

1 **Real-time dynamic modelling for the design of a cluster-**  
2 **randomized phase 3 Ebola vaccine trial in Sierra Leone**

3

4 Camacho A<sup>\*1</sup>, Eggo RM<sup>\*1</sup>, Goeyvaerts N<sup>2</sup>, Vandebosch A<sup>2</sup>, Mogg R<sup>2</sup>, Funk S<sup>1</sup>, Kucharski  
5 AJ<sup>1</sup>, Watson CH<sup>1</sup>, Vangeneugden T<sup>2</sup>, Edmunds WJ<sup>1</sup>.

6

7 1. Department of Infectious Disease Epidemiology, London School of Hygiene &  
8 Tropical Medicine, Keppel Street, London. WC1E 7HT UK

9 2. Janssen Research & Development, Beerse, Belgium

10

11 \*these authors contributed equally

12

13 **Corresponding author:** Rosalind M Eggo (r.eggo@lshtm.ac.uk)

14

15

16

17 **Abstract**

18 **Background**

19 Declining incidence and spatial heterogeneity complicated the design of phase 3 Ebola  
20 vaccine trials during the tail of the 2013-16 Ebola virus disease (EVD) epidemic in West  
21 Africa. Mathematical models can provide forecasts of expected incidence through time  
22 and can account for both vaccine efficacy in participants and effectiveness in  
23 populations. Determining expected disease incidence was critical to calculating power  
24 and determining trial sample size.

25 **Methods**

26 In real-time, we fitted, forecasted, and simulated a proposed phase 3 cluster-  
27 randomised vaccine trial for a prime-boost EVD vaccine in three candidate regions in  
28 Sierra Leone. The aim was to forecast trial feasibility in these areas through time and  
29 guide study design planning.

30 **Results**

31 EVD incidence was highly variable during the epidemic, especially in the declining  
32 phase. Delays in trial start date were expected to greatly reduce the ability to discern an  
33 effect, particularly as a trial with an effective vaccine would cause the epidemic to go  
34 extinct more quickly in the vaccine arm. Real-time updates of the model allowed  
35 decision-makers to determine how trial feasibility changed with time.

## 36 **Conclusions**

37 This analysis was useful for vaccine trial planning because we simulated effectiveness as  
38 well as efficacy, which is possible with a dynamic transmission model. It contributed to  
39 decisions on choice of trial location and feasibility of the trial. Transmission models  
40 should be utilised as early as possible in the design process to provide mechanistic  
41 estimates of expected incidence, with which decisions about sample size, location,  
42 timing, and feasibility can be determined.

43

44 **Keywords:** Ebola vaccine; transmission modelling; trial design; phase 3; epidemic;  
45 collaboration

46

47

48

## 49 **Introduction**

50

51 West Africa experienced the largest outbreak of Ebola virus disease (EVD) to date  
52 during 2013-16. This epidemic resulted in more than 25,000 cases and 10,000 deaths.  
53 As the epidemic unfolded in 2014, development of candidate vaccines was accelerated,  
54 including evaluation in phase 1-2 studies and phase 3 planning. However, the rapidly  
55 changing incidence both geographically and in time posed major challenges to the  
56 design and planning of phase 3 trials. Typical study design calculations do not allow for  
57 varying infection rates within and between communities over time, which is especially  
58 problematic during the tail of an epidemic, when few cases occur. Computer simulations  
59 employing empirical statistical models can mitigate some of these concerns however  
60 they require accurate assumptions on incidence, heterogeneity and, in addition, do not  
61 capture the mechanism of an outbreak. Moreover, an effective vaccine used widely in a  
62 given area (as would be the case in large-scale, population-based vaccine trials) could in  
63 itself further reduce the incidence.

64 Dynamic models of EVD transmission were developed during the epidemic to  
65 understand the patterns of spread of the virus and predict the course of the outbreak  
66 [1-4]. If these models are appropriately parameterised and updated, then they can be  
67 used to predict incidence and how it may change in space and time [5]. In addition,  
68 dynamic models can account for both the direct and indirect effect of vaccine-induced  
69 immunity and its impact on the transmission dynamics. That is, they can be used to  
70 assess the extent that the trial itself may affect the transmission dynamics.

71 Collaboration between the Centre for Mathematical Modelling of Infectious Disease  
72 (CMMID) at the London School of Hygiene & Tropical Medicine and Janssen Research &  
73 Development (Janssen R&D) was established to rapidly extend a mathematical model of  
74 EVD [3] to simulate a cluster-randomized phase 3 vaccine trial in Sierra Leone. The  
75 dynamic model and trial simulations were updated in real-time to match the latest  
76 incidence data available. This collaboration thus enabled a real-time, dynamic  
77 assessment of the feasibility of a potential phase 3 trial, which ultimately was  
78 implemented as a safety and immunogenicity study: EBOVAC-Salone (NCT02509494).  
79 This paper describes how the model was used to inform the planning of the trial as well  
80 as the decision-making to abandon the effectiveness part of the protocol.

81

## 82 **Methods**

### 83 **Collaboration**

84 Collaboration was initiated between CMMID and Janssen R&D in February 2015. Janssen  
85 R&D was seeking a partner to guide study design and feasibility planning of a phase 3  
86 effectiveness trial for their heterologous prime-boost vaccine regimen (Ad26.ZEBOV as  
87 prime and MVA-BN®-Filo 28 days later as boost), for which phase 1 trials were on-  
88 going.

89

90 CMMID had previously developed mathematical models of EVD transmission to assess  
91 the potential for large outbreaks [6], impact of community care centres on the evolving  
92 epidemic [7], and bed capacity in Sierra Leone [3]. In addition, CMMID members liaised  
93 with WHO on the design and analysis of the WHO EVD vaccine trial [8].

94

95 Collaboration offered a unique opportunity to explore the use of a dynamic transmission  
96 model to evaluate study feasibility. In this paper, we present the model-based incidence  
97 projections and trial simulations from 15<sup>th</sup> February 2015, similar to those sent from  
98 LSHTM to the team at Janssen on a weekly basis from February 2015 to May 2015.  
99 These were in turn employed by the clinical study team to evaluate and guide power  
100 calculations, study design as well as trial feasibility. To illustrate the impact of the  
101 evolving epidemic, an update of the projections and simulations at the end of April is  
102 provided as supplementary materials.

103

### 104 **Vaccine trial design**

105 A large-scale cluster-randomized phase 3 trial was designed to evaluate the  
106 effectiveness of prime-boost vaccine regimen against laboratory-confirmed EVD in an  
107 outbreak setting in Sierra Leone [9]. Sierra Leone is administratively divided into  
108 districts, districts into chiefdoms, and chiefdoms into sections. A trial cluster would be a  
109 section. With vaccine availability at time of study design of up to four hundred thousand  
110 doses of both prime and boost vaccine, approximately 160 clusters of 5000 participants  
111 (800,000 in total) were to be assigned in a 1:1 ratio to immediate vaccination versus no  
112 vaccination (control), whereby vaccination would be offered to the control group after  
113 effectiveness was established.

114 Initially, feasibility, statistical power, and type I error of the trial were evaluated using  
115 simulations which assumed constant incidence through time [9]. Control incidence  
116 assumptions of 3, 5, 10, 20 and 40 EVD cases per arm per month (400,000 person-  
117 months) were evaluated, with allowance for heterogeneity between clusters based on  
118 CMMID projections and simulations. However, the rapidly changing epidemic dynamics  
119 in early 2015 meant that these static predictions were unlikely to capture the  
120 epidemiological picture.

