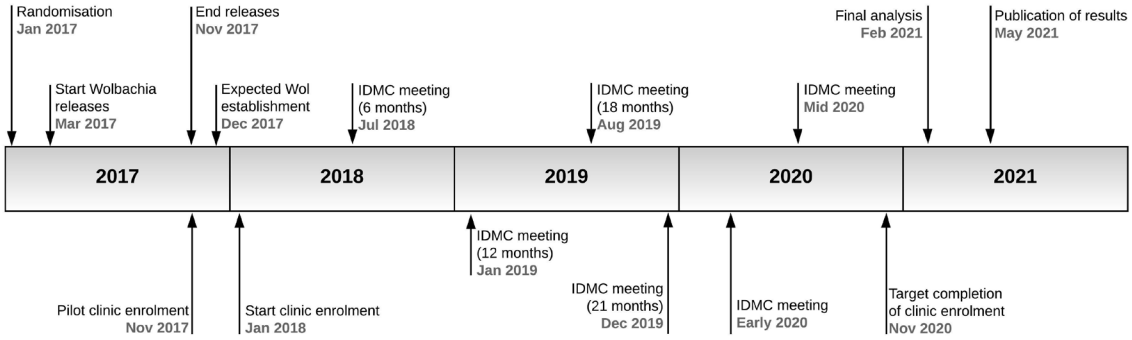


Figure 7. Study time line. Wol: *Wolbachia;* IDMC: Independent Data Monitoring Committee

4.5. Primary and secondary outcome measures

4.5.1. Primary outcome: dengue

The primary outcome measure will be virologically confirmed dengue virus infection in patients reporting febrile illness. Participants will be classified as dengue cases for the primary analysis if plasma samples collected 1-4 days after onset of fever test positive for dengue virus NS1 antigen (BioRad Platelia NS1 ELISA) and/or dengue virus nucleic acid by RT-qPCR. The infecting serotype will be determined by DENV serotype-specific RT-PCR, and participants

with a known serotype will be included in a secondary analysis to estimate serotype-specific efficacy, as described in section 11.3.1.

4.5.2. Secondary outcomes: chikungunya and Zika

Secondary outcome measures include chikungunya and Zika virus infection in patients reporting febrile illness. Participants will be classified as virologically-confirmed chikungunya cases if chikungunya nucleic acid is detected in plasma samples by RT-qPCR. Participants will be classified as virologically-confirmed Zika virus cases if Zika virus nucleic acid is detected in plasma samples by RT-qPCR. A reduction in chikungunya and/or Zika virus infections association with *Wolbachia* deployment will be measured in the same way as for dengue, as described in section 11.3.2.

5. Study setting

The study will be conducted in Yogyakarta City and Bantul District, both located in the province of Yogyakarta Special Region, Indonesia. Yogyakarta City is 32 km^2 in size and had a population of 408,000 in 2015. The study site is 26 km^2 in size, including 24 km^2 within Yogyakarta City, and extending into 2km^2 of the adjacent administrative area, Bantul District, to the south of Yogyakarta City (see Figure 6). The study site is a continuous urban area, with a total population of approximately 350,000 and an average population density of 13,460 persons per km^2. The annual dengue incidence rate in Yogyakarta ranged between 83-390 cases per 100,000 population during the years 2006-2014. Even though dengue cases are reported all year, the high dengue season usually begins in December, peaks in March, and tapers off between June and July.

The study site will be subdivided into twenty-four contiguous clusters, each approximately 1km^2 in size (range 0.7km^2-1.65km^2). Among the 24 clusters, 12 will be randomised to receive *Wolbachia* deployments and 12 will be untreated. Participant enrolment to measure the efficacy endpoint will be conducted at a network of health clinics throughout the study site.

6. Study intervention

6.1. Randomisation method

Constrained randomisation will be used to prevent a chance imbalance in the baseline characteristics or spatial distribution of treated and untreated clusters.

Covariate constrained (also referred to as "restricted") randomisation ensures balance and minimises loss of statistical power without the need for large numbers of strata, and is an excellent method to achieve balance when the number of clusters is small [54,61,62].

The approach to randomisation follows the method outlined by Hayes and Moulton [54]. First, all potential allocations of twelve intervention and twelve non-intervention clusters will be identified amongst the twenty-four clusters. Next, each allocation will be assessed against pre-defined balance criteria. All potential allocations that satisfy these balancing criteria will be retained, and non-balanced possibilities rejected. Finally, a single allocation pattern will be randomly selected from within the restricted list of balanced possibilities.

Traditionally, balancing variables include those that may be potentially confounding covariates; may impact sample size; or are relevant for logistics [54]. To avoid bias the TND approach also requires there is no association between the probability of other febrile illness (OFI) and the intervention [60]. To account for this given the propensity for spatial clustering of OFIs, randomisation will also balance on numbers of patients presenting to Puskesmas with undifferentiated fever during the preceding two years.

Sample size includes both the number of clusters and the total number of sampled individuals within each treatment arm [62], and is included because precision and power is maximised when sample sizes in treatment arms are similar. Balancing covariates are listed in Table 1.

Table 1. Balancing covariates

Reason for balance	**Balancing covariates**
Potential confounders	1. Age - % under 15 years 2. Average dengue incidence rate over most recent 3 years 3. Education - % completed high school, as a proxy for socioeconomic status
Potential sources of bias	4. Incidence of other febrile illness presenting to Puskesmas in 2014-15
Sample size	5. Number of clusters 6. Cluster population
Logistics	7. Total cluster area (km^2) 8. Total size of non-release area in cluster 9. Four spatial strata (to minimise number of contiguous areas)

6.2. *Wolbachia* deployment strategy

Wolbachia-infected *Ae. aegypti* will be deployed by setting Mosquito Release Containers (MRCs) in residential and non-residential properties throughout the intervention clusters. An MRC is a small plastic tub containing approximately 80 *Ae. aegypti* eggs, Tetramin food and water. Adult mosquitoes emerge from small holes in the side of the MRC, approximately 7-12 days after the MRC is deployed.

6.2.1. Density and duration of release

Clusters have been mapped and overlaid with 50-meter grids. Up to five properties in each grid square in the clusters allocated to *Wolbachia* deployments will host an MRC. Entomology field staff will set up the MRCs in an outdoor area, by filling the container with water and adding the eggs and the Tetramin food to the container. Each MRC will be serviced every two weeks by adding a new batch of eggs, fresh food and water, at which point the previous release event will be classified as OK or Fail. Criteria for a fail include the container tipping over or missing, or presence of predators in the container.

As a quality assurance, a subset of MRCs will be checked and the number of emerged adult mosquitoes will be recorded based on the number of pupal skins and dead adults per container. The average emergence rate in each cluster will be used as a multiplier together with the number of eggs per MRC to estimate the number of adult mosquitoes released per cluster. Based on previous field work, the expected emergence rate is approximately 60%, and so a 50-meter grid network of MRCs (400-2000 per km^2) with 100-150 eggs per MRC is expected to equate to 30,000 – 150,000 adults released per km^2, in each release week.

Eight fortnightly releases per cluster are planned initially (total duration of 16 weeks).
If *Wolbachia* prevalence in trapped mosquitoes is ≤60% in a specific cluster at the end of the release, the release will continue in that particular cluster until the threshold of 60% is reached.

Once *Wolbachia* prevalence has reached ≥60% in a cluster and releases have stopped, there will be no remediation with additional releases if *Wolbachia* prevalence drops below 60% in the future.

6.2.2. Timeline for completion of releases

Wolbachia will be deployed through rolling releases across treatment clusters within a 6 to 9-month period, with the aim of achieving *Wolbachia* establishment throughout treatment clusters within twelve months (from the start of the release). Deployments will continue in a cluster until the cluster-level *Wolbachia* prevalence in trapped *Ae. aegypti* reaches a pre-defined threshold of 60%. Deployment will then stop in that cluster and monitoring of *Wolbachia* prevalence in trapped mosquitoes will continue throughout the study period.

6.2.3. Handling individual and community-level refusal to release

Permission will be sought from community leaders (heads of Kelurahan administrative areas) prior to randomisation. If permission is not granted for a given Kelurahan or part of a Kelurahan, that area will be excluded when drawing cluster boundaries, and therefore from the study area for randomisation. Residents of these excluded areas who present to a study clinic will not be enrolled into the study, as they will not meet the inclusion criterion of residence within the study area.

Where individual householders refuse to host a mosquito release container, field staff will not release at the individual's home. Another release location within the same 50m^2 grid will be sought instead. Individuals from this household will still be eligible to participate in the clinic-based efficacy study.

6.3. *Wolbachia* monitoring strategy

6.3.1. Trapping method and density: during and post-deployment

A network of BG-Sentinel adult mosquito traps (BioGents) will be established throughout intervention clusters prior to the commencement of releases, evenly spaced throughout residential areas at a density of approximately 16 traps/km^2. A BG trap network of the same density (16 traps/ km^2) will be established also in untreated clusters prior to the commencement of the clinical study. BG traps will be serviced weekly, with trapped mosquitoes screened for *Wolbachia* at weekly, fortnightly or monthly intervals throughout the duration of the trial, depending on the stage of release and establishment. Mosquitoes will be bio-banked in the intervening weeks when screening is not done.

6.3.2. Laboratory methods for mosquito ID and screening

Trapped mosquitoes will be identified using microscopy, based on morphological criteria that allow differentiation of adult *Ae. aegypti* from other mosquito species present in Yogyakarta.

Ae. aegypti collected from a single BG trap will be preserved together, but separated by female and male, in tubes containing 80% Ethanol.

After identification, samples will be sent to the diagnostic laboratory and individual mosquitoes (male and female) will be homogenised in a buffer solution to extract DNA and screened using quantitative PCR assay to detect the presence of *Wolbachia* and to confirm the species as *Ae. aegypti*. For each tube tested, corresponding to male or female *Ae. aegypti* from a single BG trap, the data recorded will include the number tested, the number positive by *Ae. aegypti* PCR, and the number positive by *Wolbachia* PCR.

6.3.3. Definition of establishment

Establishment is defined by ≥80% *Wolbachia* prevalence in trapped *Ae. aegypti* (aggregated across all traps in the cluster), for two consecutive screening events. For the purposes of measuring the efficacy endpoint in the primary intention-to-treat analysis, *Wolbachia* will be considered established throughout intervention clusters one month after completing releases in the last cluster.

7. Selection and enrolment of participants

A lag period between the *Wolbachia* releases and the start of clinical surveillance is planned to provide sufficient time for *Wolbachia* to establish in the wild *Ae. aegypti* population. Study processes for enrolling patients presenting with febrile illness will be established at a network of primary care clinics (Puskesmas) throughout the study area. The clinic-based enrolment will operate in a pilot capacity from approximately September 2017, with a staged implementation across study clinics. The pilot period will be considered complete when study processes have been successfully implemented in all clinics. Recruitment will be continuous, with the dataset for the secondary 'per-protocol' analysis including all participants enrolled following the completion of the pilot period. The dataset for the primary 'intention-to-treat' analysis will include only participants enrolled after *Wolbachia* is considered established in the treated clusters (defined as one month after completion of releases in the last cluster). Participants will be enrolled from within the population of patients (aged between 3-45 years old) presenting with undifferentiated fever of 1-4 days duration. All patients meeting the below inclusion criteria and providing written informed consent will be eligible for enrolment. Recruitment will continue for up to 36 months, at which point a sufficient sample size for

intention-to-treat analysis is expected to have been achieved. Recruitment will continue until November 2020, even if an adequate minimum sample size is reached before this time, unless early termination is recommended by the independent data monitoring committee (IDMC; see section 13.5.

7.1. Recruitment procedures

All eligible participants meeting study inclusion criteria will be invited to enroll continuously throughout the study period. Recruitment will occur during normal clinic hours. Recruitment rates in each clinic and across the study site as a whole will be monitored monthly, including review of the screening logs to identify the proportion of eligible participants who did not consent to participate. The field coordinator will make regular visits to low-enrolling clinics to identify clinic-based, patient-based or other causes for low recruitment, and put measures in place to address these.

7.1.1 Screening log book

All patients presenting with febrile illness will be screened against the study inclusion criteria by trained staff. All eligible febrile individuals will be recorded in a screening log and invited to participate. Participation status (consent/decline) will be recorded against each participant in the log.

7.2. Informed consent procedures

Written informed consent will be sought from participants or their g uardian (parents or vertical guardian) where the participant is a minor by trained local staff, after explaining the study objectives, processes, data and sample collection and the participant has had an opportunity to ask questions. A verbal explanation of the written Explanatory Statement will be provided to all participants in the local language. In addition, participants aged between 13 and 17 years will be invited to sign an assent form indicating they understand the research and agree to participate. The documents used for participant recruitment will be in Bahasa Indonesia and all recruitment procedures will be conducted in the local language.

7.3. Inclusion criteria

Participants must meet the following inclusion criteria:

i) Fever (either self-reported or objectively measured, e.g. (tympanic membrane temperature ≥38^{o}C)) with a date of onset between 1-4 days prior to the day of presentation.

ii) Aged between 3-45 years old.

iii) Resided in the study area every night for the 10 days preceding illness onset.

7.4. Exclusion criteria

Participants will not be eligible for inclusion if any of the following are identified:

i) Localising features suggestive of a specific diagnosis other than an arboviral infection, e.g. severe diarrhea, otitis, pneumonia

ii) Prior enrollment in the study within the previous 4 weeks.

An individual presenting to the clinic on repeat occasions for different febrile episodes will be eligible for enrollment during each different episode. However, an individual may only be enrolled once during a single illness episode, which we define as illness occurring within 4 weeks of a previous febrile episode.

8. Data and sample collection procedures

8.1. Data to be collected

A unique identifier will be assigned to each participant at enrollment. Basic demographic details, eligibility against the inclusion criteria and illness onset date will be recorded in a standardised case report form. Table 2 summarises the data and samples to be collected from each participant. Data and samples are collected at a single time point at enrolment, with no longitudinal follow up of participants except for a phone call to establish their status at 14 to 21 days post-enrolment.

Table 2: Summary of data and samples collected

Data/sample type	**Purpose**
Demographic data (e.g. name, sex, date of birth, address, contact phone number)	To uniquely identify participants; describe demographic characteristics of study

	population; ascertain study eligibility; permit follow-up at 14 to 21-days post-enrolment.
Dengue vaccination	To determine proportion of participants who have received the dengue vaccine.
Illness history data (e.g. symptoms, date of onset)	To ascertain study eligibility
Travel history in past 10 days (e.g. home and other places visited, including durations and geolocations)	To determine proportion of time spent in *Wolbachia*-treated and untreated clusters, for per-protocol analysis
3 ml venous blood sample	For DENV, chikungunya virus (CHIKV), Zika virus (ZIKV) diagnostic testing, in order to classify case/control status

8.1.1. Travel history

A brief travel history interview will be conducted at enrolment to determine the main places visited by each participant within the 10 days prior to illness onset, i.e. the incubation period for dengue. Thus, travel data are collected retrospectively. These data will be used to determine the proportion of time spent in *Wolbachia*-treated and untreated areas, for the per-protocol analysis. Because laboratory diagnostics are done retrospectively, interviewers will be blinded to the case/control status of the participant at the time of collection, which will avoid interviewer bias during collection of travel histories. However, potential bias in reporting of travel history between participants living in *Wolbachia*-treated areas and those living in untreated areas cannot be excluded, and will be minimised by the use of trained interviewers and standardised interview methods for eliciting travel histories.

8.1.2. Geolocation of participants residence and visited locations

The address of participants' residence and other locations visited during the 10 days prior to illness onset will be recorded during the travel history interview as above. The coordinates of locations visited will be verified by geo-locating on a map, and these geolocations retained for per-protocol analysis.

8.2. Data handling and record keeping

8.2.1. Roles and responsibilities of clinic staff and study staff

A study nurse will be stationed at each Puskesmas to ensure consistency in screening, recruitment and consent, data collection, sampling and transfer of specimens to laboratory. Clinical management and diagnostic testing will be provided by health center staff in accordance with standard of care. The field trial coordinator will oversee study processes in all participating Puskesmas clinics, to ensure adherence to the study protocol and standard operating procedures with respect to inclusion and exclusion criteria, informed consent procedures, case report form completion and the handling of samples and data.

8.3. Clinical sampling procedures

A single 3 ml venous blood will be collected from all consenting participants, on the day of enrolment. Topical anaesthetic will be available for use if required, applied locally to the skin at the planned site of venipuncture according to the manufacturer's recommendations and the local standard of care. Blood samples from all participants will be transferred to the project laboratory on the day of collection and batch-tested within one month to determine case or control status (see Figure 8).

In the situation where a consenting participant has already had blood collected for clinical investigations on the day of enrolment, a second blood sample will not be collected for research purposes. Rather, the residual blood sample will be retained and used for the study investigations. In this situation, the sample volume may be less than the usual 3ml collected from study participants, and consent to participate in the trial may be obtained after the clinical blood sampling has occurred (but before the residual blood sample is retrieved from the clinic laboratory).