121

## 122 **Transmission model for trial**

123 The transmission model extended a previously published model for transmission of EVD  
124 [3]. It was a stochastic compartmental model, where the population was divided into  
125 classes (Figure 1): Susceptible ( $S$ ), Exposed ( $E$ ), Infectious not yet notified ( $I$ ), Infectious  
126 and notified ( $J$ ) and Removed ( $R$ , for recovered and immune, or dead). The infectious  
127 compartment was split in two sub-compartments  $I$  and  $J$  in order to account for a delay  
128 of (on average) 4.8 days to notification of new cases [4]. The model was extended to  
129 mimic the trial design closely, but modelling cluster-level randomization was not  
130 possible because there was insufficient data available at this spatial scale for fitting. It is  
131 often difficult to predict the tails of epidemics, which are characterized by small, local  
132 outbreaks, and stochastic variation. Instead, we assumed a 1:1 randomization at the  
133 district level and treated the clusters as independent units.

134 Susceptible people were assumed to be recruited to the trial for the length of the accrual  
135 time,  $T_r$ , by entering either the vaccine ( $V_S$ ) or control ( $C_1$ ) arms. An average of 2 weeks  
136 after receiving the prime, vaccinated participants entered the compartment,  $V_P$ , where  
137 they were assumed to have a reduced risk of infection,  $\sigma_p$ . On receipt of the boost  
138 vaccine, they were assumed to enter  $V_B$ , and immediately gain the target vaccine  
139 efficacy,  $\sigma_b$  (Figure 1). Control participants were assumed to proceed from  $C_1$  to  $C_2$  at

140 the same rate as  $V_S$  to  $V_P$  to maintain comparability. Parameters that govern rates of  
 141 transition are given in Table 1. To account for external influences on transmission –  
 142 such as variation in human behavior and introduction of control measures – we  
 143 assumed that the transmission rate could change over time; the extent and direction of  
 144 change was estimated during the model fitting process [3]. Hypothetical vaccine efficacy  
 145 values were defined in February 2015 for the power calculation of the effectiveness  
 146 trial. These values were conservative estimates, chosen to ensure that the planned trial  
 147 would have sufficient power in the event of unpredicted changes in incidence, and to  
 148 decrease the risk of the study. These hypothetical assumptions are only working  
 149 hypotheses and do not necessarily reflect the potential effect of this candidate vaccine,  
 150 and these hypothetical values need to be assessed in the future.

151  
 152  
 153

Table 1. Parameters used in the model. Most values are fixed based on literature values, while transmission rate is estimated.

Parameter	Description	Value	Reference
$\beta_t$	Time varying effective contact rate	Estimated	Estimated
$\lambda_t$	Time varying force of infection	$\frac{\beta_t(I_r + J_r)}{N_r}, r \in (V, C)$	[1]
$1/\epsilon$	Average latent period	9.4 days	[1]
$1/v_1$	Average infectious period before notification	4.8 days	[1]
$1/v_2$	Average infectious period after notification	6.4 days	[1]
$R_t$	Time-varying reproduction number	$\beta_t^*(1/v_1 + 1/v_2)$	-
$N_r$	Total number of subjects recruited in each arm	170,000 (Kambia) 230,000 (Port Loko) 400,000 (Western Area)	Fixed - -
$T_r$	Accrual time	12 weeks	Fixed
$r_t$	Linear recruitment rate in each arm	$N_r/T_r$	-
$1/k_p$	Average time between prime vaccination and onset of protection	14 days	Fixed
$1/k_p + 1/k_b$	Average time between prime and boost vaccination	28 days	Fixed
$1/\gamma$	Average duration of vaccine protection	11 months	Fixed
$\sigma_p$	Hypothetical vaccine efficacy for prime vaccine	50%	Fixed
$\sigma_b$	Hypothetical vaccine efficacy for prime + boost	60%, 80%, 90%	Fixed

154  
 155  
 156

### Incidence data

157 The model was fitted to weekly confirmed and probable EVD incidence data from three  
 158 districts in Sierra Leone (Kambia, Port Loko, and Western Area) that had on-going  
 159 epidemics in February 2015 and were therefore candidate areas for a potential vaccine  
 160 trial. Data were drawn from the WHO and Sierra Leone situation reports and ran from  
 161 25<sup>th</sup> May 2014 until the date of fitting and forecast [10,11]. We used Bayesian methods

162 to fit the model to the data, namely particle Markov Chain Monte Carlo, which allows  
163 parameter estimation in a stochastic framework.

164

## 165 **Forecasting**

166 We sampled the reproduction number ( $R_t$ ) 5,000 times at the last fitted data point, and  
167 forecasted the epidemic until extinction under the assumption that the reproduction  
168 number did not change from that time. We retained only forecasts that went extinct by  
169 1<sup>st</sup> January 2016 because all regions showed waning epidemics, and although  
170 persistence for a further year was possible, it was deemed unlikely (Figure S2). Sampled  
171 reproduction numbers therefore usually lie below 1 (Figure S1). Updated estimates of  
172 the reproduction number distribution made in April 2015 have very little density above  
173 1, which suggests this was a reasonable assumption.

174 The forecasted persistence probability at each point of time  $t$  was defined as the  
175 probability that at least one infectious individual remains in the arm at that time, and  
176 was computed empirically by summing over the  $N$  forecast trajectories that went extinct  
177 by 1<sup>st</sup> January 2016:

$$P(t) = \frac{1}{N} \sum_{i=1}^N 1_{\{E_i(t)+I_i(t)>0\}}$$

178 with  $i$  the index of the sample from the posterior.

179 We evaluated the impact of trial start date by comparing the simulated number of cases  
180 in the vaccine or control arms. We used a one-sided Wilcoxon signed rank test to test for  
181 the pairwise difference in the total number of cases between the vaccine and control  
182 arm. In addition, we assessed the effect of vaccine efficacy on trial success. The  
183 persistence probability is calculated as the proportion of 5000 epidemic simulations  
184 that are non-extinct at each time point.

185

## 186 **Results**

### 187 **Model fits and projections**

188 Using data as of 15<sup>th</sup> February 2015, we fitted the model to weekly confirmed and  
189 probable EVD cases (Figure 2). At that time, the epidemic was in the tail phase, which is  
190 clear from the rapidly decreasing persistence probability. Visually, incidence data that  
191 has since been observed show excellent agreement with the forecasted epidemics  
192 (Figure 2), where 65%, 59%, and 65% of weekly values lie within the 50% credible  
193 interval (CI) in Kambia, Port Loko, and Western Area respectively. 92%, 92%, and 94%  
194 of points lie within the 95% CI.

195

196 Short-term model projections were contrasted with the static model assumptions from  
197 the power analysis of [9] (Figure 2). In [9] the control incidence of 5 per 400,000  
198 person-months was identified as a threshold of sufficient statistical power to initiate the  
199 trial. The model indicates that the incidence could drop below this threshold between  
200 June and August for all three districts, though the estimates are subject to great  
201 uncertainty as can be seen from the wide 95% credible intervals. Already in April-May,  
202 the incidence was possibly too low to initiate the trial.