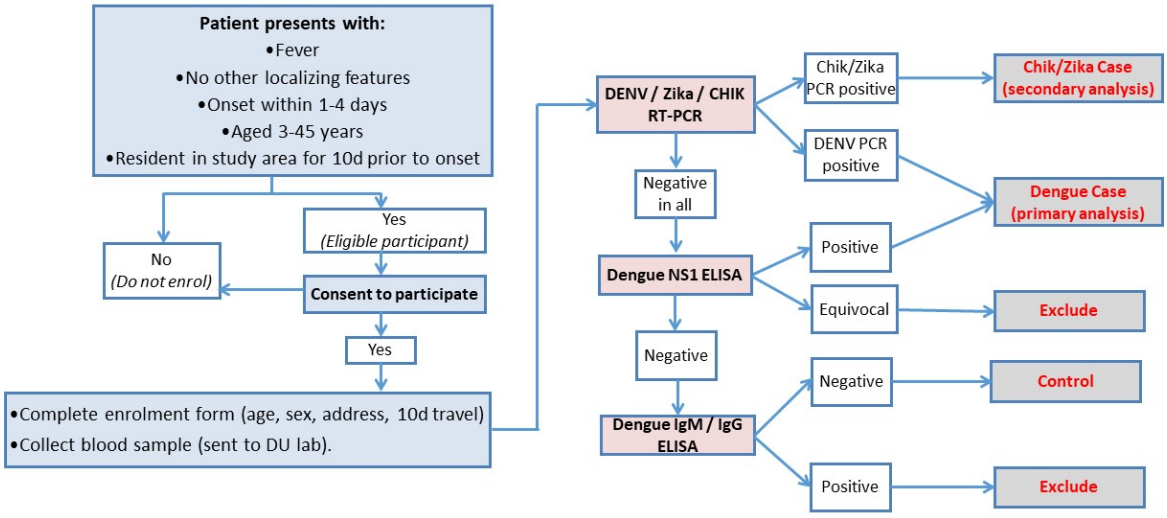


Figure 8. Flowchart of data and sample collection procedures and diagnostic algorithm

9. Laboratory assessments

9.1. Diagnostic testing for dengue, chikungunya and Zika

RT-qPCR is the gold standard method of diagnosing arboviral infections in the first few days of illness. We will use an internally controlled triplex RT-qPCR assay to detect DENV, CHIK and Zika viruses in plasma samples from all enrolled participants. Dengue NS1 Platelia ELISA (BioRad) and IgM and IgG capture ELISA (Panbio, Australia) will be performed according to the manufacturers' instructions. Samples that return a test result of "equivocal" in serology or NS1 testing will be reported as "equivocal" without retesting.

All samples positive for DENV in the triplex RT-qPCR will be tested in a serotype-specific RT-qPCR to determine the infecting serotype.

All research diagnostic investigations will be performed by the Eliminate Dengue Project diagnostics reference laboratory at the Universitas Gadjah Mada. External quality assurance panels will be used to monitor the performance of the molecular diagnostic tests.

9.2. Batch testing procedures

Diagnostic specimens will be tested in batch fashion in such a way as to maximise the throughput and minimise the cost of testing.

9.3. Sample handling and storage procedures

Clinical specimens will be collected and transferred to the reference laboratory according to standard operating procedures. All diagnostic specimens will be processed and stored on the same day as sample receipt and plasma stored at minus 80°C. Consent will be sought from participants to store their blood sample for potential use in the future to study mosquito-related infectious diseases in Indonesia. The sample will be stored for a minimum of 5 years. Sample use for future research will only be conducted after the Human Research Ethics Committee approved it and aimed only for scientific purposes.

9.4. Reporting of results

Diagnostic test results will not be reported back to individual participants since the testing will be performed in a research laboratory, not a certified diagnostic laboratory, and the batch processing of samples will mean that results are not available in time to inform clinical management. Participants will be managed according to standard clinical practice by the treating clinicians.

Given the potential risk of congenital Zika virus syndrome in a developing fetus exposed to Zika virus, we will report back to the primary care clinic a line listing of participants with positive results in Zika virus PCR at least once per month, so that standard procedures for follow up of patients at risk of Zika virus infection can be followed at the clinician's discretion. The remaining blood specimen will be made available for forwarding to a diagnostic laboratory if the clinician requests it.

9.5. Case/control classification algorithm

Dengue cases are defined as patients with virologically-confirmed DENV infection, meeting the clinical criteria for enrolment and also with a positive result in NS1 ELISA and/or DENV RT-qPCR.

Controls are patients meeting the clinical criteria for enrolment, but with negative test results for CHIK RT-qPCR, Zika RT-qPCR, DENV NS1 ELISA, DENV RT-qPCR and DENV IgM and IgG ELISA (see Figure 8).

For the secondary endpoints, participants meeting the definition of a dengue case above, and with a known infecting DENV serotype, will be included in the secondary analysis of serotype-

specific efficacy. Zika or chikungunya cases are defined as patients with virologically confirmed Zika or chikungunya infections, meeting the clinical criteria for enrolment and also with a positive result in Zika RT-qPCR or CHIK RT-qPCR, respectively, and controls are defined as above.

10. Monitoring of unintended adverse effects of *Wolbachia* releases

Given the well-established safety profile of *Wolbachia*-infected *Ae. aegypti*, we do not anticipate any adverse effects associated with *Wolbachia* deployment during this trial. In order to demonstrate that the deployment is not associated with any excess of a severe adverse outcome, we will follow up all enrolled participants (test-positive cases and test-negative controls) by telephone within 14 to 21 days post-enrolment to ascertain their health status, recorded categorically as recovered/died, and whether or not they were ever hospitalised during this illness. Follow up of participants who do not have a phone number will be done by home visit. Any death of a study participant within 14 to 21 days of enrolment will be classified as a serious adverse event (SAE) and reported to the TSC, IDMC and UGM and Monash University ethics committees within 7 days of ascertainment. The proportion of participants in each arm that were hospitalised or died will be reviewed by the Independent Data Monitoring Committee each time they meet (see Section 13.4), and at any other time at the request of the Trial Steering Committee or other agencies. The Trial Steering Committee will also be kept up-to-date on the community disposition to the intervention.

11. Statistical methods

11.1. Sample size estimation

It is estimated that approximately 1000 cases plus four times as many controls will be sufficient to detect a 50% reduction in dengue incidence with 80% power. The estimate relies on several assumptions, outlined below.

There are no published formulae to estimate sample size for the proposed study design, ie. a cluster randomised trial with a test-negative design, where the intervention effect is estimated from outcome-based sampling of test-positive and test-negative patients and

ascertainment of their exposure status. Randomisation provides a basis of inference in comparing intervention clusters with control clusters as, under the null hypothesis, there should be no difference with regard to the relative appearance of test-positives and negatives in clusters, on average, across the two arms. Thus we propose as the primary analytical approach a comparison of the exposure odds among test-positive cases versus test-negative controls (for data aggregated across all clusters), with the null hypothesis that the odds of residence in a *Wolbachia*-treated cluster is the same among test-positive cases as test-negative controls. The resulting odds ratio thus provides an estimation of the intervention effect and, as demonstrated previously, provides an unbiased estimate of the relative risk providing that the key assumptions underlying the TND are upheld, as outlined in section 4.2.

A secondary approach employs as a summary measure for a group-level analysis the proportion of test-positive cases amongst all tested participants in each cluster, with a comparison of the average of these proportions in the intervention arm versus the untreated arm forming the basis of hypothesis testing for intervention effect. The null hypothesis is that the average proportion of total enrolled participants that are cases is the same in treated and untreated study arms. The alternative hypothesis is that the proportion of enrolled participants that are cases is lower in the *Wolbachia* treated arm than the untreated arm.

Simulations were used to estimate the power to detect a range of intervention effect sizes using the two methods above, assuming 12 clusters per arm, a total of 1000 true dengue cases enrolled and 4000 non-dengue controls. Empirical data on population, historical dengue incidence and incidence of other febrile illness in the 24 study clusters were used to define the baseline characteristics for the simulated scenarios. Nine overlapping two-year windows of dengue data (2003-2014) were sourced from the Yogyakarta surveillance system. Data for other febrile illness during 2014-2015 were sourced from individual Puskesmas using ICD10 codes for non-localising fever (fever of unknown origin R50; Typhus A75.9; and acute infection due to bacteria at an unspecified site A49). We randomly allocated half the clusters to receive the intervention; this random allocation was repeated one million times, and only those allocations were kept in which the balancing criteria specified in the constrained randomization methods were met (n=247 balanced allocations, and thus 494 possible distinct randomizations of intervention allocation). Dengue case numbers per cluster were either kept at baseline values (for the simulation at the null; ie RR=1) or reduced proportionately (for

simulations of intervention effects of RR=0.6, 0.5, 0.4, 0.3). For each of these five 'true' effect sizes, applied to each of the 247 balanced allocations, the 'observed' effect size was calculated from the simulated data by the two methods outlined above; i) aggregated odds ratio for residence in a treated cluster among cases versus controls, and ii) t-test for comparison of the average cluster summary proportions (cases/cases+controls) between study arms. Statistical inference, from the t-test directly, or, for the odds ratios using permutation distribution approximations with standard errors adjusted to account appropriately for the clustered nature of the data, respectively, was used to calculate the proportion of constrained random allocations that yielded a significant result. This provided an estimate of Type I error at the null, and power away from the null (Table 3). Both of these approaches thus are using approximations to the exact permutation distribution.[63] In practice, the appropriate reference distribution for inference will be based on the set of 247 potential balanced allocations.

Table 3: Percent of random allocations that yield significant results on simulated data

Risk Ratio	*T-test*		*Odds Ratio test*	
	Constrained	*Random*	*Constrained*	*Random*
1	0.13	5	1	7
0.6	48	49	61	57
0.5	81	75	89	82
0.4	97	93	99	96
0.3	100	95	100	100

The results show that constrained randomization is somewhat conservative at the null but generally increases power moderately. The odds ratio test is more powerful than the t-test approach, and will thus be used as the primary analysis with the additional attraction of being standard for the traditional test-negative design.

A re-estimation of sample size requirements was conducted in January 2019 after one year of recruitment. The initial power calculation used 1000 dengue cases and 4000 non-dengue controls allocated to each cluster based on historical proportions of dengue cases and other febrile illnesses, assuming no variation in the proportion of cases by cluster. This method was

found to overestimate power for small samples by not taking into account randomness in the sampling. The sample size re-estimation included power estimates for 200, 400, 600, 800 and 1000 dengue cases with 4 times as many controls allocated to each cluster by sampling from a multinomial distribution, which incorporated added randomness by allowing the proportion of cases allocated to each cluster to vary across simulations. The re-estimation found that 400 dengue cases plus four times as many controls would be sufficient to detect a 50% reduction in dengue incidence with 80% power.

Additional simulations were conducted in September 2019 to assess the potential impact on power if a number of untreated clusters were 'lost' to *Wolbachia* contamination. For the target minimum observed effect size of 50% (RR=0.5) and 400 enrolled dengue cases, contamination of 3 untreated clusters (assuming that contaminated clusters experience the full intervention effect for 1 out of the 3 years of trial recruitment) is expected to result in a ~7% loss of power, and contamination of 6 clusters to result in a ~14% loss of power.

11.2. Analysis plan for primary endpoint

11.2.1. Intention-to-treat analysis

The intention-to-treat (primary) analysis will consider *Wolbachia* exposure as a binary classification based on residence in a cluster allocated to *Wolbachia* deployment or not. Residence will be defined as the primary place of residence during the 10 days prior to illness onset. The intention-to-treat analysis will be performed on data acquired during the case surveillance period, i.e. the period commencing when *Wolbachia* is deemed to have been established throughout intervention clusters, defined as one month after completion of releases in the last cluster.

The intervention effect will be estimated from an aggregate odds ratio comparing the exposure odds (residence in a *Wolbachia*-treated cluster) among test-positive cases versus test-negative controls (for data aggregated across all clusters), using the constrained permutation distribution as the foundation for inference. The null hypothesis is that the odds of residence in a *Wolbachia*-treated cluster is the same among test-positive cases as test-negative controls. The resulting odds ratio provides an unbiased estimate of the RR providing that the key assumptions underlying the TND are upheld, as outlined in section 4.2. To note,

since the constrained permutation distribution used for statistical inference contains only the 247 potential allocations (494 distinct randomisations) that meet all balancing criteria, the most extreme odds ratio in the distribution would carry a two-sided p-value of ~0.004 (1/494*2). Therefore p<0.004 is the minimum threshold at which statistical significance can be evaluated in this design.

A secondary group-level analysis will be performed using a cluster-level summary measure of the proportion of test-positive individuals amongst all tested individuals in each cluster. The difference in the average proportion of test positives between the intervention clusters and untreated clusters will be used to test the null hypothesis of no intervention effect using the t-test statistic but basing inference on the exact permutation distribution. These average proportions in each arm can be used to derive an estimate of the RR of dengue in treated versus untreated clusters, which is a much more intuitive effect measure, using a method described in detail elsewhere[63]. Briefly, we can substitute the estimated difference in the proportions, *d* into the formula $d = \frac{1}{1+(\frac{r}{2})(1+RR)} - \frac{RR}{RR+(\frac{r}{2})(1+RR)}$, where *r* is the average number of test negatives for every test positive, which yields a quadratic equation for the unknown *RR.* Only one solution is plausible so that this then yields an estimate of *RR*, along with the appropriately transformed confidence interval (from that associated with *d*).

11.2.2. Per-protocol analysis

The per-protocol analysis will consider *Wolbachia* exposure as a quantitative index based on measured *Wolbachia* prevalence in local *Ae. aegypti* mosquitoes in the participant's cluster of residence, and in locations visited by the participant during the period 3-10 days prior to illness onset. The per-protocol analysis therefore allows for *Wolbachia* exposure to vary in a location over time, and also accounts for human mobility, in terms of the exposure-time that individuals spend outside their cluster of residence as reported in the travel history interview at enrolment. This analysis can also account for the temporal matching of dengue cases and test-negative controls: risk sets of cases and controls will be defined by frequency matching enrolled confirmed dengue cases to arbovirus-negative controls with illness onset in the same calendar month. In the unlikely event that a minimum of four controls cannot be found for a case within the same calendar month, the window for matching can be extended until four controls are identified, for that case only. The per-protocol analysis will include all participants

enrolled from the commencement of the main phase of clinic-based sampling (i.e. excluding the pilot phase, but including participants enrolled before *Wolbachia* was established in treated clusters) unless no cases arise in the corresponding quarter.

Participants will be asked about their mobility during the ten days prior to illness onset using a structured interview administered at enrolment. This will record the duration of time spent at home, work or school, and up to three other most-visited locations during daylight hours (5am – 9pm) in the ten-day period. The geographic coordinates of those locations will be derived by geo-locating them on a digital map, with the assistance of the respondent. A weighted '*Wolbachia* exposure index' (WEI) will be defined for each participant, as follows. The aggregate *Wolbachia* prevalence for each cluster will be calculated each month from all *Ae. aegypti* trapped in that cluster. The WEI for each participant will then be calculated by multiplying the cluster-level *Wolbachia* prevalence (in the month of participant enrolment) at each of the locations visited, by the proportion of time spent at each location, to give a value on a continuous scale from 0 to 1. The process of calculating WEI will be conducted blinded to participants' case/control status, by partitioning the travel history data from the laboratory diagnostic data, to remove any possibility of observer bias.

An additional per-protocol analysis will be conducted in which the WEI is calculated using only the cluster-level *Wolbachia* prevalence in the participant's cluster of residence (in the month of participant enrolment), ignoring the participant's recent travel history. This recognises that dengue exposure risk may be higher at home versus other locations, rather than assuming an even distribution of exposure risk across daytime hours and locations visited.

Cases and controls will be classified by strata of their WEI (e.g. 0-0.2; 0.2-0.4; 0.4-0.6; 0.6-0.8; 0.8-1). This acknowledges that the WEI is not a highly precise measure, and serves to reduce error in exposure classification. Both inference methods described above will be extended to allow for this individual level covariate using regression approaches and extension of the permutation-derived inference used to test the null[64]. For a time-adjusted analysis, a Cox proportional hazard model will be fitted, incorporating the temporal risk sets and using a shared frailty for cluster membership. Such models yield an estimate, and associated confidence interval, for the incidence rate ratio (IRR, the relative hazard). The WEI strata will be included as categorical variables to calculate stratum-specific IRRs (relative to the baseline

0-0.2 stratum). This will allow examination of a 'dose response' relationship. An additional benefit of transforming WEI to a categorical variable is that it avoids any assumption of linearity in the dose response relationship.

11.3. Analysis of secondary objectives

11.3.1. DENV serotype-specific efficacy of *Wolbachia* deployment

In laboratory experiments, the degree to which *Wolbachia* reduces the DENV transmission potential of *Ae. aegypti* is dependent on the infecting virus serotype, with DENV1 transmission least affected [52]. A secondary analysis will estimate the serotype-specific efficacy of *Wolbachia* deployments in reducing symptomatic dengue virus infection with a known infecting serotype, for each of the four serotypes in turn, or as many as are detected in the study population. The same intention-to-treat and per-protocol analyses will be used as described for the primary endpoint above, with case populations restricted to each of the DENV serotypes in turn, and with the same control population as for the primary analysis.

11.3.2. Impact of *Wolbachia* deployment on Zika and chikungunya

There exists no baseline data on the prevalence of Zika or chikungunya infection among febrile patients presenting to primary health care clinics in Yogyakarta City, from which to estimate the expected number of cases; therefore, these secondary analyses are exploratory only and not subject to any formal sample size or power calculations. Blood samples from enrolled participants will be tested by Zika and chikungunya PCR for the purpose of defining arbovirus-negative controls for the primary analysis, as described above. These results will permit estimation of the prevalence of virologically confirmed Zika virus and chikungunya virus infection among the study population of ambulatory febrile patients presenting to primary health care.

If virologically confirmed Zika or chikungunya cases are detected, a secondary analysis will estimate the efficacy of *Wolbachia* deployments in reducing the incidence of symptomatic virologically confirmed Zika virus and chikungunya virus infection. The same enrolled patient population will be used to analyse all three arbovirus endpoints (dengue, Zika and chikungunya), and the same intention-to-treat and per-protocol analyses will be used as described for the primary (dengue) endpoint above. For Zika and chikungunya, the cases will be defined as enrolled participants who test positive by Zika or chikungunya PCR, respectively, and the controls will be those who test negative to all three arboviruses. Cases and controls

will be matched by month of illness onset, as described above (11.2.1). Statistical methods will be as described above (11.2.2 and 11.2.3).

11.3.3. Impact of *Wolbachia* deployment on notified dengue cases

The existing system for routine notification of dengue cases in Yogyakarta City is based on hospital-reporting of cases diagnosed clinically as Dengue Hemorrhagic Fever (DHF), which historically have not been accompanied by supportive laboratory testing. Since March 2016, hospitals have been encouraged to record a serological testing result, where available, on the report form, and also to report cases diagnosed clinically as Dengue Fever where there is a confirmatory NS1-positive test result. A separate reporting system, established in March 2016, collates data on the number of NS1 rapid tests performed – and number positive – in primary health clinics (Puskesmas) across the city. Both of these reporting systems include address information for notified cases.

We will collate data from these two reporting systems on a monthly basis from 2016-2020, aggregated by Kelurahan of residence, to monitor trends in reported dengue incidence across the City and by Kelurahan, before, during and after *Wolbachia* deployment.

The impact of *Wolbachia* deployment on DHF case notifications will be evaluated using an interrupted time series analysis of monthly DHF notifications by kelurahan, before and after *Wolbachia* releases (January 2006 – December 2020). Methods will be developed and validated to classify area-level *Wolbachia* exposure status in a way that aligns with the administrative (kelurahan) boundaries by which dengue cases are reported.

11.3.4. Human mobility in Yogyakarta and implications for measuring efficacy of *Wolbachia* deployment

Understanding the level and distribution of routine movements among the study population is critical to the success of this study design. 'Contamination' by human mobility between study arms may lead to a dilution of the true intervention effect in the ITT analysis, and will influence the degree to which the per-protocol analysis can retain comparison groups with different levels of *Wolbachia* exposure after taking into account participant's crude movement patterns. The data captured through the travel history interview will be analysed to quantify the geographical extent and duration of participants' travel outside the home, and

to estimate the proportion of their daytime ('at risk') hours that participants randomised to treated and untreated arms *actually* spend in *Wolbachia*-treated and untreated areas, overall and by age group. An age-stratified analysis will describe the proportion of participants' time (5am – 9pm) spent at home versus away from home, estimate the distribution of participants' time as a function of increasing distance from home, and identify the predominant non-home locations at which participants in different age groups spend their daytime hours. This information can inform the design of future trials of cluster-randomised household-based interventions, by estimating the optimal size of the clusters needed to account for the majority of daily movements and providing information on the degree to which a true intervention effect might be diluted by movement of participants between treatment arms.

11.3.5. *Wolbachia*-mediated effects on *Aedes* species abundance

The AWED trial provides an opportunity to explore whether fitness costs associated with *Wolbachia* infection of *Ae. aegypti* that have been identified in laboratory environments (e.g. egg survivorship) manifest as a lower population size of adult mosquitoes in areas where *Wolbachia* is established versus untreated areas. We propose to use the AWED trial as a basis to measure and compare the number of adult *Aedes* mosquitoes caught in *Wolbachia*-treated versus untreated clusters. This secondary analysis uses only existing data from BG trap mosquito collections, and no patient samples.

Up to three years of weekly BG collections conducted as part of routine *Wolbachia* monitoring in the AWED trial are available, providing counts of *Ae. aegypti*, *Ae. albopictus* and other species in each trap. Additional BG collections from Yogyakarta City are available from prior to the commencement of *Wolbachia* mosquito releases in 2017. Poisson regression will be used to test the null hypothesis (H_0) of no difference in the abundance of each species by treatment arm. The clustered sampling of *Ae. aegypti* mosquitoes by intervention cluster and BG trap will be accounted for in the analysis, by stratifying on RCT cluster and including BG trap as a random effect in a mixed-effect Poisson regression.

12. Data management

12.1. Data collection and coding

Field data on *Wolbachia* deployment and monitoring will be captured through standardised electronic data capture forms deployed on mobile devices. When connected to the internet,

the devices will sync with a web-based Core Data Repository and all new data will be uploaded.

Data collected from participants in the clinical study will be similarly captured through standardised electronic data capture forms and digital mapping interfaces, deployed either on mobile devices or through web-based applications on desktop or laptop computers. Laboratory diagnostic results will be captured directly from laboratory assay output.

Validity controls will be applied at the point of data capture into electronic forms, by predefining value ranges, specifying categorical option lists, and minimising the use of free text fields. The use of carefully designed electronic forms will facilitate the coding of participant responses at the point of data collection.

12.2. Data storage and security

Field data on *Wolbachia* release and monitoring will be stored in the Core Data Repository, a custom designed relational database hosted on an Australian web-based server. A parallel custom designed database to store clinical study data is in development, using the same infrastructure as the Core Data Repository. Clinical study data will be entered into electronic data capture forms which feed directly into the clinical database, or uploaded to the database from laboratory assay output.

In order to maintain blinding of research staff and data managers, measures will be put in place to ensure the datasets identifying participant's exposure status (cluster of residence and clusters visited during 10 days prior to illness) will remain unlinked and partitioned from the dataset that classifies their case/control status until the final analysis. In the event that the Independent Data Monitoring Committee requires data to be unblinded following the interim analysis, a single member of the World Mosquito Program, Monash University data management unit will be responsible for linking the participant dataset to the exposure status.

Role based, tiered access permissions will be used to control access to the Core Data Repository, clinical database and associated data capture applications. User logs will document the activities of all users. Security of the web-hosted databases will be assured by the security processes of the cloud service (Amazon Web Services), namely: automated

backups and database snapshots, high-level availability and 24/7 incident response and detection. The overall Core Data architecture has been subject to a security audit by Monash University's IT operations, eSolutions.

12.3. Data quality assurance

Quality control in the form of logic and consistency checks will be applied at several stages of data capture and management: i) at the point of data capture into an electronic form; ii) at the point of upload into the web-based database; and iii) during routine monitoring processes by internal and external data monitors. An audit trail will be preserved within the database to capture the history of any changes made to data records after their initial capture.

12.4. Study record retention

All data relating to the trial, including field entomology and epidemiological data, will be retained indefinitely, and for a minimum of 5 years after study completion, in accordance with ICH-GCP requirements.

13. Ethical considerations and trial governance

13.1. Summary of governance structure

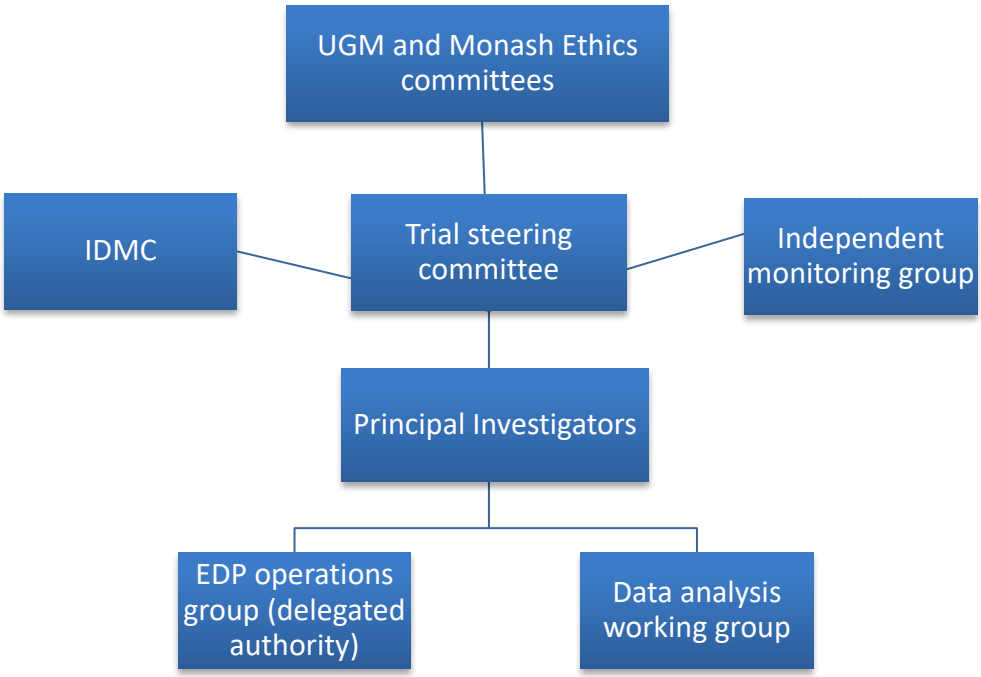


Figure 9: Trial governance structure

The Principal Investigator (PI) from Universitas Gadjah Mada, Yogyakarta, supported by the Chief Investigator from Monash University, will be responsible for ensuring the study is performed in compliance with the approved protocol and the principles of Good Clinical Practice.

The Trial Steering Committee (TSC), chaired by the PIs, will include one or more co-investigators and one or more members who are independent of the investigators and sponsors. The TSC will provide overall supervision of the trial, including monitoring of recruitment progress, and will consider and act upon (as appropriate) any recommendation from the IDMC with regards to early stopping of the trial.

The Trial Operations Group will, under the delegation of the PI, be responsible for day-to-day coordination of the trial processes.

The Monitoring Group will be independent of the investigators, and will conduct periodic monitoring of study processes including data collection and storage, sample collection and chain of custody, and laboratory processes.

The Data Analysis Working Group will be chaired by the trial statistician, and will be responsible for developing the statistical methods for randomization, data cleaning and validation, and preparing and implementing the statistical analyses.

13.2. Ethical review

This protocol and the informed consent document and any subsequent modifications will be reviewed and approved by the institutional review boards (IRBs) of Universitas Gadjah Mada, Yogyakarta, and Monash University, Melbourne, prior to the commencement of the trial. A letter of protocol approval by the ethical review boards will be obtained prior to the commencement of the trial.

If any substantive changes to study processes are required after commencement of the study, a protocol amendment request will be submitted to both review boards.

13.3. Modifications to the protocol

This study will be conducted in compliance with the current approved version of the protocol. Any change to the protocol document or informed consent form that affects the scientific intent, study design, participant safety, or may affect a participant's willingness to participate in the study is considered an amendment, and therefore will be written as a protocol amendment and submitted to the ethical review boards for approval prior to becoming effective.

13.4. Independent Data Monitoring Committee

An Independent Data Monitoring Committee (IDMC) will be constituted from local and international experts in accordance with standard practice for randomised clinical trials.

The IDMC will meet at study initiation, six months following the commencement of clinic-based enrolment, and then at regular intervals, as well as any other time at the request of the TSC or other agencies. Their primary role is to safeguard the interests of the trial participants, to assess the safety and efficacy of the intervention during the trial, and to monitor the overall conduct of the trial.

The IDMC will provide recommendations about stopping or continuing the trial, and may also make recommendations relating to trial procedures, and data management and quality control. Any proposed major modifications to the study protocol will be reviewed by the IDMC, and approval for a protocol amendment will be sought from the relevant IRBs, prior to their implementation. Detailed responsibilities and terms of reference will be set out in an IDMC charter, and agreed to by all IDMC members, prior to study commencement.

13.5. Interim analyses and stopping rules

An interim analysis of the primary endpoint (intention-to-treat analysis only using the odds ratio approach based on the permutation distribution, as described in 11.2.2) was originally planned for the mid-point of the study, ie after enrolment of 500 dengue cases from an initial target sample size of 1000. Following re-estimation of statistical power in January 2019 (section 11.1) to examine a range of sample sizes using multinomial sampling, instead of deterministic sampling of 1000 cases, it was estimated that 400 virologically-confirmed dengue cases (and 1600 test-negative controls) would be sufficient to detect a 50% reduction in dengue incidence with 80% power. Given this revised power estimate, the threshold of 500 cases for conducting an interim analysis is no longer appropriate and is unlikely to be reached. Re-calculation of sample size using the observed inter-cluster distribution of participants is no longer required, as the revised methods for power estimation use sampling from a multinomial distribution which already accounts for inter-cluster variability in dengue case distribution. An interim analysis is no longer planned, unless recommended by the independent data monitoring committee. In the event of the IDMC recommending an interim analysis, the study statistician and other members of the data analysis working group will prepare the analysis code and blank tables for interim analysis. Data for analysis will be extracted by the study data management unit, retaining blinding as to the treatment arms of dengue test-positive and test-negative study participants. The study statistician will generate the tables and distribute the interim report among IDMC members.

The IDMC may recommend modification or termination of the study if analyses of data from an interim analysis indicates beyond reasonable doubt that exposure to *Wolbachia* confers a reduced risk of dengue in the intention-to-treat analysis. As detailed in the analysis methods (11.2.1), the use of the constrained permutation distribution for statistical inference means the smallest two-sided p-value that can be observed is p~0.004. The usual Haybittle-Peto

boundary[66], requiring p<0.001 at interim analysis to consider stopping for efficacy, cannot therefore be applied. Instead, p<0.01 at interim analysis will be used as guidance for considering stopping early for efficacy. The IDMC may also recommend termination if preliminary data clearly suggest that *Wolbachia* is associated with an excess of dengue (or Zika or chikungunya) cases. A less conservative p<0.05 in the direction of harm will be used as a guidance. Termination or modification may also be recommended for any other operational reason (e.g. participant enrolment rates), perceived safety concern, or external factor.

Additional criteria for early termination of the trial were introduced in October 2019 in response to increasing *Wolbachia* contamination in several untreated clusters.

i) To address the possibility that a loss of power due to contamination may compromise the ITT analysis, the trial will stop if 5 or more untreated clusters are classified as contaminated. A cluster will be defined as contaminated when the cluster level *Wolbachia* frequency is >50% for 2 monthly monitoring events within a 6-month rolling window *and* >50% of the BG traps in the cluster have detected *Wolbachia* during those monitoring events. An assessment will be made when the *Wolbachia* monitoring results in each cluster are uploaded to the Core data system each month. The participant dataset for analysis will include all those cases enrolled up until and including the date the Trial Steering Committee endorses the decision to stop the trial because this *Wolbachia* contamination threshold is breached.
ii) To address the potential for consistently low rates of virologically-confirmed dengue (VCD) case enrolment to make continued recruitment until November 2020 futile in terms of increasing sample size, and thereby statistical power, the trial will stop if the rate of enrolment of VCD cases is on average <2 per month in a 3-month rolling window, commencing 1 November 2019. i.e. if five or fewer VCD cases are enrolled in a 3-month window the trial will stop. This could theoretically occur due to a large *Wolbachia* effect size and increasing contamination of untreated clusters, or epidemiological natural history, or both. An assessment will be made based on data accrued to the first day of each month, starting 1 February 2020 but allowing for a lag in completion of diagnostic laboratory testing. The participant dataset for analysis will include all those cases enrolled up until and including the date the Trial Steering Committee endorses the decision to stop the trial for this stopping rule.

The final decision to terminate or modify the study rests with the TSC.

13.6. Confidentiality

Confidentiality of participant information will be strictly maintained at all times by the participating investigators, research staff, and the sponsoring institution. This confidentiality is extended to cover testing of biological samples in addition to the clinical, demographic and geospatial information relating to participating subjects. All laboratory specimens, reports, data collection forms and log books, and geo-located records will be identified by a coded ID number only to maintain participant confidentiality. All records that contain names or other personal identifiers, such as informed consent forms, will be stored separately from study records identified by ID numbers. All local databases will be secured with password-protected access systems. No information concerning the study or the data will be released to any unauthorised third party, without prior written approval of the sponsoring institution. Clinical or personal information will not be released without written permission of the subject, except as necessary for monitoring by an ethical review board or regulatory agencies. Reporting of study results will not be done in any way that permits identification of individual participants, or the location of their homes or other visited locations.

13.7. Participant reimbursement

A small gift and cash will be provided to participants after completion of study processes, to acknowledge their contribution and time. The value of this gift will not exceed $10 USD per participant. Participants will not be paid for their participation, nor will the study team be liable for payment of any medical costs.