203

204 We now describe the results of the vaccine trial simulations. The baseline scenario  
205 presented in this paper, unless otherwise specified, is that the trial began on 1<sup>st</sup> May  
206 2015 using a conservative working hypothesis of 50% reduction in susceptibility  
207 following prime vaccine, rising to hypothetical 80% after the boost vaccine. The  
208 populations of Kambia (population 340,000) and Port Loko (population 558,000) were  
209 smaller than the target of 400,000 participants in each arm, so we show simulations for  
210 those districts with each arm of the trial encompassing both – 170,000 in Kambia, and  
211 230,000 in Port Loko. Western Area had a large enough population that the trial could  
212 be conducted solely in that district.

213

#### 214 **Effect of the start date of the trial**

215 For forecasts made on 15<sup>th</sup> February 2015 with trial start dates on 1<sup>st</sup> May, 1<sup>st</sup> June and  
216 1<sup>st</sup> July 2015, there were fewer cases in both arms when the start date was later (Figure  
217 3), due to the continued decline of the epidemic. The later the start date, the lower the  
218 probability that the epidemic was still on-going by the start of the trial. Although the  
219 trial begins on the first of each month, delays involved in the recruitment of participants,  
220 onset of protective immunity, and time to boost vaccination mean that cases occurring  
221 in the population are not necessarily trial endpoints, and therefore do not accrue in the  
222 cumulative cases shown in Figure 3. While the difference between vaccine and control  
223 arms could have been large by December 2015, the probability of the epidemic  
224 persisting until December was very low. We found the overall number of cases was  
225 larger if the trial was in Kambia and Port Loko combined, compared to Western Area  
226 solely.

227

228 The occurrence of a vaccine trial in a candidate region would affect the persistence  
229 probability of the epidemic in that region if the vaccine were efficacious. In a declining  
230 epidemic, this would cause the epidemic to go extinct faster, reducing the persistence

231 probability. This can be seen from Figure 3, as starting the trial earlier increased the  
232 probability of earlier elimination in the vaccine arms.

233

234 Starting the trial later would result in a reduced probability of detecting a difference in  
235 the number of cases between the two arms and an increased probability of having no  
236 cases in either arm (Figure 4). In some simulations more cases would be observed in the  
237 control group in comparison to the vaccinated group (a “negative” effect). Figure 4  
238 indicates that over 5000 replicates of the trial, the distribution of the total number of  
239 cases is always significantly different between the vaccine and control arms at all  
240 starting time ( $p < 0.05$ , Wilcoxon signed rank test). However, when considering the  
241 proportion of trials with a positive or negative effect (i.e. below or above the diagonal in  
242 Fig 4), these proportions decrease and become more similar as the trial start later, with  
243 the proportion of trials with no cases increasing in the same time. While a split in favour  
244 of the vaccinated group would be expected under the assumed prime/boost effect, the  
245 simulations indicate that for a trial starting on 1<sup>st</sup> May, a negative (versus positive) effect  
246 would be observed in respectively 18% vs. 55%, 12% vs. 32% and 11% vs. 38% of  
247 simulations in Kambia, Port Loko and Western Area. For a trial starting 1<sup>st</sup> July 2015 the  
248 difference became smaller with 13% vs 15%, 4% vs 5% and 3% vs 4% of simulations for  
249 the three districts, respectively. The probability of observing no cases in either arm  
250 increased when the trial started later due to increasing stochastic extinctions  
251 (supplementary information).

252

### 253 **Effect of vaccine efficacy**

254 In simulations of this large trial, which is started in the declining phase of the epidemic,  
255 any effective vaccine causes a decrease in persistence probability of the epidemic  
256 (Figure 5). Higher hypothetical vaccine efficacy leads to quicker extinction of the  
257 epidemic, although the differences are very small. The simulations gave very  
258 comparable expected number of cases in the vaccine arm between the hypothetical  
259 vaccine efficacy values of 60%, 80% and 90%, which is shown by the grey, red, and  
260 yellow boxplots (Figure 5). This finding was consistent for Kambia, Port Loko and  
261 Western Area.

262

### 263 **Real-time updates**

264 The epidemic progressed from February to May 2015, and we updated the model fits  
265 and projections in real time. Here we display updated results as of 26<sup>th</sup> April 2015



266 (Figure 6). The persistence probability of the epidemic in the potential study areas  
267 changed as incidence in those areas decreased and the model was fitted to more  
268 available data. There was a sharp decrease in persistence probability for Port Loko, due  
269 to the decline in the epidemic occurring there. In contrast, for Kambia and Western  
270 Area, the projections from February changed very little by April.

271

272

## 273 **Discussion**

274 This close collaboration between CMMID and Janssen R&D in forecasting and planning a  
275 phase 3 Ebola vaccine trial had many key benefits: firstly, the production of up-to-date  
276 epidemic projections gave better situational awareness to the clinical study team and  
277 key decision makers at Janssen. These forecasts were based on fitting a mechanistic  
278 transmission model to the current epidemiological data, thereby providing rigorous and  
279 realistic predictions. Secondly, the mechanistic model provided a means by which to  
280 assess the feasibility of the phase 3 effectiveness trial, and how this changed through  
281 time. This is critical to trial planning, determining whether to proceed, and to  
282 understanding the effect of logistical delays or constraints on feasibility. And thirdly, by  
283 fitting mechanistic models to potential study regions individually, the forecasts provided  
284 a better understanding of the variability between candidate sites and the impact that a  
285 trial might have had on the epidemic. This allowed study team and decision makers to  
286 assess the relative probability of trial success based on geographically specific  
287 information.

288

289 Other trials planned for Ebola vaccines in various parts of West Africa faced challenges  
290 to feasibility as a result of the declining incidence [5,12]. In this study, modelling was  
291 used to help gauge the feasibility of the cluster-randomised design, by forecasting  
292 incidence in potential regions, which was then used in power calculations [9]. The  
293 dynamic transmission model could account for both vaccine efficacy in those vaccinated  
294 as well as vaccine effectiveness in the population. Trial simulations indicated how many  
295 cases to expect in the vaccine and control arms for various trial locations, start dates and  
296 hypotheses of vaccine efficacy and how this changed over time. The trial simulations  
297 thus guided decisions of trial location and feasibility. For example, the simulations  
298 indicated that Kambia was more likely to have sustained transmission compared to  
299 Western Area and Port Loko. Further, the rapid decrease of the persistence probability  
300 over time urged the vaccine development team at Janssen to explore alternative trial  
301 designs, and, partly as a result of this work, it was decided to abandon the effectiveness

302 trial protocol as planned. Instead, a safety and immunogenicity study was initiated in  
303 Kambia in October 2015.

304

305 The trial was intended to start during the declining phase of the outbreak and the  
306 assumption on potential start dates reflects realistic assumptions about operational  
307 timing. Additional work could explore the feasibility of the trial starting at earlier points  
308 of the outbreak. Importantly, this work was performed using conservative estimates of  
309 vaccine efficacy after prime and boost vaccination, which were used to calculate power  
310 of the effectiveness trial. The effect of different assumptions of potential effect of prime  
311 and boost vaccinations could be reassessed, also taking into account durability of  
312 protection as suggested by recent immunogenicity data in humans highlighting a strong  
313 and sustained immune response [13]. Future work could also integrate formal testing of  
314 vaccine efficacy [9] in the dynamic transmission modelling framework, in a similar way  
315 as for an individually randomized trial [5].