14. Dissemination and publications policy

14.1. Dissemination of trial results

The scientific integrity of the trial requires that only the results of final analyses will be disseminated publicly; there will be no dissemination of any interim analysis, unless the results lead to early stoppage of the trial. Dissemination of trial results, including any publications arising, will be subject to the prior approval of the Trial Steering Committee. Final trial results will be disseminated to community leaders, healthcare professionals, the public and other relevant stakeholders, as well as being submitted for publication in scientific journals.

14.2. Publication plan

The trial findings will be submitted for peer review and publication in an appropriate open access journal. Every attempt will be made to reduce to a minimum the interval between the completion of data collection and the release of study results. After finalising recruitment, we expect to take no more than four months to prepare the final results paper for submission.

14.3. Authorship eligibility guidelines

Named protocol contributors will be included as authors on the primary report of trial findings, assuming that they have fulfilled international criteria for authorship at the time of manuscript submission. Authors will be expected to have made a substantive contribution to the design, conduct, interpretation and reporting of the trial.

14.4. Data sharing statement

A summary of the trial protocol will be published in an open access journal prior to study commencement, and the full trial protocol will be made publicly available within one year of the conclusion of data collection. The trial will be registered on an appropriate clinical trials database prior to study commencement.

15. Funding source

This study is funded by the Tahija Foundation. The funding source had no role in the design of the study and will not have any role during its execution, analyses, interpretation of the data, or decision to submit results.

16. Conflict of interest

The Principal Investigators have no conflict of interest to declare.

17. References

1. Brady, O. J. *et al.* Refining the global spatial limits of dengue virus transmission by evidence-based consensus. *PLoS Negl. Trop. Dis.* **6,** e1760 (2012).
2. Bhatt, S. *et al.* The global distribution and burden of dengue. *Nature* **496,** 504–7 (2013).
3. Shepard, D. S., Coudeville, L., Halasa, Y. a, Zambrano, B. & Dayan, G. H. Economic impact of dengue illness in the Americas. *Am. J. Trop. Med. Hyg.* **84,** 200–7 (2011).
4. Shepard, D. S., Undurraga, E. a & Halasa, Y. a. Economic and disease burden of dengue in southeast Asia. *PLoS Negl. Trop. Dis.* **7,** e2055 (2013).

5. Shepard, D. S. *et al.* Cost-effectiveness of a pediatric dengue vaccine. *Vaccine* **22,** 1275–80 (2004).

6. Dengue Vaccine Initiative. Dengue vaccine candidates in clinical development. (2016). Available at: http://www.denguevaccines.org/vaccine-development. (Accessed: 13th June 2016)

7. L'Azou, M. *et al.* Symptomatic Dengue in Children in 10 Asian and Latin American Countries. *N. Engl. J. Med.* **374,** 1155–66 (2016).

8. Capeding, M. R. *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. *Lancet* **6736,** 1–8 (2014).

9. Villar, L. *et al.* Efficacy of a Tetravalent Dengue Vaccine in Children in Latin America. *N. Engl. J. Med.* **372,** 113–123 (2015).

10. Guy, B., Lang, J., Saville, M. & Jackson, N. Vaccination Against Dengue: Challenges and Current Developments. *Annu. Rev. Med.* **67,** 387–404 (2016).

11. Hadinegoro, S. R. *et al.* Efficacy and Long-Term Safety of a Dengue Vaccine in Regions of Endemic Disease. *N. Engl. J. Med.* **373,** 1195–206 (2015).

12. Schilte, C. *et al.* Chikungunya virus-associated long-term arthralgia: a 36-month prospective longitudinal study. *PLoS Negl. Trop. Dis.* **7,** e2137 (2013).

13. Rolph, M. S., Foo, S. S. & Mahalingam, S. Emergent chikungunya virus and arthritis in the Americas. *Lancet. Infect. Dis.* **15,** 1007–8 (2015).

14. Weaver, S. C. *et al.* Zika virus: History, emergence, biology, and prospects for control. *Antiviral Res.* **130,** 69–80 (2016).

15. World Health Organization. *Outcome of the Emergency Committee regarding clusters of microcephaly and Guillain-Barre syndrome*. (2016).

16. World Health Organization. Mosquito (vector) control emergency response and preparedness for Zika virus. (2016). Available at: http://www.who.int/neglected_diseases/news/mosquito_vector_control_response/en/. (Accessed: 18th March 2016)

17. Karyanti, M. R. *et al.* The changing incidence of dengue haemorrhagic fever in Indonesia: a 45-year registry-based analysis. *BMC Infect. Dis.* **14,** 412 (2014).

18. Graham, R. R. *et al.* A prospective seroepidemiologic study on dengue in children four to nine years of age in Yogyakarta, Indonesia I. studies in 1995-1996. *Am. J. Trop. Med. Hyg.* **61,** 412–9 (1999).

19. BPS - Statistics Indonesia. Badan Pusat Statistik. (2016). Available at: www.bps.go.id.

20. Ramadona, A. L. *et al.* Prediction of Dengue Outbreaks Based on Disease Surveillance and Meteorological Data. *PLoS One* **11,** e0152688 (2016).

21. Porter, K. R. *et al.* A serological study of Chikungunya virus transmission in Yogyakarta, Indonesia: evidence for the first outbreak since 1982. *Southeast Asian J. Trop. Med. Public Health* **35,** 408–15 (2004).

22. Kosasih, H. *et al.* Evidence for Endemic Chikungunya Virus Infections in Bandung, Indonesia. *PLoS Negl. Trop. Dis.* **7,** e2483 (2013).

23. Riswari, S. F. *et al.* Study of viremic profile in febrile specimens of chikungunya in Bandung, Indonesia. *J. Clin. Virol.* **74,** 61–5 (2016).

24. Laras, K. *et al.* Tracking the re-emergence of epidemic chikungunya virus in Indonesia. *Trans. R. Soc. Trop. Med. Hyg.* **99,** 128–41 (2005).

25. Mulyatno, K. C., Susilowati, H., Yamanaka, A., Soegijanto, S. & Konishi, E. Primary isolation and phylogenetic studies of Chikungunya virus from Surabaya, Indonesia. *Jpn. J. Infect. Dis.* **65,** 92–4 (2012).

26. Haddow, A. D. *et al.* Genetic characterization of Zika virus strains: geographic expansion of the Asian lineage. *PLoS Negl. Trop. Dis.* **6,** e1477 (2012).

27. Perkasa, A. *et al.* Isolation of Zika Virus from Febrile Patient, Indonesia. *Emerg. Infect. Dis.* **22,** 924–5 (2016).

28. Kwong, J. C., Druce, J. D. & Leder, K. Zika virus infection acquired during brief travel to Indonesia. *Am. J. Trop. Med. Hyg.* **89,** 516–7 (2013).

29. Leung, G. H. Y., Baird, R. W., Druce, J. & Anstey, N. M. Zika virus infection in Australia following a monkey bike in Indonesia. *Southeast Asian J. Trop. Med. Public Health* **46,** 460–4 (2015).

30. Esu, E., Lenhart, A., Smith, L. & Horstick, O. Effectiveness of peridomestic space spraying with insecticide on dengue transmission; systematic review. *Trop. Med. Int. Health* **15,** 619–31 (2010).

31. Pilger, D., De Maesschalck, M., Horstick, O. & San Martín, J. L. Dengue outbreak response: documented effective interventions and evidence gaps. *TropIKA* **1,** (2010).

32. Erlanger, T. E., Keiser, J. & Utzinger, J. Effect of dengue vector control interventions on entomological parameters in developing countries: a systematic review and meta-analysis. *Med. Vet. Entomol.* **22,** 203–21 (2008).

33. Bowman, L. R., Runge-Ranzinger, S. & McCall, P. J. Assessing the relationship between vector indices and dengue transmission: a systematic review of the evidence. *PLoS Negl. Trop. Dis.* **8,** e2848 (2014).

34. Bowman, L. R., Donegan, S. & McCall, P. J. Is Dengue Vector Control Deficient in Effectiveness or Evidence?: Systematic Review and Meta-analysis. *PLoS Negl. Trop. Dis.* **10,** e0004551 (2016).

35. Andersson, N. *et al.* Evidence based community mobilization for dengue prevention in Nicaragua and Mexico (Camino Verde , the Green Way): cluster randomized controlled trial. *BMJ* **351,** h3267 (2015).

36. Degener, C. M. *et al.* Evaluation of the effectiveness of mass trapping with BG-sentinel traps for dengue vector control: a cluster randomized controlled trial in Manaus, Brazil. *J. Med. Entomol.* **51,** 408–20 (2014).

37. Wilson, A. L. *et al.* Evidence-based vector control? Improving the quality of vector control trials. *Trends Parasitol.* **31,** 380–90 (2015).

38. O'Neill, S. L. *et al.* In vitro cultivation of Wolbachia pipientis in an Aedes albopictus cell line. *Insect Mol. Biol.* **6,** 33–9 (1997).

39. Stouthamer, R., Breeuwer, J. A. & Hurst, G. D. Wolbachia pipientis: microbial manipulator of arthropod reproduction. *Annu. Rev. Microbiol.* **53,** 71–102 (1999).

40. Rousset, F., Vautrin, D. & Solignac, M. Molecular identification of Wolbachia, the agent of cytoplasmic incompatibility in Drosophila simulans, and variability in relation with host mitochondrial types. *Proc. Biol. Sci.* **247,** 163–8 (1992).

41. Hilgenboecker, K., Hammerstein, P., Schlattmann, P., Telschow, A. & Werren, J. H. How many species are infected with Wolbachia?--A statistical analysis of current data. *FEMS Microbiol. Lett.* **281,** 215–20 (2008).

42. McMeniman, C. J. *et al.* Stable introduction of a life-shortening Wolbachia infection into the mosquito Aedes aegypti. *Science* **323,** 141–4 (2009).

43. Joubert, D. A. *et al.* Establishment of a Wolbachia Superinfection in Aedes aegypti Mosquitoes as a Potential Approach for Future Resistance Management. *PLoS Pathog.* **12,** e1005434 (2016).

44. Walker, T. *et al.* The wMel Wolbachia strain blocks dengue and invades caged Aedes aegypti populations. *Nature* **476,** 450–3 (2011).

45. Dutra, H. L. C. *et al.* Wolbachia Blocks Currently Circulating Zika Virus Isolates in Brazilian Aedes aegypti Mosquitoes. *Cell Host Microbe* **19,** 771–4 (2016).

46. Johnson, K. N. The Impact of Wolbachia on Virus Infection in Mosquitoes. *Viruses* **7,** 5705–17 (2015).

47. Rainey, S. M., Shah, P., Kohl, A. & Dietrich, I. Understanding the Wolbachia-mediated inhibition of arboviruses in mosquitoes: progress and challenges. *J. Gen. Virol.* **95,** 517–30 (2014).

48. Amuzu, H. E., Simmons, C. P. & McGraw, E. A. Effect of repeat human blood feeding on Wolbachia density and dengue virus infection in Aedes aegypti. *Parasit. Vectors* **8,** 246 (2015).

49. Ye, Y. H. *et al.* Wolbachia Reduces the Transmission Potential of Dengue-Infected Aedes aegypti. *PLoS Negl. Trop. Dis.* **9,** e0003894 (2015).

50. Frentiu, F. D. *et al.* Limited dengue virus replication in field-collected Aedes aegypti mosquitoes infected with Wolbachia. *PLoS Negl. Trop. Dis.* **8,** e2688 (2014).

51. Moreira, L. A. *et al.* A Wolbachia symbiont in Aedes aegypti limits infection with dengue, Chikungunya, and Plasmodium. *Cell* **139,** 1268–78 (2009).

52. Ferguson, N. M. *et al.* Modeling the impact on virus transmission of Wolbachia-mediated blocking of dengue virus infection of Aedes aegypti. *Sci. Transl. Med.* **7,** 279ra37-279ra37 (2015).

53. Wolbers, M., Kleinschmidt, I., Simmons, C. P. & Donnelly, C. a. Considerations in the design of clinical trials to test novel entomological approaches to dengue control. *PLoS Negl. Trop. Dis.* **6,** e1937 (2012).

54. Hayes, R. & Moulton, L. *Cluster randomized trials*. (Chapman & Hall/CRC, 2009).

55. Endy, T. P., Yoon, I.-K. & Mammen, M. P. Prospective cohort studies of dengue viral transmission and severity of disease. *Curr. Top. Microbiol. Immunol.* **338,** 1–13 (2010).

56. Vandenbroucke, J. P. & Pearce, N. Case-control studies: Basic concepts. *Int. J. Epidemiol.* **41,** 1480–1489 (2012).

57. Serres, G. De, Skowronski, D. M., Wu, X. W. & Ambrose, C. S. The test-negative design : validity , accuracy and precision of vaccine efficacy estimates compared to the gold standard of randomised placebo-controlled clinical trials. *Eurosurveillance* **18,** pii=20585 (2013).

58. Greenland, S. & Thomas, D. C. On the need for the rare disease assumption in case-control studies. *Am. J. Epidemiol.* **116,** 547–553 (1982).

59. Jackson, M. L. & Nelson, J. C. The test-negative design for estimating influenza vaccine effectiveness. *Vaccine* **31,** 2165–2168 (2013).

60. Haber, M. *et al.* A Probability Model for Evaluating the Bias and Precision of Influenza Vaccine Effectiveness Estimates from Case-Control Studies. *Epidemiol. Infect.* **143,** 1417–1426 (2015).

61. Smith, P., Morrow, R. & Ross, D. *Field trials of health interventions*. (Oxford University Press, 2015).

62. Ivers, N. M. *et al.* Allocation techniques for balance at baseline in cluster randomized trials: a methodological review. *Trials* **13,** 120 (2012).

63. Jewell, N. P., Dufault, S., Cutcher, Z., Simmons, C. P. & Anders, K. L. Analysis of cluster-

randomized test-negative designs: cluster-level methods. *Biostatistics* (2018). doi:10.1093/biostatistics/kxy005

64. Small, D. S., Ten Have, T. R. & Rosenbaum, P. R. Randomization Inference in a Group–Randomized Trial of Treatments for Depression. *J. Am. Stat. Assoc.* **103,** 271–279 (2008).

65. Faul, F., Erdfelder, E., Lang, A.-G. & Buchner, A. G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav. Res. Methods* **39,** 175–91 (2007).

66. Haybittle, J. L. Repeated assessment of results in clinical trials of cancer treatment. *Br. J. Radiol.* **44,** 793–7 (1971).

Summary of protocol amendments

Note: Light grey shading indicates amendments relating to trial design and analysis, including measurement of exposure and outcomes; others amendments had only operational implications.

Changes made from protocol version 1 (21 October 2016) to version 2 (24 March 2017)		
Page	**Amendment**	**Comment**
1	Modified trial short title	Short title changed to include trial acronym: 'The AWED trial – Applying *Wolbachia* to Eliminate Dengue'
1	Added trial registration identifier	Clinicaltrials.gov identifier added
1	Updated protocol contributors	Removed a contributor who has left the project
5 – 6	Updated quick reference table	To reflect amendments detailed elsewhere in the protocol
16, 21, 33, 36	Removed age-matching of controls to cases	Revised analysis plan will include age as a covariate to be adjusted for in per-protocol analysis, rather than frequency matching case-control risk sets by age as well as calendar month of enrolment.
20	Revised estimate of the expected number of eligible participants presenting to study clinics	Based on historical data, revised the estimate of the number of eligible participants likely to present to study clinics from 6000 per annum to 5000. This has no impact on the sample size estimations, as we will still enrol all eligible and consenting participants and the minimum estimated required sample size remains unchanged (1000 cases and 4000 controls).
21, 27	Clarified that recruitment will continue for 24 months even if the estimated minimum sample size is reached sooner	The sample size estimate is a minimum only, and based on a number of assumptions, so recruitment will continue for the full 24 months even if that minimum sample size is exceeded.
21	Updated *Wolbachia* deployment start date	Updated the deployment start date to March 2017 instead of February 2017

22	Removed prospective data collection from clinics on eligible patient numbers	The feasibility of achieving the minimum target sample size has already been assessed using historical data (2015) from clinics.
24	Constrained randomisation used 4 spatial strata instead of 3	In order to minimise the chance occurrence of a large number of contiguous intervention clusters.
24	Up to 5 release points per 50m grid, instead of 1	To allow for an increased density of *Wolbachia* mosquito releases in areas where it is slower to achieve and sustain a high prevalence of *Wolbachia* in local mosquito populations.
24	Increased number of eggs per release container from 80 to 100-150	To allow for an increased density of *Wolbachia* mosquito releases in areas where it is slower to achieve and sustain a high prevalence of *Wolbachia* in local mosquito populations.
25	Changed density of mosquito trap network for *Wolbachia* monitoring	Changed to a consistent density of 16 BG traps per km2 throughout the study area (intervention and untreated clusters), rather than a fixed number of traps (10) per cluster. This aims to achieve more consistent spatial representativeness in the *Wolbachia* prevalence estimates between clusters.
29	Specified that sex and contact phone number of participants will also be recorded	Sex is a standard demographic variable to be recorded, and contact phone number is required for ascertaining participants'status at 14 days post enrolment.
31	Updated diagnostic algorithm	Samples will be tested first by PCR, and only samples negative for all 3 viruses by PCR need to be tested by NS1 ELISA. Similarly only samples negative by NS1 ELISA need to be tested by IgM. The case/control classification criteria remain unchanged, this just saves unnecessary testing of samples for participants who will already be classified as a case on the basis of a positive result in a prior assay.
31	Added dengue IgG serology for test-negative samples collected during pilot phase	To determine what proportion of participants who would otherwise be classified as controls, have detectable dengue IgG indicating potentially acute secondary dengue or another cross-reactive flavivirus infection. If the prevalence is high enough, this may justify future addition of dengue IgG capture ELISA to the diagnostic algorithm.

33	Added details of serious adverse event classification	The death of a study participant within 14 days of enrolment will be classified as an SAE and reported to appropriate committees within 7 days.
33-35	Refined methods for sample size estimation	Novel simulation-based methods to estimate the required minimum sample size have been developed, because no existing formulae existed for the proposed study design. The methods described for sample size/power estimation now align with the intended methods for analysis.
36-38	Refined analysis methods	Additional detail has been added on planned group-level analyses, and the extensions to traditional statistical methods that are being developed for analysis of trial data to adequately account for both the temporal matching of cases and controls and the non-independence of study participants resident in the same intervention cluster.
40	Updated details of trial data web storage location	The custom designed web database will now be hosted solely on an Australian web-based served, rather than split between Indonesian and Australian locations.
Changes made from protocol version 2 to version 3 (13 March 2018)		
1, 5	Amend protocol title to include the trial name and acronym	The trial name and acronym were already given in the 'Short title' but have now been added to the protocol title
1, 2, 13, 43	Change of research program name	The name of our global research program has changed from Eliminate Dengue Program to World Mosquito Program.
16, 42	Additional secondary objective and analysis	Arboviral infection prevalence in Ae.aegypti mosquitoes will be assessed in intervention and control clusters
16, 41	Modified wording of human mobility secondary objective and analysis	The description of the human mobility secondary endpoint has been rephrased to better reflect its objective, and to give clarity on what analyses will be performed.
16, 28, 29	Amendment of inclusion criteria	Maximum age for eligibility amended from 30 to 45 years, based on preliminary enrolment numbers from the pilot study and discussion within the research group

27	Calculation of *Wolbachia* frequency	We continue to refine the methodology used by the program. No minimum number of *Ae aegypti* is now specified for the four-weekly screening to determine *Wolbachia* prevalence (previously minimum 100 per cluster per screening event)..
22, 27, 28, 38	Definition of *Wolbachia* establishment	Definition of *Wolbachia* establishment for the purpose of measuring the efficacy endpoint was amended from '80% of intervention clusters have two consecutive screening events with >80% Wolbachia prevalence', to one month after completing releases in the last cluster, for simplicity.
29	Change in terminology for exclusion criteria	"Localising features suggestive of an alternative diagnosis" changed to "Localising features suggestive of a specific diagnosis" for clarity
30	Additional data field	Participants will be asked about whether they have received the dengue vaccine
30	Timing of follow-up call correction	Follow-up call will be done within 14-21 days after enrolment (rather than 14 days strictly) to allow a one week window for follow-up completion
31	Topical anaesthetic for blood draws	An option has been added for use of a topical anaesthetic when drawing the blood sample.
32	Updated Figure	Figure 8 updated to reflect change in age inclusion criteria, include IgG test, and include action based on equivocal test results
32	Equivocal test result	Equivocal test results will be reported as "equivocal" without retesting
33	Blood sample storage	With participants' consent, blood samples will be stored for future research for a minimum of 5 years. This is added to the protocol to align with the text in the explanatory statement for participants.
34	Follow up of safety endpoints	Participants who do not have a phone number will be followed up by home visit.
35	Mention community-level adverse events	To clarify that the TSC will be kept up-to-date on the community disposition to the intervention.

35, 37	Primary analysis approach	We now make explicit that the aggregate odds ratio is the primary approach for the ITT analysis.
36	Power calculation simulations	We now clarify that there were in fact 494 possible distinct randomisations that met balancing criteria, since for each of the 247 balanced allocations there were two possibilities for which arm received the intervention.
37	Analysis approach	We make explicit that the constrained permutation distribution will be used for statistical inference.
7, 37, 39	Temporal matching of cases and controls	Our development of statistical methods for analysis of the trial data has progressed, and we now clarify that in the ITT analysis all cases and controls will be used, without time matching. The per protocol analysis will still include temporal matching of cases and controls.
39	Adjustment for mobility in per-protocol analysis	We have clarified the description of how the weighted *Wolbachia* exposure index will be calculated, for the per protocol analysis.
45	Update organogram of governance structure	Governance structure updated to align with IDMC recommendations.
47	Interim analyses and stopping rules	The interim analysis plan and stopping rules have been modified to reflect developments in the statistical approach, namely that $p<0.005$ is the smallest achievable p-value.
49	Funder information	Addition of details of funder and role of funder and sponsor in study design and analysis
50	Financial and competing interests	Addition of disclosure for financial and competing interests
-	ICF, Assent form and Explanatory statement	English versions of the ICF, assent form and explanatory statement were removed from the appendices to avoid the need to resubmit the protocol where changes were made only to the explanatory statement, consent and/or assent forms. These will be submitted as separate documents along with protocol amendments where necessary.
1	Addition of a new protocol contributor	A new epidemiologist joined the global World Mosquito Program team.

Changes made from protocol version 3 to version 4 (28 August 2018)		
44	Addition of cash reimbursement to participants	One change was made to the protocol and explanatory statements to explain that participants will be reimbursed in the form of a small gift and cash (previously with a small gift only) which will not exceed US$10 in value per participant. This change arose at the recommendation of independent clinical research monitors.
Changes made from protocol version 4 to version 5.1 (16 October 2019)		
5, 15, 21, 31, 32, 38	Added new secondary endpoint of serotype-specific DENV infection, and new secondary objective to measure serotype-specific efficacy.	The primary endpoint of the trial includes virologically-confirmed DENV infections of any (or unknown) serotype combined, based on the detection of DENV RNA in a pan-dengue RT-PCR, or detection of NS1 antigen. We now make explicit the use of a second serotype-specific RT-PCR to determine the infecting DENV serotype in samples positive in the pan-dengue PCR, and to estimate serotype-specific efficacy of the *Wolbachia* intervention.
5, 27	Study duration now up to 36 months for participant enrolment.	A 12-month extension to the trial duration has been approved by the IDMC, to account for lower than expected dengue incidence in Yogyakarta (and elsewhere in Indonesia) during the first year of the study period.
6, 33	Removed wording about re-calculating sample size using observed inter-cluster distribution of participants.	The power re-estimation included in this protocol amendment (see page 35) uses sampling from a multinomial distribution which already accounts for inter-cluster variability in dengue case distribution, so this is no longer necessary.
6	Removed clause about extending trial to reach minimum sample size.	As the trial duration has now been formally extended to 36 months, no further extension intended.
15, 39	Removed original secondary objective #3	Findings so far indicate that the overall prevalence of dengue virus-infected Aedes aegypti mosquitoes is too low for this secondary objective to be feasible, within the resources available.
15, 41	Additional secondary objective	The AWED trial provides an opportunity to explore whether fitness costs associated with *Wolbachia* infection of *Ae. aegypti* that have been identified in laboratory environments (fecundity, egg

		survivorship) manifest as a lower population size of adult mosquitoes in areas where *Wolbachia* is established versus untreated areas. We propose to use the AWED trial as a basis to measure and compare the population size of adult mosquitoes in *Wolbachia*-treated versus untreated cluster. Note that this secondary analysis uses only field mosquito collections, and no patient samples or data.
16	Updated the map of the study area	The previous map had inaccuracies and lacked clarity.
21	Updated study time line	The time line was updated to reflect the extension of trial recruitment by 12 months, to the end of 2020.
27	Clarification of the term ‘guardian’	Clarification that informed consent for underage participants will be sought from a parent or vertical guardian (i.e. grandparent).
30	Use of residual clinical blood sample	In a situation where a consenting participant has already had blood collected for clinical investigations on the day of enrolment, a second blood sample will not be collected for research purposes. Rather, the residual blood sample will be retained and used for study investigations. Two implications of this are noted: 1) the sample volume may be lower than the usual 3ml, and 2) informed consent may be obtained from the participant after the clinical blood sampling has occurred (but before retrieval of the residual sample for use in the study).
6, 36	Power / sample size re-estimation	Statistical power for the trial was re-estimated using a multinomial simulation method, with a range of sample sizes and effect sizes, as compared to the original power estimation which used a deterministic simulation of a range of effect sizes but with a fixed sample size. This revised method better accounts for randomness in sampling, including variability in the distribution of dengue cases and non-dengue febrile patients between clusters. This indicates that adequate power (>=80%) can be achieved with a substantially lower sample size than the 1000 dengue cases assumed for the original power simulation scenario. There is >80% power to detect

		an effect size of RR=0.5 with a sample of 400 cases and 4x400 test-negative controls. Additional simulations explored the potential impact on power if a number of untreated clusters are 'lost' to *Wolbachia* contamination. Assuming 400 enrolled dengue cases and a true effect size of 50% (RR=0.5), contamination of 3 or 6 untreated clusters is expected to result in a ~7% and ~14% loss of statistical power, respectively.
6, 38	Additional per protocol analysis	An additional per-protocol analysis will be conducted in which the WEI is calculated using only the cluster-level *Wolbachia* prevalence in the participant's cluster of residence (in the month of participant enrolment), ignoring the participant's recent travel history. This recognises that dengue exposure risk may be higher at home versus other locations, rather than assuming an even distribution of exposure risk across daytime hours and locations visited
40	Methods added to address existing secondary objective #3	The impact of *Wolbachia* deployment on DHF case notifications will be evaluated using an interrupted time series analysis of monthly DHF notifications by kelurahan, before and after *Wolbachia* releases (January 2006 – December 2020). Methods will be developed and validated to classify area-level *Wolbachia* exposure status in a way that aligns with the administrative (kelurahan) boundaries by which dengue cases are reported.
46	Change to planned interim analysis	An interim analysis was originally planned for the mid-point of the study, ie after enrolment of 500 dengue cases from an initial target sample size of 1000. The re-estimation of statistical power described in this protocol amendment indicates that the trial will be adequately powered even with a smaller sample size, and the threshold of 500 cases is unlikely to be reached. The multinomial sampling method used in the power re-estimation means it is no longer necessary to re-calculate sample size using the observed inter-cluster distribution of participants, as originally stated.
47	Two new trial stopping rules	Additional criteria for early termination of the trial have been proposed in response to increasing *Wolbachia* contamination in several untreated clusters. The first rule addresses the possibility that

		a loss of power due to contamination may compromise the ITT analysis, and would see the trial stop if 5 or more untreated clusters are classified as contaminated. A cluster will be defined as contaminated when the cluster level *Wolbachia* frequency is >50% for 2 monthly monitoring events within a 6-month rolling window *and* >50% of the BG traps in the cluster have detected *Wolbachia* during those monitoring events. The second rule addresses the potential for consistently low rates of virologically-confirmed dengue case enrolment to make continued recruitment until November 2020 futile in terms of increasing sample size (statistical power), and would see the trial stop if five or fewer VCD cases are enrolled in any 3-month rolling window (commencing 1 Nov 2019).

Statistical Analysis Plan

Applying *Wolbachia* to Eliminate Dengue (AWED): A non-blinded cluster randomised controlled trial to assess the efficacy of *Wolbachia*-infected mosquito deployments to reduce dengue incidence in Yogyakarta, Indonesia

Table of Contents

1. Objectives

1.1. Primary Objective

To assess the efficacy of community-based deployments of *Wolbachia*-infected *Ae. aegypti* mosquitoes in reducing the incidence of symptomatic, virologically-confirmed dengue cases of any severity in Yogyakarta residents aged 3-45 years in release (intervention) areas, relative to non-release (untreated) areas.

1.2. Secondary Objectives

- To measure the efficacy of the *Wolbachia* method against each of the four DENV serotypes.
- To measure the efficacy of the *Wolbachia* method in reducing the incidence of symptomatic virologically-confirmed Zika virus and chikungunya virus infection in intervention areas, relative to untreated areas, and
- To quantify the impact of *Wolbachia* deployments on notifications of dengue haemorrhagic fever (DHF) cases to the Yogyakarta district health office

2. Study Design

2.1. Type of Study

The AWED trial is a parallel two-arm non-blinded cluster randomised controlled trial conducted in a single site in Yogyakarta City, Indonesia. The study site was subdivided into twenty-four contiguous clusters, approximately 1km^2 in size (range 0.7km^2-1.65km^2). Clusters were randomly allocated in a 1-to-1 ratio to receive *Wolbachia* deployments or no intervention, such that 12 clusters received *Wolbachia* deployments and 12 received no intervention (see **Figure 1**).

There are no buffer areas between clusters, but natural borders were used to define cluster boundaries as much as possible, to limit the spatial spread of *Wolbachia* from intervention clusters into untreated areas, and of wild-type mosquitoes in *Wolbachia*-treated clusters. Exclusion areas were minimised, and any areas within the study site where releases were not possible for reasons of logistics, public acceptance or absence of mosquito populations were pre-specified prior to randomisation and balanced between study arms. No attempt is made

to alter the routine dengue prevention and vector control activities conducted by public and private agencies throughout the study area (intervention and untreated clusters). The capacity of the disease surveillance system to detect (and thus respond to) dengue has been enhanced across the city through increased availability of diagnostic kits, which have been supplied to primary care clinics and hospitals since March 2016 by the World Mosquito Program (previously Eliminate Dengue Project) Indonesia, to support efforts to enhance the surveillance of dengue across Yogyakarta.

The impact of *Wolbachia* deployments on dengue incidence will be assessed by comparing the exposure distribution (probability of living in a *Wolbachia*-treated area) among virologically-confirmed dengue cases presenting to a network of public primary clinics (Puskesmas), against the exposure distribution among patients with febrile illness of non-arboviral aetiology presenting to the same network of clinics in the same temporal windows. Dengue cases and arbovirus-negative controls are sampled concurrently from within the population of patients presenting with febrile illness to the study clinic network, with case or control status classified retrospectively based on the results of laboratory diagnostic testing. By recruiting participants from within the population of patients presenting to clinics with febrile illness – with dengue test-positive patients classified as cases and test-negative patients classified as controls – the controls are necessarily drawn from the same source population as the cases, thus avoiding common pitfalls that can introduce selection bias [1]. In this situation, the odds ratio is an unbiased estimate of the rate ratio in the source population over the period of participant enrolment (the 'risk' period), without the need for any rare disease assumption [2,3].

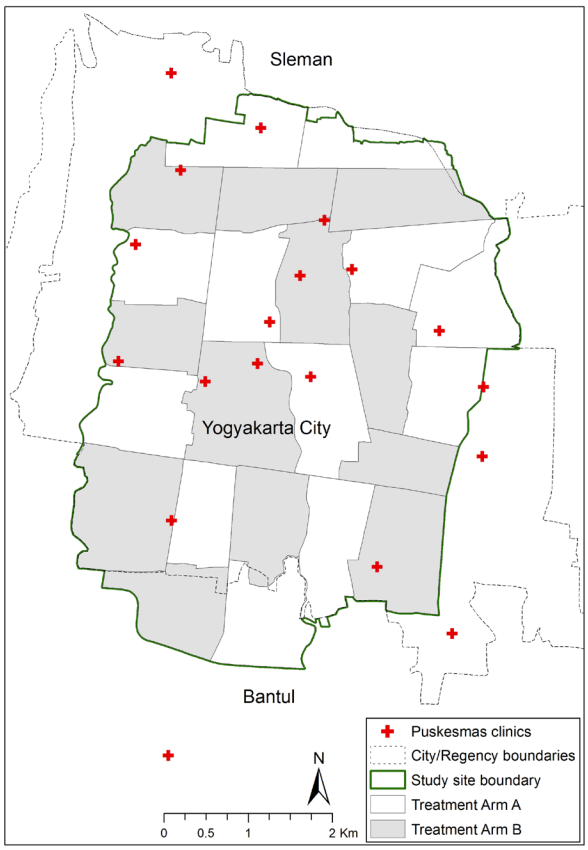


Figure 1. Map of study area, cluster boundaries, and Puskesmas clinics. The study area is outlined in green. The 12 clusters in each treatment arm are shown in grey and white. The location of the Puskesmas clinics at which trial recruitment is conducted are shown by red crosses.

2.2. Study Participants

The study population for measurement of the efficacy endpoint is the population of patients resident in the study area, presenting to the network of participating Puskesmas with febrile illness, and meeting the eligibility criteria as described in **Table 1**. Based on two years of historic data collated from the network of health clinics in the study area, it was estimated that at least 5000 patients per year present to these clinics with febrile illness (range 200-1500 per clinic per annum). We will enroll all participants presenting to any of the participating clinics who meet the eligibility criteria. Following laboratory testing and classification of participants' diagnostic status, all cases and those controls enrolled within the same calendar month as any case will be retained in the dataset for analysis.