316

317 Although the model was able to accurately forecast the incidence by district, a key  
318 limitation was that we could not model cluster-level randomization due to insufficient  
319 data at this scale. Also, although the model could suggest which was likely to be the most  
320 favourable setting for the trial from an epidemiological point of view, in reality, there  
321 may be logistical constraints such as local customs, staff availability, manufacturing  
322 capacity and storage, trials running in parallel, and other factors that would affect trial  
323 feasibility. The scenarios that we explored were considered realistic at the start of the  
324 collaboration.

325

326 The model structure used here does not explicitly include different transmission  
327 mechanisms such as during unsafe burials [14,15]. Instead, we used a flexible,  
328 stochastic, transmission rate to capture the combined effect of these different  
329 transmission components. The precise contribution of different factors was likely to  
330 have changed over the course of the epidemic. For example, by November 2014, trained  
331 burial teams and a safe burial command centre were established in Kambia and Port  
332 Loko (having been previously established in Western Area) [16], so the risk of  
333 transmission due to unsafe practices was likely decreased by these interventions.  
334 However, the model was able to capture overall patterns of disease transmissions that  
335 occur as a result of changes in transmission routes.

336

337 While the use of mechanistic transmission models in evaluating vaccination programs is  
338 well established, their use in trial design, planning, and analysis, is a relatively new and  
339 growing area of research [17]. Designing interventions to reduce influenza transmission  
340 gives different preferred trial designs whether the goal is achieving power or taking  
341 account of economic constraints [18,19]. Modelling has been used to propose new trials  
342 for HIV antiviral treatment in serodiscordant couples [20], and has been used  
343 specifically for vaccine trials for malaria [21,22], intestinal helminths [23], wildlife  
344 vaccines [24], and nasopharyngeal bacteria [25]. For Ebola vaccine trials, a semi-  
345 mechanistic model developed during the epidemic addressed the feasibility of a  
346 proposed phase 3 trial in high risk individuals [12].

347

348 Our collaboration represents a novel example of close collaboration between modellers  
349 and trial planners to guide the design of a phase 3 trial during an epidemic. We  
350 delivered up-to-the-minute projections for both the epidemic and trial feasibility from  
351 academic researchers to industry partners. This type of information is critical to trial  
352 planning and clinical development, and mathematical models of disease transmission  
353 should be integrated into trial design at the earliest possible stage.

354

355

## 356 **Acknowledgements**

357 We acknowledge Stefan Flasche (LSHTM) and Benoit Callendret, Wim Parys, Guillermo  
358 Herrera-Taracena (Janssen) for valuable discussion.

359

## 360 **Funding source**

361 The authors acknowledge funding from the Innovative Medicines Initiative 2 (IMI2)  
362 Joint Undertaking, under grant agreement 115854 (RME, WJE), the Research for Health  
363 in Humanitarian Crises (R2HC) Programme, managed by Research for Humanitarian  
364 Assistance (Grant 13165) (AC, SF, AJK), and the Norwegian Institute of Public Health  
365 (CHW, WJE). IMI2 receives support from the European Union's Horizon 2020 research  
366 and innovation programme and the European Federation of Pharmaceutical Industries  
367 and Associations (EFPIA). The funders had no influence on the content of this analysis  
368 or decision to publish.

369

## 370 **Author contribution**

371 AC, SF, WJE, AJK developed the model. AC programmed the model and ran all  
372 simulations. AC, RME made the figures. AC, RME, SF, AJK, CHW, NG, TV, RM, AV, WJE  
373 analysed the outputs and interpreted the model. RME, AC, NG wrote the first draft of the  
374 paper. All authors contributed to the draft and have approved the final version.

375

## 376 **Conflict of interest**

377 AC, WJE and CHW have acted as unpaid advisors to the WHO on Ebola vaccination and  
378 report travel and accommodation paid for by the WHO to attend meetings. WJE is a co-  
379 investigator on, and RME is funded by, the European Commission Innovative Medicines  
380 Initiative-funded EBOVAC trial of the Johnson & Johnson prime-boost Ebola vaccine  
381 candidate. WJE's partner is an epidemiologist at GlaxoSmithKline, in a role unrelated to  
382 the company's development of an Ebola vaccine. AC and CHW have acted as unpaid  
383 advisors to the EBOVAC trial, for which CHW reports travel and accommodation paid for  
384 by the EBOVAC consortium to attend a meeting. NG, AV and TV are employees of Janssen  
385 R&D. RM was employee of Janssen R&D during the time this work was done and is now  
386 employee of Merck.

387

388

## 389 **Figure captions**

390

391 Figure 1. Model diagram. Susceptible people were recruited to the trial by entering either the vaccine  
392 ( $V_S$ ) or control ( $C_1$ ) arms. Two weeks after receiving the prime, vaccinated participants developed  
393 protective immunity ( $\sigma_p$ ), and entered the compartment  $V_p$ . On receipt of the boost vaccine, they  
394 enter  $V_B$ , and immediately gain the target vaccine efficacy ( $\sigma_b$ ). Control participants proceed from  $C_1$   
395 to  $C_2$  at the same rate as  $V_S$  to  $V_p$  to maintain comparability. Transitions have Erlang-distributed  
396 waiting periods with shape equal to 2, apart from  $S$  to  $E$ ,  $I$  to  $J$ , and  $J$  to  $R$ , which are exponentially  
397 distributed as well as  $S$  to  $V_S$  or  $C_1$ , which are step-wise processes.

398

399 Figure 2. Epidemic in Kambia, Port Loko, and Western Area, Sierra Leone. Upper panels: Time points  
400 marked by dotted lines correspond with simulated trial start dates; 1<sup>st</sup> May 2015, 1<sup>st</sup> June 2015, and  
401 1<sup>st</sup> of July 2015. Filled red circles are weekly EVD cases to which the model was fitted (blue line, with  
402 dark shaded region showing 50% credible interval and light region showing 95% interval) and empty  
403 circles displays data after that date (not fitted). Grey areas show forward-simulations of possible  
404 epidemic trajectories generated by the model, conditioned on extinction by 1 January 2016. Middle  
405 panels: Projections of the weekly number of reported cases rescaled to per 400,000 subjects.  
406 Horizontal dashed blue lines correspond to the static model incidence assumptions in [9], of 3, 5, 10,  
407 20, and 40 reported cases per 400,000 person-months. Lower panels: persistence probability in each  
408 area.

409

410 Figure 3. Effect of start date on number of cases in vaccine and control arms, and persistence  
411 probability in each region, stratified by start date of the trial, for the baseline scenario. Cumulative  
412 cases are only shown for trajectories that persist until that month. Where no boxplot is shown, all  
413 trajectories were extinct by that month.

414

415 Figure 4. Distribution of total cases observed in each arm of the trial, stratified by start date, for the  
416 baseline scenario. Note that the colour scale (which indicates the number of simulations) is

417 logarithmic. In simulations above the diagonal, more cases occurred in the vaccine arm. The p-values  
418 are from a one-sided Wilcoxon signed rank test for fewer cases in the vaccine arm. There is one  
419 simulated point not shown, where the number of cases in the control arm is 100, and in the vaccine  
420 arm is 93, which occurred in Port Loko.

421

422 Figure 5. Effect of vaccine efficacy on number of cases in vaccine and control arms and persistence  
423 probability, for a trial starting on 1<sup>st</sup> May 2015. Forecasts start on 15<sup>th</sup> February 2015. Cumulative  
424 cases are only shown for trajectories that persist until that month. When no boxplot is shown, this  
425 indicates that all trajectories were extinct by that month.

426

427 Figure 6. Updated estimates of persistence probability on 15<sup>th</sup> February, and 26<sup>th</sup> April 2015 in each  
428 potential trial region. Figures show simulated trial under the baseline scenario.

429

430

431

432

## References

433

[1] Agua-Agum J, Ariyaratnam A, Aylward B, Blake IM, Brennan R, Cori A, et al. West  
434 African Ebola epidemic after one year--slowing but not yet under control. *N Engl J*  
435 *Med* 2015;372:584–7. doi:10.1056/NEJMc1414992.

436

[2] Lewnard JA, Ndeffo Mbah ML, Alfaro-Murillo JA, Altice FL, Bawo L, Nyenswah TG,  
437 et al. Dynamics and control of Ebola virus transmission in Montserrado, Liberia: a  
438 mathematical modelling analysis. *Lancet Infect Dis* 2014;14:1189–95.  
439 doi:10.1016/S1473-3099(14)70995-8.

440

[3] Camacho A, Kucharski A, Aki-Sawyer Y, White MA, Flasche S, Baguelin M, et al.  
441 Temporal Changes in Ebola Transmission in Sierra Leone and Implications for  
442 Control Requirements: a Real-time Modelling Study. *PLoS Curr* 2015;7.  
443 doi:10.1371/currents.outbreaks.406ae55e83ec0b5193e30856b9235ed2.

444

[4] WHO Ebola Response Team. Ebola Virus Disease in West Africa — The First 9  
445 Months of the Epidemic and Forward Projections. *N Engl J Med* 2014;371:1481–  
446 95. doi:10.1056/NEJMoa1411100.

447

[5] Camacho A, Eggo RM, Funk S, Watson CH, Kucharski AJ, Edmunds WJ. Estimating  
448 the probability of demonstrating vaccine efficacy in the declining Ebola epidemic:  
449 a Bayesian modelling approach. *BMJ Open* 2015;5:e009346.  
450 doi:10.1136/bmjopen-2015-009346.

451

[6] Camacho A, Kucharski AJ, Funk S, Breman J, Piot P, Edmunds WJ. Potential for  
452 large outbreaks of Ebola virus disease. *Epidemics* 2014.  
453 doi:10.1016/j.epidem.2014.09.003.

454

[7] Kucharski AJ, Camacho A, Flasche S, Glover RE, Edmunds WJ, Funk S. Measuring  
455 the impact of Ebola control measures in Sierra Leone. *Proc Natl Acad Sci U S A*  
456 2015;112:14366–71. doi:10.1073/pnas.1508814112.

457

[8] Henao-Restrepo AM, Longini IM, Egger M, Dean NE, Edmunds WJ, Camacho A, et  
458 al. Efficacy and effectiveness of an rVSV-vectored vaccine expressing Ebola  
459 surface glycoprotein: interim results from the Guinea ring vaccination cluster-  
460 randomised trial. *Lancet* 2015;386:857–66. doi:10.1016/S0140-6736(15)61117-  
461 5.

462

[9] Vandebosch A, Mogg R, Goeyvaerts N, Truyers C, Greenwood B, Watson-Jones D,  
463 et al. Simulation-guided phase 3 trial design to evaluate vaccine effectiveness to  
464 prevent Ebola virus disease infection: Statistical considerations, design rationale,  
465 and challenges. *Clin Trials* 2016;13:57–65. doi:10.1177/1740774515621059.

466

[10] World Health Organisation. Ebola Situation Reports n.d.  
467 <http://apps.who.int/ebola/ebola-situation-reports> (accessed June 7, 2015).

468

[11] Leone R of S. Ministry of Health and Sanitation n.d.  
469 [http://health.gov.sl/?page\\_id=583](http://health.gov.sl/?page_id=583).

470

[12] Bellan SE, Pulliam JRC, Pearson CAB, Champredon D, Fox SJ, Skrip L, et al.

471 Statistical power and validity of Ebola vaccine trials in Sierra Leone: a simulation  
472 study of trial design and analysis. *Lancet Infect Dis* 2015;15:703–10.  
473 doi:10.1016/S1473-3099(15)70139-8.

474 [13] Milligan ID, Gibani MM, Sewell R, Clutterbuck EA, Campbell D, Plested E, et al.  
475 Safety and Immunogenicity of Novel Adenovirus Type 26- and Modified Vaccinia  
476 Ankara-Vectored Ebola Vaccines: A Randomized Clinical Trial. *JAMA*  
477 2016;315:1610–23. doi:10.1001/jama.2016.4218.

478 [14] LEGRAND J, GRAIS RF, BOELLE PY, VALLERON AJ, FLAHAULT A. Understanding  
479 the dynamics of Ebola epidemics. *Epidemiol Infect* 2007;135:610.  
480 doi:10.1017/S0950268806007217.

481 [15] Fang L-Q, Yang Y, Jiang J-F, Yao H-W, Kargbo D, Li X-L, et al. Transmission  
482 dynamics of Ebola virus disease and intervention effectiveness in Sierra Leone.  
483 *Proc Natl Acad Sci* 2016;113:4488–93. doi:10.1073/pnas.1518587113.

484 [16] Nielsen CF, Kidd S, R.M. SA, Edward D, Mermin J, H. KP. Improving Burial  
485 Practices and Cemetery Management During an Ebola Virus Disease Epidemic —  
486 Sierra Leone, 2014. *MMWR* 2015;64:20–7.

487 [17] Lessler J, Edmunds WJ, Halloran ME, Hollingsworth TD, Lloyd AL. Seven  
488 challenges for model-driven data collection in experimental and observational  
489 studies. *Epidemics* 2015;10:78–82. doi:10.1016/j.epidem.2014.12.002.

490 [18] Klick B, Leung GM, Cowling BJ. Optimal design of studies of influenza  
491 transmission in households. I: Case-ascertained studies. *Epidemiol Infect*  
492 2012;140:106–14. doi:10.1017/S0950268811000392.

493 [19] Klick B, Nishiura H, Leung GM, Cowling BJ. Optimal design of studies of influenza  
494 transmission in households. II: comparison between cohort and case-ascertained  
495 studies. *Epidemiol Infect* 2014;142:744–52. doi:10.1017/S0950268813001623.

496 [20] Cori A, Ayles H, Beyers N, Schaap A, Floyd S, Sabapathy K, et al. HPTN 071  
497 (PopART): a cluster-randomized trial of the population impact of an HIV  
498 combination prevention intervention including universal testing and treatment:  
499 mathematical model. *PLoS One* 2014;9:e84511.  
500 doi:10.1371/journal.pone.0084511.

501 [21] Valim C, Mezzetti M, Maguire J, Urdaneta M, Wypij D. Estimation of vaccine  
502 efficacy in a repeated measures study under heterogeneity of exposure or  
503 susceptibility to infection. *Philos Trans A Math Phys Eng Sci* 2008;366:2347–60.  
504 doi:10.1098/rsta.2008.0044.

505 [22] White MT, Griffin JT, Drakeley CJ, Ghani AC. Heterogeneity in malaria exposure  
506 and vaccine response: implications for the interpretation of vaccine efficacy  
507 trials. *Malar J* 2010;9:82. doi:10.1186/1475-2875-9-82.

508 [23] Alexander N, Cundill B, Sabatelli L, Bethony JM, Diemert D, Hotez P, et al.  
509 Selection and quantification of infection endpoints for trials of vaccines against  
510 intestinal helminths. *Vaccine* 2011;29:3686–94.  
511 doi:10.1016/j.vaccine.2011.03.026.