Table 1. Participant eligibility criteria

Inclusion criteria	Exclusion criteria
1. Fever (either self-reported or objectively measured, e.g. tympanic membrane temperature ≥37.5ºC) with a date of onset between 1-4 days prior to the day of presentation.	1. Localising features suggestive of a specific diagnosis other than an arboviral infection, e.g. severe diarrhea, otitis, pneumonia.
2. Aged between 3-45 years old.	2. Prior enrollment in the study within the previous 4 weeks.
3. Resided in the study area every night for the 10 days preceding illness onset.	

2.3. Expected Duration of Study

The clinic-based sampling of febrile patients commenced in pilot phase in September 2017, with active enrolment in all clinics by December 2017. Enrolment will continue for up to 36 months, unless early termination is recommended by the independent data monitoring committee (IDMC).

3. Analysis Endpoints

3.1. Primary Efficacy Endpoint: Dengue

The primary outcome measure will be virologically-confirmed dengue virus infection in patients reporting febrile illness. Participants will be classified as dengue cases for the primary analysis if plasma samples collected 1-4 days after onset of fever test positive for dengue virus nucleic acid by RT-qPCR and/or dengue virus NS1 antigen (BioRad Platelia NS1 ELISA) (see **Figure 2**).

3.2. Secondary Efficacy Endpoint: DENV serotype-specific

For each participant who tests positive for dengue by RT-qPCR, the infecting serotype will be determined by DENV serotype-specific RT-PCR, and participants with a known serotype will be included in a secondary analysis to estimate serotype-specific efficacy, as described in section 11.3.1.

3.3. Secondary Efficacy Endpoints: Chikungunya and Zika

Secondary outcome measures include chikungunya and Zika virus infection in patients reporting febrile illness. Participants will be classified as virologically-confirmed chikungunya cases if chikungunya nucleic acid is detected in plasma samples by RT-qPCR (see **Figure 2**). Participants will be classified as virologically-confirmed Zika virus cases if Zika virus nucleic acid is detected in plasma samples by RT-qPCR.

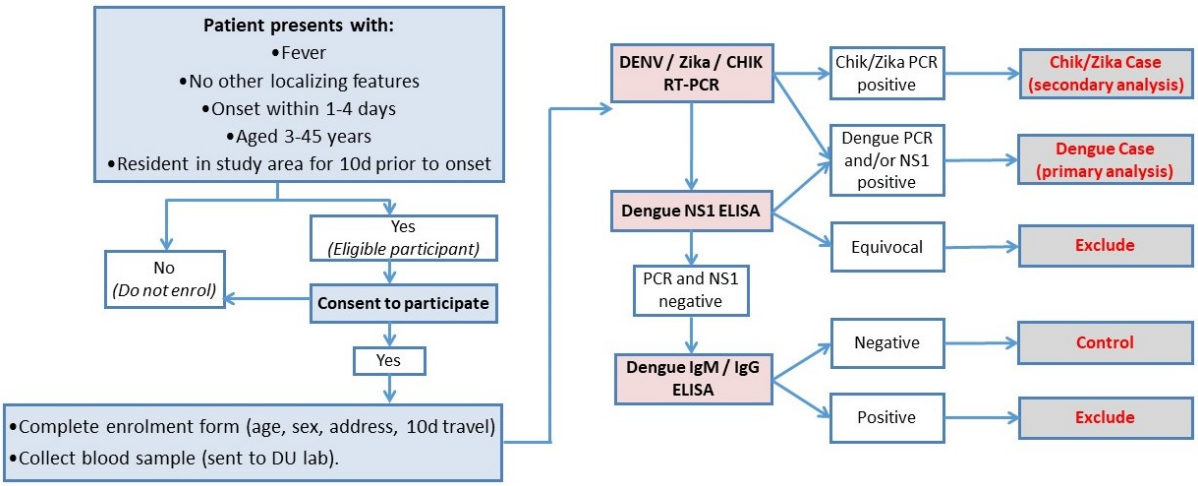


Figure 2. Flowchart of data and sample collection procedures and diagnostic algorithm.

4. Monitoring of *Wolbachia* prevalence in local *Ae. aegypti* populations

A network of BG-Sentinel adult mosquito traps (BioGents) has been in place throughout intervention and untreated clusters for the duration of the trial, evenly spaced throughout residential areas at a density of approximately 16 traps/km^2. BG traps are serviced weekly, with trapped mosquitoes screened for *Wolbachia* at weekly intervals during releases, fortnightly intervals after completion of releases, and monthly intervals since *Wolbachia* establishment (≥80% prevalence for two consecutive screening events). Mosquitoes are bio-banked in the intervening weeks when screening is not done. Trapped mosquitoes are identified by microscopy, and individual *Ae.aegypti* mosquitoes (male and female) are screened using quantitative PCR to detect the presence of *Wolbachia* and to confirm the species as *Ae. aegypti*.

5. Monitoring of unintended adverse effects of *Wolbachia* releases

In order to demonstrate that the deployment is not associated with any excess of a severe adverse outcome, we follow up all enrolled participants by telephone within 14 to 21 days

post-enrolment to ascertain their health status, recorded categorically as recovered/died, and whether or not they were ever hospitalised during this illness. Any death of a study participant within 14 to 21 days of enrolment is classified as a serious adverse event (SAE).

6. Sample Size Estimation

It was initially estimated that enrolment of approximately 1000 cases plus four times as many controls would be sufficient to detect a 50% reduction in dengue incidence with 80% power. Simulations were used to estimate the power to detect a range of intervention effect sizes, assuming 12 clusters per arm, a total of 1000 true dengue cases enrolled and 4000 non-dengue controls.

A re-estimation of sample size requirements was conducted in January 2019 after one year of recruitment. The initial power calculation used 1000 dengue cases and 4000 non-dengue controls allocated to each cluster based on historical proportions of dengue cases and other febrile illnesses, assuming no variation in the proportion of cases by cluster. This method was found to overestimate power for small samples by not taking into account randomness in the sampling. The sample size re-estimation included power estimates for 200, 400, 600, 800 and 1000 dengue cases with 4 times as many controls allocated to each cluster by sampling from a multinomial distribution, which incorporated added randomness by allowing the proportion of cases allocated to each cluster to vary across simulations. The re-estimation found that 400 dengue cases plus four times as many controls would be sufficient to detect a 50% reduction in dengue incidence with 80% power.

Additional simulations were conducted in September 2019 to assess the potential impact on power if a number of untreated clusters were 'lost' to *Wolbachia* contamination. For the target minimum observed effect size of 50% (Relative Risk (RR)=0.5) and 400 enrolled dengue cases, contamination of 3 untreated clusters (assuming that contaminated clusters experience the full intervention effect for 1 out of the 3 years of trial recruitment) is expected to result in a ~7% loss of power, and contamination of 6 clusters to result in a ~14% loss of power.

7. Statistical Analysis Method

7.1. General Considerations

This SAP was developed on the information provided in AWED Protocol version 5.1 dated 16 October 2019.

All statistical analyses will be generated using Stata version 14.0 or higher, or R (R Foundation for Statistical Computing, Austria).

A blinded data review will be conducted to assess the accuracy and completeness of the study database, prior to unblinding of the cluster intervention allocations. The appropriateness of planned statistical analyses will be assessed on a blinded set of 1000 observations comprised of exposure and demographic data from 1000 randomly selected participants combined with diagnostic results from a separate 1000 randomly selected participants. Exposure information and diagnostic results are stored in separate tables within the database. By merging exposure and outcome information from different randomly selected sets of 1000 participants we aim to avoid accidental unblinding of the data.

7.2. Analysis Sets

The dataset for analysis will retain all enrolled virologically-confirmed dengue cases, and all test-negative controls that are matched to a case by calendar month of enrolment. Unmatched controls will not be used for the primary analysis.

The analysis will be performed on data acquired during the case surveillance period, that is the period commencing when *Wolbachia* is deemed to have been established throughout intervention clusters, defined as one month after completion of releases in the last cluster (i.e. 8 January 2018). Cases and controls enrolled prior to 8 January 2018 will be excluded from the analysis dataset.

7.3. Status of potential participants

The status of all potential participants that were screened for enrolment will be summarized descriptively, according to the following categories, overall and by treatment arm:

- Number screened

- Number of screened patients that met eligibility criteria
- Number of eligible patients that consented to participate
- Number of consenting participants enrolled in the trial
- Number of enrolled participants successfully followed up for safety endpoints
- Number of enrolled participants for whom a blood sample was available for diagnostic testing
- Number of enrolled participants included in datasets for ITT and PP analysis

7.4. Demographic Characteristics

Participants' age and sex will be summarized descriptively overall, and by treatment arm, diagnostic category, inclusion/exclusion from analysis, and follow-up status.

7.5. Analysis Plan for Primary Efficacy Endpoint

Intention-to-Treat Analysis

The intention-to-treat (primary) analysis will consider *Wolbachia* exposure as a binary classification based on residence in a cluster allocated to *Wolbachia* deployment or not. Residence will be defined as the primary place of residence during the 10 days prior to illness onset.

The intervention effect will be estimated from an aggregate odds ratio comparing the exposure odds (residence in a *Wolbachia*-treated cluster) among test-positive cases versus test-negative controls (for data aggregated across all clusters), using the constrained permutation distribution as the foundation for inference. The null hypothesis is that the odds of residence in a *Wolbachia*-treated cluster is the same among test-positive cases as test-negative controls. The resulting odds ratio provides an unbiased estimate of the RR providing that the key assumptions underlying the TND are upheld (i.e. that test-negative controls are allowed to include participants who may test positive for dengue at any other time during the study period, and the distribution of non-dengue febrile illness is not associated with the intervention status). To note, since the constrained permutation distribution used for statistical inference contains only the 247 potential allocations (494 distinct randomisations) that meet all balancing criteria, the most extreme odds ratio in the distribution would carry a two-sided p-value of ~0.004 (1/494*2). Therefore, $p<0.004$ is the minimum threshold at

which statistical significance can be evaluated in this design. An exploratory analysis will estimate the intervention effect over time, by calculating the aggregate odds ratio at 12 months and 24 months into the ITT case surveillance period based on the cumulative test-positive cases and test-negative controls enrolled up to that point in time.

An additional group-level analysis will be performed using a cluster-level summary measure of the proportion of test-positive individuals amongst all tested individuals in each cluster. The difference in the average proportion of test positives between the intervention clusters and untreated clusters will be used to test the null hypothesis of no intervention effect using the t-test statistic but basing inference on the exact permutation distribution. These average proportions in each arm can be used to derive an estimate of the RR of dengue in treated versus untreated clusters, which is a much more intuitive effect measure, using a method described in detail elsewhere [4]. Briefly, we can substitute the estimated difference in the proportions, *d* into the formula $d = \frac{1}{1+(\frac{r}{2})(1+RR)} - \frac{RR}{RR+(\frac{r}{2})(1+RR)}$, where *r* is the simply the ratio of the total number of test negatives to the total number of test positives, which yields a quadratic equation for the unknown *RR.* Only one solution is plausible so that this then yields an estimate of *RR*, along with the appropriately transformed confidence interval (from that associated with *d*).

Per-protocol analysis

The per-protocol analysis will consider *Wolbachia* exposure as a quantitative index based on measured *Wolbachia* prevalence in local *Ae. aegypti* mosquitoes in the participant's cluster of residence, and in locations visited by the participant during the period 3-10 days prior to illness onset. The per-protocol analysis therefore allows for *Wolbachia* exposure to vary in a location over time, and also accounts for human mobility, in terms of the exposure-time that individuals spend outside their cluster of residence as reported in the travel history interview at enrolment. This analysis can also account for the temporal matching of dengue cases and test-negative controls: risk sets of cases and controls will be defined by frequency matching enrolled confirmed dengue cases to arbovirus-negative controls enrolled in the same calendar month.

Participants are asked about their mobility during the ten days prior to illness onset using a structured interview administered at enrolment. This records the duration of time spent at

home, work or school, and other locations visited during daylight hours (5am – 9pm) in the ten-day period. The geographic coordinates of those locations are derived by geo-locating them on a digital map, with the assistance of the respondent. A weighted '*Wolbachia* exposure index' (WEI) will be defined for each participant, as follows. The aggregate *Wolbachia* prevalence for each cluster will be calculated each month from all *Ae. aegypti* trapped in that cluster. For any calendar month where mosquito collection was not done, the average of the cluster-level *Wolbachia* prevalence in the one previous and one subsequent month will be used. The WEI for each participant will then be calculated by multiplying the cluster-level Wolbachia prevalence (in the calendar month of participant enrolment) at each of the locations visited, by the proportion of time spent at each location, to give a value on a continuous scale from 0 to 1. For visited locations within the quasi-experimental study area, the measured kelurahan-level *Wolbachia* prevalence from the screening event closest in time to the participant's enrolment will be used. Visited locations outside of both the AWED study area and the quasi-experimental study area will be assumed to have a *Wolbachia* prevalence of zero. The process of calculating WEI will be conducted blinded to participants' case/control status, by partitioning the travel history data from the laboratory diagnostic data, to remove any possibility of observer bias.

An additional per-protocol analysis will be conducted in which the WEI is calculated using only the cluster-level *Wolbachia* prevalence in the participant's cluster of residence (in the calendar month of participant enrolment), ignoring the participant's recent travel history. This recognises that dengue exposure risk may be higher at home versus other locations, rather than assuming an even distribution of exposure risk across daytime hours and locations visited.

Cases and controls will be classified by strata of their WEI: 0-<0.2; 0.2-<0.4; 0.4-<0.6; 0.6-<0.8; and 0.8-1. This acknowledges that the WEI is not a highly precise measure, and serves to reduce error in exposure classification. The ITT methods described above will be extended to allow for this individual level covariate using a regression approach [5], adjusted for time. A mixed effects logistic regression model will be fitted, incorporating time as random effect and with another random effect for cluster membership. Such models yield an estimate, and associated confidence interval, for the relative risk. The WEI strata will first be included as an ordinal covariate and the slope of the WEI variable will be tested for a difference from

zero. The WEI strata will additionally be included as a nominal (unordered) covariate to calculate stratum-specific IRRs (relative to the baseline 0-<0.2 stratum). This will allow examination of a 'dose response' relationship. An additional benefit of including WEI as a nominal variable is that it avoids any assumption of linearity in the dose response relationship.

7.6. Analysis of Secondary Efficacy Endpoints

DENV serotype-specific efficacy of Wolbachia *deployment*

In laboratory experiments, the degree to which *Wolbachia* reduces the DENV transmission potential of *Ae. aegypti* is dependent on the infecting virus serotype, with DENV1 transmission least affected [6]. A secondary analysis will estimate the serotype-specific efficacy of *Wolbachia* deployments in reducing symptomatic dengue virus infection with a known infecting serotype, for each of the four serotypes in turn, or as many as are detected in the study population. The same intention-to-treat and per-protocol analyses will be used as described for the primary endpoint above, with case populations restricted to each of the DENV serotypes in turn, and with the same control population as for the primary analysis.

Impact of Wolbachia deployment on Zika and chikungunya

There exists no baseline data on the prevalence of Zika or chikungunya infection among febrile patients presenting to primary health care clinics in Yogyakarta City, from which to estimate the expected number of cases; therefore, these secondary analyses are exploratory only and not subject to any formal sample size or power calculations. Blood samples from enrolled participants will be tested by Zika and chikungunya PCR for the purpose of defining arbovirus-negative controls for the primary analysis, as described above. These results will permit estimation of the prevalence of virologically confirmed Zika virus and chikungunya virus infection among the study population of ambulatory febrile patients presenting to primary health care.

If ≥20 virologically confirmed Zika or chikungunya cases are detected, a secondary analysis will estimate the efficacy of *Wolbachia* deployments in reducing the incidence of symptomatic virologically confirmed Zika virus and chikungunya virus infection. The same enrolled patient population will be used to analyse all three arbovirus endpoints (dengue, Zika and chikungunya), and the same intention-to-treat and per-protocol analyses will be used as

described for the primary (dengue) endpoint above. For Zika and chikungunya, the cases will be defined as enrolled participants who test positive by Zika or chikungunya PCR, respectively, and the controls will be those who test negative to all three arboviruses. Cases and controls will be matched by month of enrolment, as described above. If <20 cases of either Zika or chikungunya are detected there will be no formal analysis, only a descriptive analysis of the temporal and spatial distribution of cases.

Impact of Wolbachia deployment on notified dengue cases

The existing system for routine notification of dengue cases in Yogyakarta City is based on hospital-reporting of cases diagnosed clinically as Dengue Hemorrhagic Fever (DHF), which historically have not been accompanied by supportive laboratory testing. Since March 2016, hospitals have been encouraged to record a serological testing result, where available, on the report form, and also to report cases diagnosed clinically as Dengue Fever where there is a confirmatory NS1-positive test result. A separate reporting system, established in March 2016, collates data on the number of NS1 rapid tests performed – and number positive – in Puskesmas across the city. Both of these reporting systems include address information for notified cases.

We will collate data from these two reporting systems on a monthly basis, aggregated by kelurahan of residence, to monitor trends in reported dengue incidence across the city and by kelurahan, before, during and after *Wolbachia* deployment.

The impact of *Wolbachia* deployment on DHF case notifications will be evaluated using an interrupted time series analysis of monthly DHF notifications by kelurahan, before, during and after *Wolbachia* releases. Methods will be developed and validated *a priori* to classify area-level *Wolbachia* exposure status in a way that aligns with the kelurahan boundaries by which dengue cases are reported. A separate statistical analysis plan will be developed for this endpoint and the results will be reported in a secondary publication, subsequent to the publication of the main trial results.

7.7. Monitoring of Safety Endpoints

The safety endpoints of hospitalisation and death will be summarised by treatment arm. Any difference in the distribution of these two safety endpoints between treatment arms will be

evaluated from an aggregate odds ratio comparing the exposure odds (residence in a *Wolbachia*-treated cluster) among those with versus without the endpoint (for data aggregated across all clusters), using the constrained permutation distribution as the foundation for inference, and from the relative risk of hospitalisation in the intervention versus untreated clusters, derived from a comparison between treatment arms of the mean proportion of hospitalised participants among total participants in each cluster. These analyses will be repeated among VCD cases only, to compare the distribution of hospitalisations of VCD cases between treatment arms.

7.8. Interim Analysis

The trial protocol states that an interim analysis will be conducted at the mid-point of the study, i.e. after enrolment of 500 dengue cases with an initial target sample size of 1000. Re-estimation of statistical power conducted in January 2019 showed that the trial has 80% power to detect a reduction in dengue incidence greater than or equal to 50%, for a minimum sample of 400 virologically-confirmed dengue cases. This finding demonstrates that the trial is likely to be adequately powered even though it will not reach the original target of 1000 dengue cases prior to its revised completion date in August 2020. The original plan of conducting an interim analysis after enrolment of 500 dengue cases is therefore no longer appropriate. The IDMC and Trial Steering Committee decided in November 2019 that no interim analysis will be done for this study.

8. Differences between protocol and SAP

Differences between the protocol and the SAP listed in the table below will be resolved in the next protocol amendment (version 6).

SAP section, page	Text in SAP	Difference from protocol
Figure 2, page 7	All blood samples are tested by RT-PCR and NS1	Only samples that are PCR negative for dengue, chikungunya and Zika are subsequently tested using NS1.

7.5, page 12	This records the duration of time spent at home, work or school, and other locations visited during daylight hours (5am – 9pm) in the ten-day period.	Protocol stated '…and up to three other most-visited locations…', but in practice all locations visited for ≥1 hour were recorded.
7.2, page 9	The same analysis dataset will be used for ITT and PP analysis, restricted to cases and controls enrolled from one month after the completion of releases (i.e. 8 Jan 2018).	In the protocol, the PP analysis dataset includes all cases and controls enrolled from the start of full clinic enrollment. In practice controls from Dec 2017 would be excluded due to no cases, so the only difference from ITT would be inclusion of participants enrolled 1–7 Jan 2018. For simplicity, align PP dataset with ITT dataset.
7.2, page 9	The dataset for analysis will retain all enrolled virologically-confirmed dengue cases, and all test-negative controls that are matched to a case by calendar month of enrolment.	The protocol states that cases and controls will be matched on calendar month of illness onset.
7.5, page 11	This analysis can also account for the temporal matching of dengue cases and test-negative controls: risk sets of cases and controls will be defined by frequency matching enrolled confirmed dengue cases to arbovirus-negative controls enrolled in the same calendar month.	
7.6, page 14	Cases and controls will be matched by month of enrolment, as described above.	

7.6, page 14	For the analysis of Zika and chikungunya secondary endpoints, added a caveat that if <20 cases of either disease are detected then no formal analysis will be undertaken, only a descriptive analysis of the temporal and spatial distribution of cases.	
7.8, page 15	This finding demonstrates that the trial is likely to be adequately powered even though it will not reach the original target of 1000 dengue cases prior to its revised completion date in August 2020.	The protocol states that the revised completion date is November 2020.

9. References

1. De Serres G, Skowronski DM, Wu XW, Ambrose CS. The test-negative design: validity, accuracy and precision of vaccine efficacy estimates compared to the gold standard of randomised placebo-controlled clinical trials. Euro Surveill 2013;18.
2. Greenland S, Thomas DC. On the need for the rare disease assumption in case-control studies. Am J Epidemiol 1982;116:547-53.
3. Vandenbroucke JP, Pearce N. Case-control studies: basic concepts. Int J Epidemiol 2012;41:1480-9.
4. Jewell NP, Dufault S, Cutcher Z, Simmons CP, Anders KL. Analysis of cluster-randomized test-negative designs: cluster-level methods. Biostatistics 2019;20:332-46.
5. Small DS, Ten Have TR, Rosenbaum PR. Randomization Inference in a Group–Randomized Trial of Treatments for Depression. Journal of the American Statistical Association 2008;103:271-9.
6. Ferguson NM, Kien DT, Clapham H, et al. Modeling the impact on virus transmission of Wolbachia-mediated blocking of dengue virus infection of Aedes aegypti. Sci Transl Med 2015;7:279ra37.

Statistical Analysis Plan

Applying *Wolbachia* to Eliminate Dengue (AWED): A non-blinded cluster randomised controlled trial to assess the efficacy of *Wolbachia*-infected mosquito deployments to reduce dengue incidence in Yogyakarta, Indonesia

Table of Contents

1. Objectives

1.1. Primary Objective

To assess the efficacy of community-based deployments of *Wolbachia*-infected *Ae. aegypti* mosquitoes in reducing the incidence of symptomatic, virologically-confirmed dengue cases of any severity in Yogyakarta residents aged 3-45 years in release (intervention) areas, relative to non-release (untreated) areas.

1.2. Secondary Objectives

- To measure the efficacy of the *Wolbachia* method against each of the four DENV serotypes.
- To measure the efficacy of the *Wolbachia* method in reducing the incidence of symptomatic virologically-confirmed Zika virus and chikungunya virus infection in intervention areas, relative to untreated areas, and
- To quantify the impact of *Wolbachia* deployments on notifications of dengue haemorrhagic fever (DHF) cases to the Yogyakarta district health office

2. Study Design

2.1. Type of Study

The AWED trial is a parallel two-arm non-blinded cluster randomised controlled trial conducted in a single site in Yogyakarta City, Indonesia. The study site was subdivided into twenty-four contiguous clusters, approximately 1km^2 in size (range 0.7km^2-1.65km^2). Clusters were randomly allocated in a 1-to-1 ratio to receive *Wolbachia* deployments or no intervention, such that 12 clusters received *Wolbachia* deployments and 12 received no intervention (see **Figure 1**).

There are no buffer areas between clusters, but natural borders were used to define cluster boundaries as much as possible, to limit the spatial spread of *Wolbachia* from intervention clusters into untreated areas, and of wild-type mosquitoes in *Wolbachia*-treated clusters. Exclusion areas were minimised, and any areas within the study site where releases were not possible for reasons of logistics, public acceptance or absence of mosquito populations were pre-specified prior to randomisation and balanced between study arms. No attempt is made

to alter the routine dengue prevention and vector control activities conducted by public and private agencies throughout the study area (intervention and untreated clusters). The capacity of the disease surveillance system to detect (and thus respond to) dengue has been enhanced across the city through increased availability of diagnostic kits, which have been supplied to primary care clinics and hospitals since March 2016 by the World Mosquito Program (previously Eliminate Dengue Project) Indonesia, to support efforts to enhance the surveillance of dengue across Yogyakarta.

The impact of *Wolbachia* deployments on dengue incidence will be assessed by comparing the exposure distribution (probability of living in a *Wolbachia*-treated area) among virologically-confirmed dengue cases presenting to a network of public primary clinics (Puskesmas), against the exposure distribution among patients with febrile illness of non-arboviral aetiology presenting to the same network of clinics in the same temporal windows. Dengue cases and arbovirus-negative controls are sampled concurrently from within the population of patients presenting with febrile illness to the study clinic network, with case or control status classified retrospectively based on the results of laboratory diagnostic testing. By recruiting participants from within the population of patients presenting to clinics with febrile illness – with dengue test-positive patients classified as cases and test-negative patients classified as controls – the controls are necessarily drawn from the same source population as the cases, thus avoiding common pitfalls that can introduce selection bias [1]. In this situation, the odds ratio is an unbiased estimate of the rate ratio in the source population over the period of participant enrolment (the 'risk' period), without the need for any rare disease assumption [2,3].

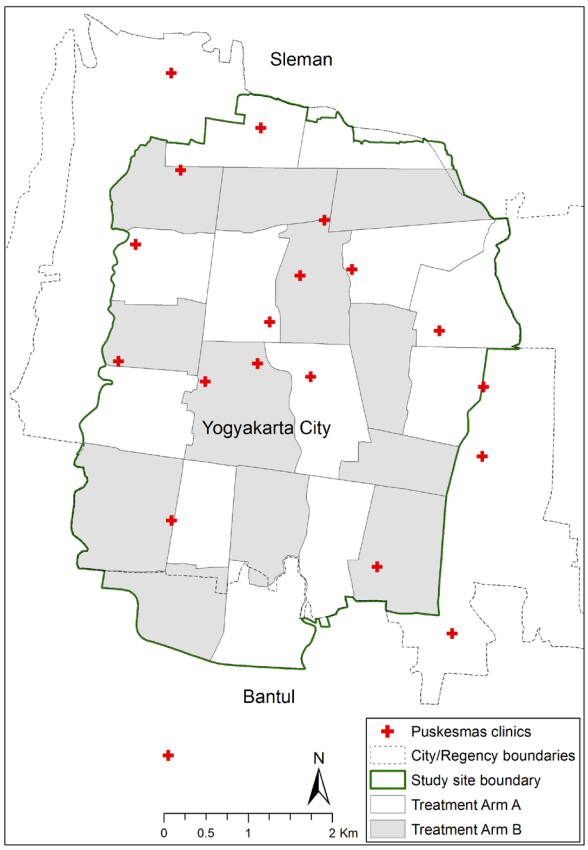


Figure 1. Map of study area, cluster boundaries, and Puskesmas clinics. The study area is outlined in green. The 12 clusters in each treatment arm are shown in grey and white. The location of the Puskesmas clinics at which trial recruitment is conducted are shown by red crosses.

2.2. Study Participants

The study population for measurement of the efficacy endpoint is the population of patients resident in the study area, presenting to the network of participating Puskesmas with febrile illness, and meeting the eligibility criteria as described in **Table 1**. Based on two years of historic data collated from the network of health clinics in the study area, it was estimated that at least 5000 patients per year present to these clinics with febrile illness (range 200-1500 per clinic per annum). We will enroll all participants presenting to any of the participating clinics who meet the eligibility criteria. Following laboratory testing and classification of participants' diagnostic status, all cases and those controls enrolled within the same calendar month as any case will be retained in the dataset for analysis.

Table 1. Participant eligibility criteria

Inclusion criteria	Exclusion criteria
1. Fever (either self-reported or objectively measured, e.g. tympanic membrane temperature ≥37.5ºC) with a date of onset between 1-4 days prior to the day of presentation.	1. Localising features suggestive of a specific diagnosis other than an arboviral infection, e.g. severe diarrhea, otitis, pneumonia.
2. Aged between 3-45 years old.	2. Prior enrollment in the study within the previous 4 weeks.
3. Resided in the study area every night for the 10 days preceding illness onset.	

2.3. Expected Duration of Study

The clinic-based sampling of febrile patients commenced in pilot phase in September 2017, with active enrolment in all clinics by December 2017. Enrolment will continue for up to 36 months, unless early termination is recommended by the independent data monitoring committee (IDMC).

3. Analysis Endpoints

3.1. Primary Efficacy Endpoint: Dengue

The primary outcome measure will be virologically-confirmed dengue virus infection in patients reporting febrile illness. Participants will be classified as dengue cases for the primary analysis if plasma samples collected 1-4 days after onset of fever test positive for dengue virus nucleic acid by RT-qPCR and/or dengue virus NS1 antigen (BioRad Platelia NS1 ELISA) (see **Figure 2**). A predefined exploratory analysis will evaluate hospitalised virologically-confirmed dengue cases as an outcome measure (a pragmatic proxy indicator for disease severity).

3.2. Secondary Efficacy Endpoint: DENV serotype-specific

For each participant who tests positive for dengue by RT-qPCR, the infecting serotype will be determined by DENV serotype-specific RT-PCR, and participants with a known serotype

will be included in a secondary analysis to estimate serotype-specific efficacy, as described in section 11.3.1.

3.3. Secondary Efficacy Endpoints: Chikungunya and Zika

Secondary outcome measures include chikungunya and Zika virus infection in patients reporting febrile illness. Participants will be classified as virologically-confirmed chikungunya cases if chikungunya nucleic acid is detected in plasma samples by RT-qPCR (see **Figure 2**). Participants will be classified as virologically-confirmed Zika virus cases if Zika virus nucleic acid is detected in plasma samples by RT-qPCR.

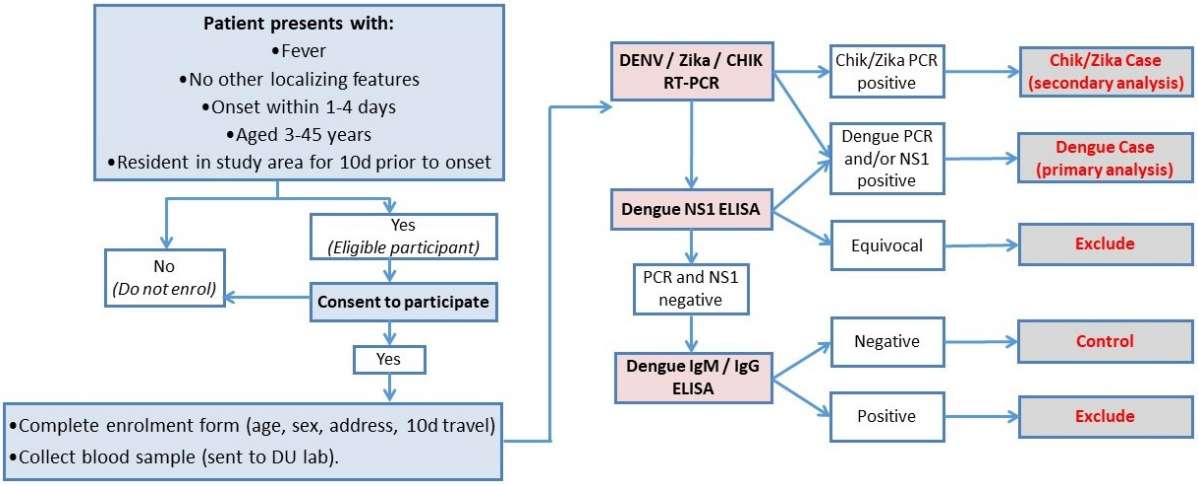


Figure 2. Flowchart of data and sample collection procedures and diagnostic algorithm.

4. Monitoring of *Wolbachia* prevalence in local *Ae. aegypti* populations

A network of BG-Sentinel adult mosquito traps (BioGents) has been in place throughout intervention and untreated clusters for the duration of the trial, evenly spaced throughout residential areas at a density of approximately 16 traps/km^2. BG traps are serviced weekly, with trapped mosquitoes screened for *Wolbachia* at weekly intervals during releases, fortnightly intervals after completion of releases, and monthly intervals since *Wolbachia* establishment (≥80% prevalence for two consecutive screening events). Mosquitoes are bio-banked in the intervening weeks when screening is not done. Trapped mosquitoes are identified by microscopy, and individual *Ae.aegypti* mosquitoes (male and female) are screened using quantitative PCR to detect the presence of *Wolbachia* and to confirm the species as *Ae. aegypti*.

5. Monitoring of unintended adverse effects of *Wolbachia* releases

In order to demonstrate that the deployment is not associated with any excess of a severe adverse outcome, we follow up all enrolled participants by telephone within 14 to 21 days post-enrolment to ascertain their health status, recorded categorically as recovered/died, and whether or not they were ever hospitalised during this illness. Any death of a study participant within 14 to 21 days of enrolment is classified as a serious adverse event (SAE).

6. Sample Size Estimation

It was initially estimated that enrolment of approximately 1000 cases plus four times as many controls would be sufficient to detect a 50% reduction in dengue incidence with 80% power. Simulations were used to estimate the power to detect a range of intervention effect sizes, assuming 12 clusters per arm, a total of 1000 true dengue cases enrolled and 4000 non-dengue controls.

A re-estimation of sample size requirements was conducted in January 2019 after one year of recruitment. The initial power calculation used 1000 dengue cases and 4000 non-dengue controls allocated to each cluster based on historical proportions of dengue cases and other febrile illnesses, assuming no variation in the proportion of cases by cluster. This method was found to overestimate power for small samples by not taking into account randomness in the sampling. The sample size re-estimation included power estimates for 200, 400, 600, 800 and 1000 dengue cases with 4 times as many controls allocated to each cluster by sampling from a multinomial distribution, which incorporated added randomness by allowing the proportion of cases allocated to each cluster to vary across simulations. The re-estimation found that 400 dengue cases plus four times as many controls would be sufficient to detect a 50% reduction in dengue incidence with 80% power.