512 [24] Calenge C, Rossi S. Bayesian modelling of hunting data may improve the  
513 understanding of host-parasite systems: wild boar diseases and vaccination as an  
514 example. *J Theor Biol* 2014;343:32–43. doi:10.1016/j.jtbi.2013.11.011.

515 [25] Scott P, Herzog SA, Auranen K, Dagan R, Low N, Egger M, et al. Timing of bacterial  
516 carriage sampling in vaccine trials: a modelling study. *Epidemics* 2014;9:8–17.  
517 doi:10.1016/j.epidem.2014.08.003.

518  
519  
520

**Figure 1**

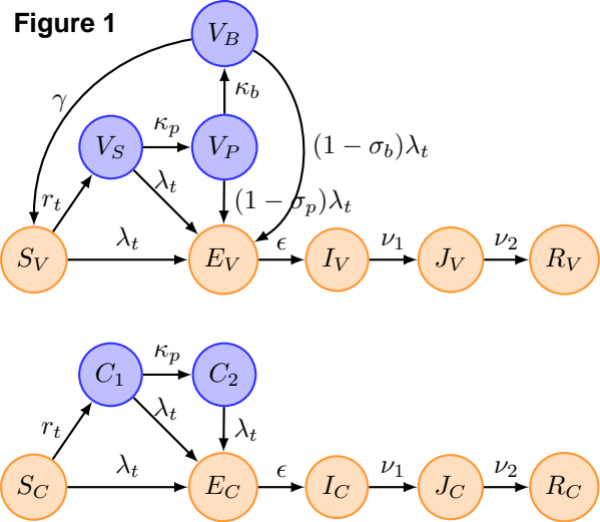


Figure 2

**Data**      **Model**      **Credible Intervals**  
● fitted   ○ not fitted   ■ fit   ■ forecast   ■ 50%   ■ 95%

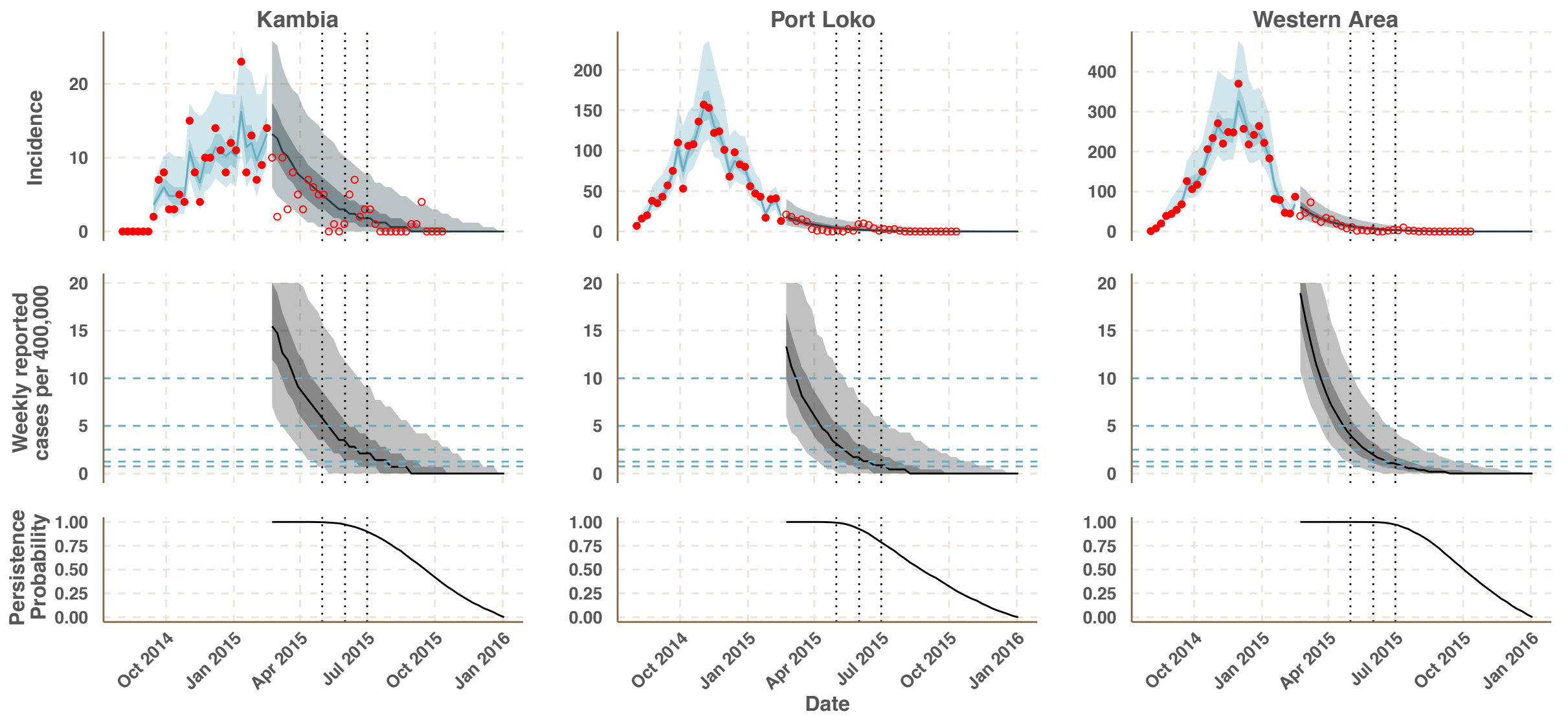




Figure 3

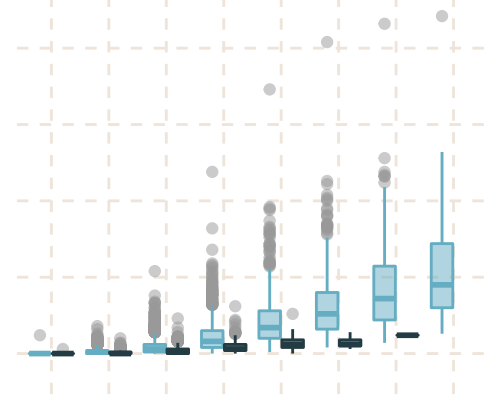
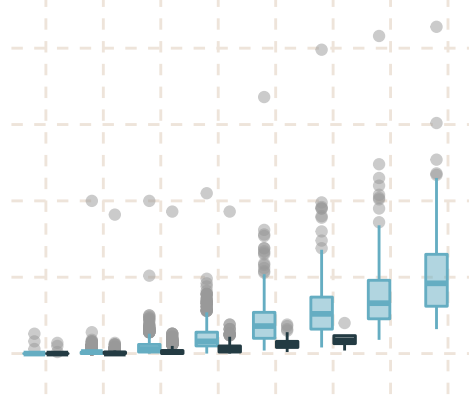
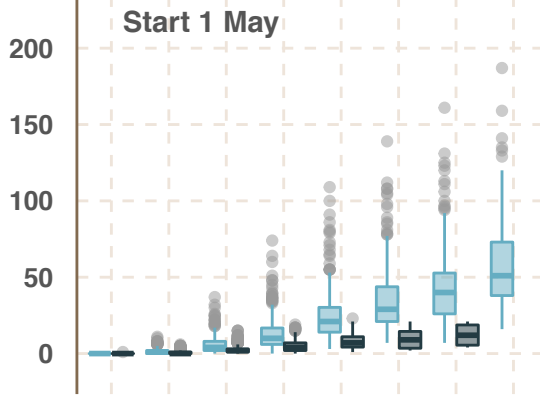
Arm control intervention 1 May intervention 1 June intervention 1 July