Additional simulations were conducted in September 2019 to assess the potential impact on power if a number of untreated clusters were ‘lost’ to *Wolbachia* contamination. For the target minimum observed effect size of 50% (Relative Risk (RR)=0.5) and 400 enrolled dengue cases, contamination of 3 untreated clusters (assuming that contaminated clusters experience the full intervention effect for 1 out of the 3 years of trial recruitment) is expected

to result in a ~7% loss of power, and contamination of 6 clusters to result in a ~14% loss of power.

7. Statistical Analysis Method

7.1. General Considerations

This SAP was developed on the information provided in AWED Protocol version 5.1 dated 16 October 2019.

All statistical analyses will be generated using Stata version 14.0 or higher, or R (R Foundation for Statistical Computing, Austria).

A blinded data review will be conducted to assess the accuracy and completeness of the study database, prior to unblinding of the cluster intervention allocations. The appropriateness of planned statistical analyses will be assessed on a blinded set of 1000 observations comprised of exposure and demographic data from 1000 randomly selected participants combined with diagnostic results from a separate 1000 randomly selected participants. Exposure information and diagnostic results are stored in separate tables within the database. By merging exposure and outcome information from different randomly selected sets of 1000 participants we aim to avoid accidental unblinding of the data.

7.2. Analysis Sets

The dataset for analysis will retain all enrolled virologically-confirmed dengue cases, and all test-negative controls that are matched to a case by calendar month of enrolment. Unmatched controls will not be used for the primary analysis.

The analysis will be performed on data acquired during the case surveillance period, that is the period commencing when *Wolbachia* is deemed to have been established throughout intervention clusters, defined as one month after completion of releases in the last cluster (i.e. 8 January 2018). Cases and controls enrolled prior to 8 January 2018 will be excluded from the analysis dataset.

7.3. Status of potential participants

The status of all potential participants that were screened for enrolment will be summarized descriptively, according to the following categories, overall and by treatment arm:

- Number screened
- Number of screened patients that met eligibility criteria
- Number of eligible patients that consented to participate
- Number of consenting participants enrolled in the trial
- Number of enrolled participants successfully followed up for safety endpoints
- Number of enrolled participants for whom a blood sample was available for diagnostic testing
- Number of enrolled participants included in datasets for ITT and PP analysis

7.4. Demographic Characteristics

Participants' age and sex will be summarized descriptively overall, and by treatment arm, diagnostic category, inclusion/exclusion from analysis, and follow-up status.

7.5. Analysis Plan for Primary Efficacy Endpoint

Intention-to-Treat Analysis

The intention-to-treat (primary) analysis will consider *Wolbachia* exposure as a binary classification based on residence in a cluster allocated to *Wolbachia* deployment or not. Residence will be defined as the primary place of residence during the 10 days prior to illness onset.

The intervention effect will be estimated from an aggregate odds ratio comparing the exposure odds (residence in a *Wolbachia*-treated cluster) among test-positive cases versus test-negative controls (for data aggregated across all clusters), using the constrained permutation distribution as the foundation for inference. The null hypothesis is that the odds of residence in a *Wolbachia*-treated cluster is the same among test-positive cases as test-negative controls. The resulting odds ratio provides an unbiased estimate of the RR providing that the key assumptions underlying the TND are upheld (i.e. that test-negative controls are allowed to include participants who may test positive for dengue at any other time during the study period, and the distribution of non-dengue febrile illness is not associated with the

intervention status). To note, since the constrained permutation distribution used for statistical inference contains only the 247 potential allocations (494 distinct randomisations) that meet all balancing criteria, the most extreme odds ratio in the distribution would carry a two-sided p-value of ~0.004 (1/494*2). Therefore, p<0.004 is the minimum threshold at which statistical significance can be evaluated in this design. An exploratory analysis will estimate the intervention effect over time, by calculating the aggregate odds ratio at 12 months and 24 months into the ITT case surveillance period based on the cumulative test-positive cases and test-negative controls enrolled up to that point in time. Efficacy of the intervention will be calculated as 100*(1-aggregate odds ratio). For clarity in reporting of study results, primacy will be given to the aggregate odds ratio approach.

An additional group-level analysis will be performed using a cluster-level summary measure of the proportion of test-positive individuals amongst all tested individuals in each cluster. The difference in the average proportion of test positives between the intervention clusters and untreated clusters will be used to test the null hypothesis of no intervention effect using the t-test statistic but basing inference on the exact permutation distribution. These average proportions in each arm can be used to derive an estimate of the RR of dengue in treated versus untreated clusters, which is a much more intuitive effect measure, using a method described in detail elsewhere [4]. Briefly, we can substitute the estimated difference in the proportions, *d* into the formula $d = \frac{1}{1+(\frac{r}{2})(1+RR)} - \frac{RR}{RR+(\frac{r}{2})(1+RR)}$, where *r* is the simply the ratio of the total number of test negatives to the total number of test positives, which yields a quadratic equation for the unknown *RR.* Only one solution is plausible so that this then yields an estimate of *RR*, along with the appropriately transformed confidence interval (from that associated with *d*).

Per-protocol analysis

The per-protocol analysis will consider *Wolbachia* exposure as a quantitative index based on measured *Wolbachia* prevalence in local *Ae. aegypti* mosquitoes in the participant's cluster of residence, and in locations visited by the participant during the period 3-10 days prior to illness onset. The per-protocol analysis therefore allows for *Wolbachia* exposure to vary in a location over time, and also accounts for human mobility, in terms of the exposure-time that individuals spend outside their cluster of residence as reported in the travel history interview at enrolment. This analysis can also account for the temporal matching of dengue cases and

test-negative controls: risk sets of cases and controls will be defined by frequency matching enrolled confirmed dengue cases to arbovirus-negative controls enrolled in the same calendar month.

Participants are asked about their mobility during the ten days prior to illness onset using a structured interview administered at enrolment. This records the duration of time spent at home, work or school, and other locations visited during daylight hours (5am – 9pm) in the ten-day period. The geographic coordinates of those locations are derived by geo-locating them on a digital map, with the assistance of the respondent. A weighted '*Wolbachia* exposure index' (WEI) will be defined for each participant, as follows. The aggregate *Wolbachia* prevalence for each cluster will be calculated each month from all *Ae. aegypti* trapped in that cluster. For any calendar month where mosquito collection was not done, the average of the cluster-level *Wolbachia* prevalence in the one previous and one subsequent month will be used. The WEI for each participant will then be calculated by multiplying the cluster-level Wolbachia prevalence (in the calendar month of participant enrolment) at each of the locations visited, by the proportion of time spent at each location, to give a value on a continuous scale from 0 to 1. For visited locations within the quasi-experimental study area, the measured kelurahan-level *Wolbachia* prevalence from the screening event closest in time to the participant's enrolment will be used. Visited locations outside of both the AWED study area and the quasi-experimental study area will be assumed to have a *Wolbachia* prevalence of zero. The process of calculating WEI will be conducted blinded to participants' case/control status, by partitioning the travel history data from the laboratory diagnostic data, to remove any possibility of observer bias.

An additional per-protocol analysis will be conducted in which the WEI is calculated using only the cluster-level *Wolbachia* prevalence in the participant's cluster of residence (in the calendar month of participant enrolment), ignoring the participant's recent travel history. This recognises that dengue exposure risk may be higher at home versus other locations, rather than assuming an even distribution of exposure risk across daytime hours and locations visited.

Cases and controls will be classified by strata of their WEI: 0-<0.2; 0.2-<0.4; 0.4-<0.6; 0.6-<0.8; and 0.8-1. This acknowledges that the WEI is not a highly precise measure, and serves

to reduce error in exposure classification. The ITT methods described above will be extended to allow for this individual level covariate using a regression approach [5], adjusted for time. A mixed effects logistic regression model will be fitted, incorporating time as random effect and with another random effect for cluster membership. Such models yield an estimate, and associated confidence interval, for the relative risk. Efficacy will then be calculated as 100*(1-RR). The WEI strata will first be included as an ordinal covariate and the slope of the WEI variable will be tested for a difference from zero. The WEI strata will additionally be included as a nominal (unordered) covariate to calculate stratum-specific IRRs (relative to the baseline 0-<0.2 stratum). This will allow examination of a 'dose response' relationship. An additional benefit of including WEI as a nominal variable is that it avoids any assumption of linearity in the dose response relationship.

7.6. Analysis of Secondary Efficacy Endpoints

DENV serotype-specific efficacy of* Wolbachia *deployment

In laboratory experiments, the degree to which *Wolbachia* reduces the DENV transmission potential of *Ae. aegypti* is dependent on the infecting virus serotype, with DENV1 transmission least affected [6]. A secondary analysis will estimate the serotype-specific efficacy of *Wolbachia* deployments in reducing symptomatic dengue virus infection with a known infecting serotype, for each of the four serotypes in turn, or as many as are detected in the study population. The same intention-to-treat and per-protocol analyses will be used as described for the primary endpoint above, with case populations restricted to each of the DENV serotypes in turn, and with the same control population as for the primary analysis.

Impact of Wolbachia deployment on Zika and chikungunya

There exists no baseline data on the prevalence of Zika or chikungunya infection among febrile patients presenting to primary health care clinics in Yogyakarta City, from which to estimate the expected number of cases; therefore, these secondary analyses are exploratory only and not subject to any formal sample size or power calculations. Blood samples from enrolled participants will be tested by Zika and chikungunya PCR for the purpose of defining arbovirus-negative controls for the primary analysis, as described above. These results will permit estimation of the prevalence of virologically confirmed Zika virus and chikungunya virus infection among the study population of ambulatory febrile patients presenting to primary health care.

If ≥20 virologically confirmed Zika or chikungunya cases are detected, a secondary analysis will estimate the efficacy of *Wolbachia* deployments in reducing the incidence of symptomatic virologically confirmed Zika virus and chikungunya virus infection. The same enrolled patient population will be used to analyse all three arbovirus endpoints (dengue, Zika and chikungunya), and the same intention-to-treat and per-protocol analyses will be used as described for the primary (dengue) endpoint above. For Zika and chikungunya, the cases will be defined as enrolled participants who test positive by Zika or chikungunya PCR, respectively, and the controls will be those who test negative to all three arboviruses. Cases and controls will be matched by month of enrolment, as described above. If <20 cases of either Zika or chikungunya are detected there will be no formal analysis, only a descriptive analysis of the temporal and spatial distribution of cases.

Impact of Wolbachia deployment on notified dengue cases

The existing system for routine notification of dengue cases in Yogyakarta City is based on hospital-reporting of cases diagnosed clinically as Dengue Hemorrhagic Fever (DHF), which historically have not been accompanied by supportive laboratory testing. Since March 2016, hospitals have been encouraged to record a serological testing result, where available, on the report form, and also to report cases diagnosed clinically as Dengue Fever where there is a confirmatory NS1-positive test result. A separate reporting system, established in March 2016, collates data on the number of NS1 rapid tests performed – and number positive – in Puskesmas across the city. Both of these reporting systems include address information for notified cases.

We will collate data from these two reporting systems on a monthly basis, aggregated by kelurahan of residence, to monitor trends in reported dengue incidence across the city and by kelurahan, before, during and after *Wolbachia* deployment.

The impact of *Wolbachia* deployment on DHF case notifications will be evaluated using an interrupted time series analysis of monthly DHF notifications by kelurahan, before, during and after *Wolbachia* releases. Methods will be developed and validated *a priori* to classify area-level *Wolbachia* exposure status in a way that aligns with the kelurahan boundaries by which dengue cases are reported. A separate statistical analysis plan will be developed for

this endpoint and the results will be reported in a secondary publication, subsequent to the publication of the main trial results.

7.7. Monitoring of Safety Endpoints

The safety endpoints of hospitalisation and death will be summarised by treatment arm. Any difference in the distribution of these two safety endpoints between treatment arms will be evaluated from an aggregate odds ratio comparing the exposure odds (residence in a *Wolbachia*-treated cluster) among those with versus without the endpoint (for data aggregated across all clusters), using the constrained permutation distribution as the foundation for inference, and from the relative risk of hospitalisation in the intervention versus untreated clusters, derived from a comparison between treatment arms of the mean proportion of hospitalised participants among total participants in each cluster. These analyses will be repeated among VCD cases only, to compare the distribution of hospitalisations of VCD cases between treatment arms.

7.8. Interim Analysis

The trial protocol states that an interim analysis will be conducted at the mid-point of the study, i.e. after enrolment of 500 dengue cases with an initial target sample size of 1000. Re-estimation of statistical power conducted in January 2019 showed that the trial has 80% power to detect a reduction in dengue incidence greater than or equal to 50%, for a minimum sample of 400 virologically-confirmed dengue cases. This finding demonstrates that the trial is likely to be adequately powered even though it will not reach the original target of 1000 dengue cases prior to its revised completion date in August 2020. The original plan of conducting an interim analysis after enrolment of 500 dengue cases is therefore no longer appropriate. The IDMC and Trial Steering Committee decided in November 2019 that no interim analysis will be done for this study.

8. Differences between protocol and SAP

Differences between the approved protocol (version 5.1) and the SAP are listed in the table below.

SAP section, page	Text in SAP	Difference from protocol
Figure 2, page 7	All blood samples are tested by RT-PCR and NS1	Only samples that are PCR negative for dengue, chikungunya and Zika are subsequently tested using NS1.
7.5, page 12	This records the duration of time spent at home, work or school, and other locations visited during daylight hours (5am – 9pm) in the ten-day period.	Protocol stated '…and up to three other most-visited locations…', but in practice all locations visited for ≥1 hour were recorded.
7.2, page 9	The same analysis dataset will be used for ITT and PP analysis, restricted to cases and controls enrolled from one month after the completion of releases (i.e. 8 Jan 2018).	In the protocol, the PP analysis dataset includes all cases and controls enrolled from the start of full clinic enrollment. In practice controls from Dec 2017 would be excluded due to no cases, so the only difference from ITT would be inclusion of participants enrolled 1–7 Jan 2018. For simplicity, align PP dataset with ITT dataset.
7.2, page 9	The dataset for analysis will retain all enrolled virologically-confirmed dengue cases, and all test-negative controls that are matched to a case by calendar month of enrolment.	The protocol states that cases and controls will be matched on calendar month of illness onset.
7.5, page 11	This analysis can also account for the temporal matching of dengue cases and test-negative controls: risk sets of cases and controls will be defined by frequency matching enrolled confirmed dengue cases to arbovirus-	

	negative controls enrolled in the same calendar month.	
7.6, page 14	Cases and controls will be matched by month of enrolment, as described above.	
7.6, page 14	For the analysis of Zika and chikungunya secondary endpoints, added a caveat that if <20 cases of either disease are detected then no formal analysis will be undertaken, only a descriptive analysis of the temporal and spatial distribution of cases.	
7.8, page 15	This finding demonstrates that the trial is likely to be adequately powered even though it will not reach the original target of 1000 dengue cases prior to its revised completion date in August 2020.	The protocol states that the revised completion date is November 2020.

9. References

1. De Serres G, Skowronski DM, Wu XW, Ambrose CS. The test-negative design: validity, accuracy and precision of vaccine efficacy estimates compared to the gold standard of randomised placebo-controlled clinical trials. Euro Surveill 2013;18.
2. Greenland S, Thomas DC. On the need for the rare disease assumption in case-control studies. Am J Epidemiol 1982;116:547-53.
3. Vandenbroucke JP, Pearce N. Case-control studies: basic concepts. Int J Epidemiol 2012;41:1480-9.
4. Jewell NP, Dufault S, Cutcher Z, Simmons CP, Anders KL. Analysis of cluster-randomized test-negative designs: cluster-level methods. Biostatistics 2019;20:332-46.
5. Small DS, Ten Have TR, Rosenbaum PR. Randomization Inference in a Group–Randomized Trial of Treatments for Depression. Journal of the American Statistical Association 2008;103:271-9.
6. Ferguson NM, Kien DT, Clapham H, et al. Modeling the impact on virus transmission of Wolbachia-mediated blocking of dengue virus infection of Aedes aegypti. Sci Transl Med 2015;7:279ra37.

Summary of SAP amendments

Changes made from original SAP (version 1.3, 24 February 2020) to final SAP (version 1.6, 28 May 2020)		
Page	**Amendment**	**Comment**
6	Addition to primary efficacy endpoint	An exploratory analysis was added to evaluate efficacy against hospitalised virologically-confirmed dengue cases (a subset of the primary efficacy endpoint), as a pragmatic proxy indicator for disease severity.
11	Clarification to the intention-to-treat analysis	Made explicit that efficacy would be calculated as 100*(1-aggregate odds ratio), and that of the two ITT analyses described (arm-level aggregate odds ratio and cluster-level test-positive fraction) the aggregate odds ratio would be reported as the primary result.
13	Clarification to the per-protocol analysis	Made explicit that efficacy would be calculated as 100*(1-relative risk).