Kambia

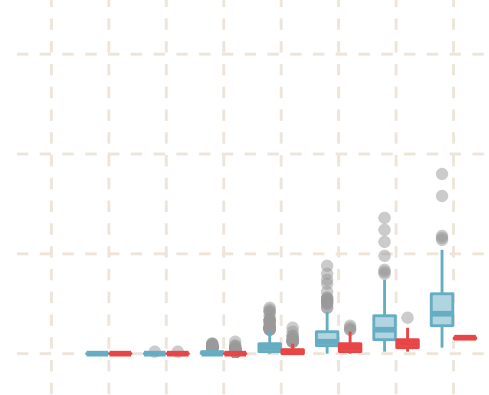
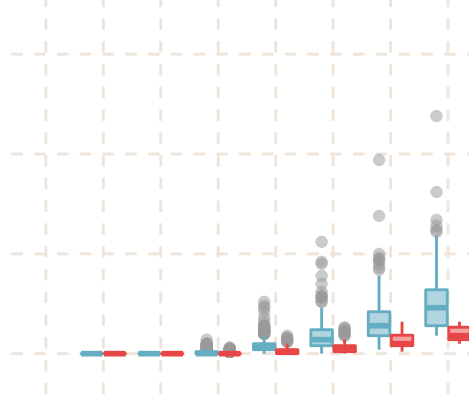
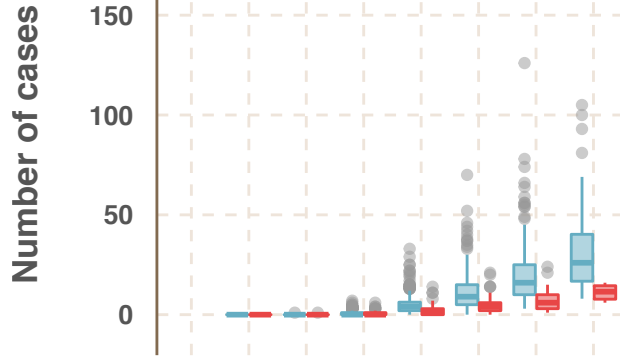
Port Loko

Western Area

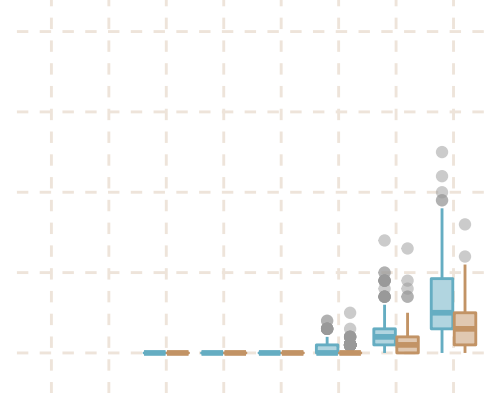
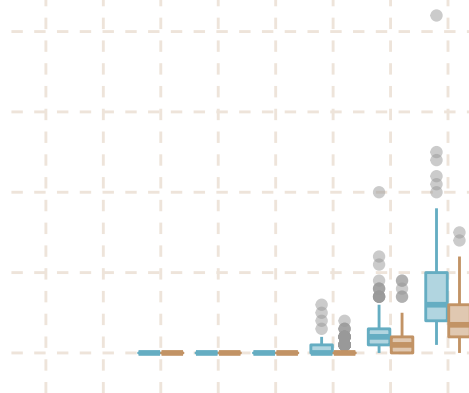
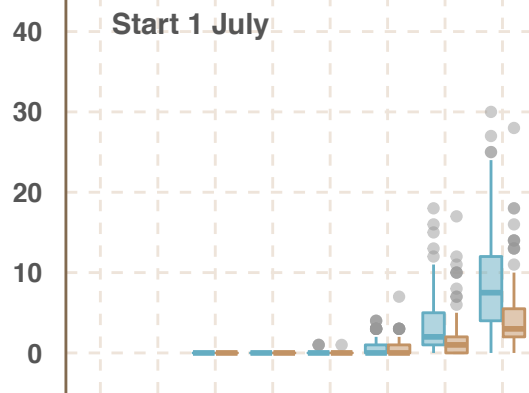
Start 1 May



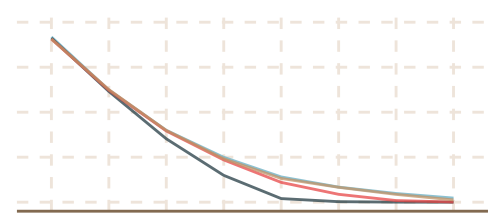
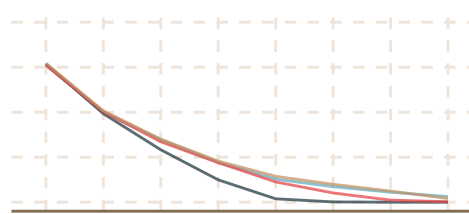
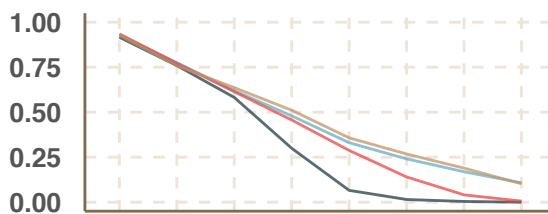
Start 1 June



Start 1 July



Persistence probability



May Jun Jul Aug Sep Oct Nov Dec

May Jun Jul Aug Sep Oct Nov Dec

May Jun Jul Aug Sep Oct Nov Dec

Month of 2015

Figure 4

Number of simulations 1 10 100 1000

Kambia

Port Loko

Western Area

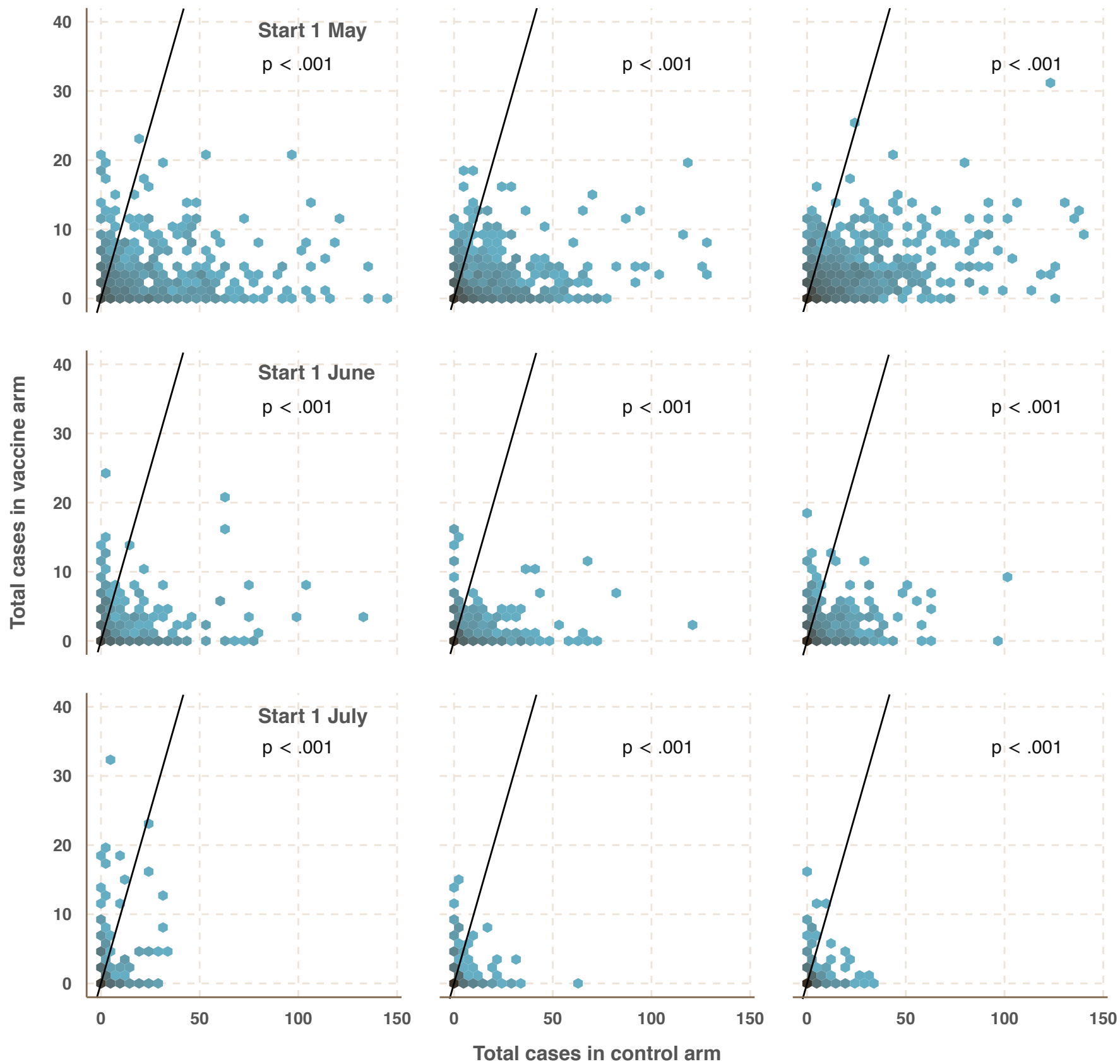


Figure 5

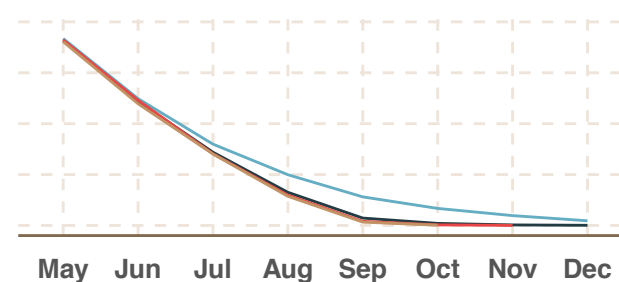
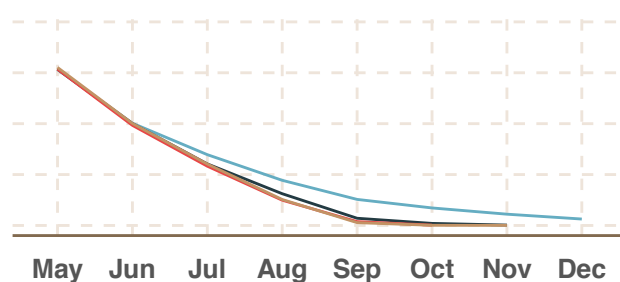
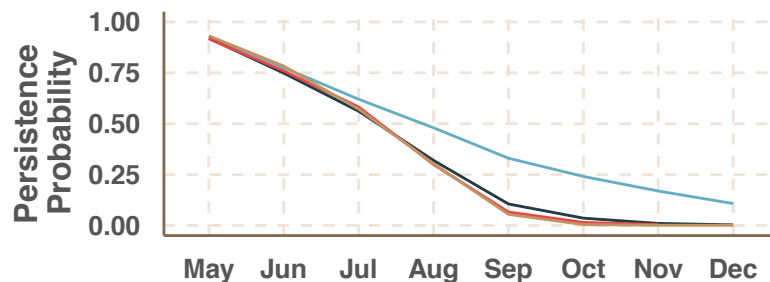
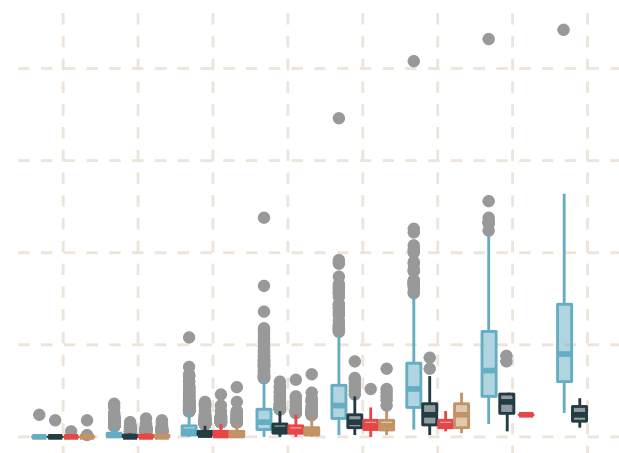
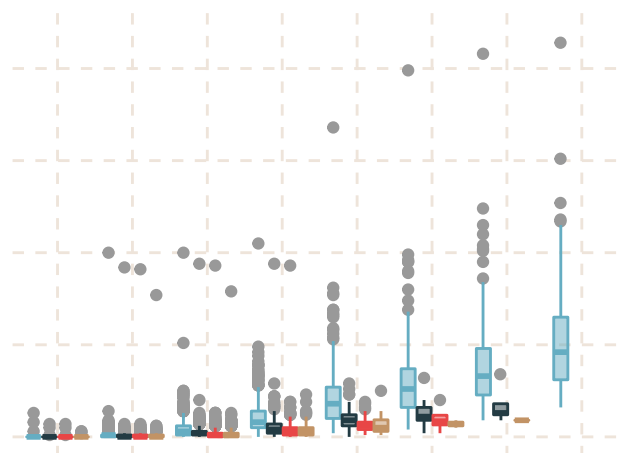
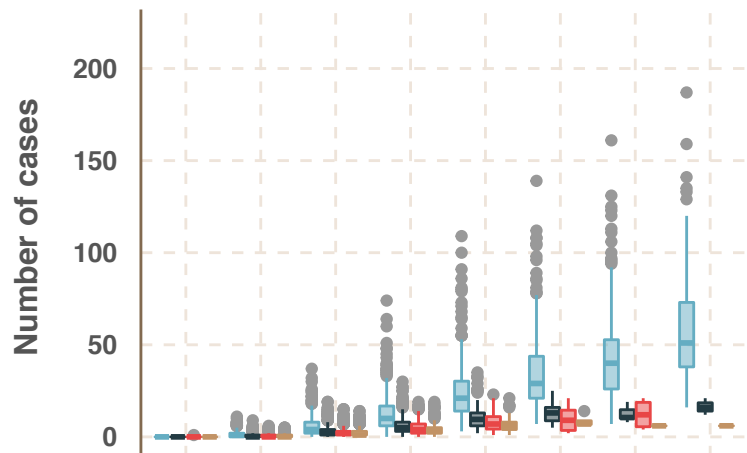
### Hypothetical vaccine efficacy

Control 60% 80% 90%

#### Kambia

#### Port Loko

#### Western Area



Month

Figure 6

