**Simulation-guided phase 3 trial design to evaluate vaccine effectiveness to prevent Ebola Virus Disease infection:  statistical considerations, design rationale and challenges**

An Vandebosch1, Robin Mogg3, Nele Goeyvaerts1, Carla Truyers1; Brian Greenwood4, Debby Watson-Jones4, Guillermo Herrera-Taracena2, Wim Parys1, Tony Vangeneugden1

1Janssen Research and Development, Beerse Belgium

2Janssen Research and Development LLC, Horsham, Pennsylvania, US

3Janssen Research and Development LLC, Spring house, Pennsylvania, US

4London School of Hygiene and Tropical Medicine, London, UK

[word count 4267]

**Running head: Challenges evaluating Ebola Vaccine Effectiveness**

**Keywords: Ebola, Modeling and Simulation, Phase 3 design, Statistical Challenges, Vaccine Efficacy**

This project has received funding from the Innovative Medicines Initiative 2 Joint Undertaking under grant agreement No 115854. This Joint Undertaking receives support from the European Union’s Horizon 2020 research and innovation programme and the European Federation of Pharmaceutical Industries and Association.

**Corresponding author:**

**An Vandebosch,**

**Turnhoutseweg 30,**

**2340 Beerse**

**Belgium**

**avandebo@its.jnj.com**

**+32-14-60 5083**

**Abstract**

Starting in December 2013, West Africa was overwhelmed with the deadliest outbreak of Ebola virus known to date, resulting in more than 27,500 cases and 11,000 deaths. In response to the epidemic, development of a heterologous prime-boost vaccine regimen was accelerated and involved preparation of a phase 3 effectiveness study. While individually randomized controlled trials are widely acknowledged as the gold-standard for demonstrating the efficacy of a candidate vaccine, there was considerable debate on the ethical appropriateness of these designs in the context of an epidemic. A suitable phase 3 trial must convincingly ensure unbiased evaluation with sufficient statistical power. In addition, efficient evaluation of a vaccine candidate is required so that an effective vaccine can be immediately disseminated. This manuscript aims to present the statistical and modeling considerations, design rationale and challenges encountered due to the emergent, epidemic setting that led to the selection of a cluster-randomized phase 3 study design under field conditions.

**Background**

The unprecedented size and severity of the 2014-15 Ebola outbreak in West Africa1 led to the declaration of a public health emergency of international concern by the WHO in August 2014. No licensed vaccine, treatment or cure exists for this disease leading to immense pressure to rapidly implement clinical research. Accelerated development of a heterologous prime-boost regimen in which Ad26.ZEBOV primes a filovirus-specific immune response followed by MVA-BN-Filo to boost the immune response involved preparation of a phase 3 effectiveness study.2 Individually randomized controlled trials are widely acknowledged as the gold-standard for demonstrating the efficacy of a candidate vaccine. However there was considerable debate on the ethics of such randomization in the context of the epidemic.3-7 Alternative designs were advocated, including a cluster-randomized trial and a stepped-wedge design.4,5 Three principles to consider for an efficacy design were recommended by Lipsitch et al3: blocked randomization with analysis matched by small centers (blocks), stepped rollout for logistical or political purposes and adaptive designs allowing flexibility in sample size. When carefully implemented, these principles may offer ethical and logistical advantages in an epidemic setting, where Ebola incidence varies considerably by time and location.

A suitable phase 3 trial must ensure unbiased evaluation with sufficient statistical power. This manuscript presents the statistical considerations, design rationale and challenges encountered that led to the selection of a cluster-randomized controlled phase 3 study design. In preparation for the phase 3 trial, computer simulations were used to guide decisions about study design in an uncertain environment. Using available, historical incidence data, aspects affecting a cluster-randomized trial were evaluated, including cluster size, statistical power, type I error rate control and trial adaptation rules under a variety of scenarios with varying Ebola incidence and heterogeneity. The study was designed at the peak of the epidemic and initially submitted for review to health authorities. However, the protocol was abandoned due to the waning epidemic and a final version was not endorsed by any regulatory authority. Nevertheless, we believe that the findings of the approach taken with design and analysis could potentially provide useful insights for future Ebola or other emerging infectious diseases outbreaks.

**Challenges of conducting a vaccine trial during an Ebola outbreak**

There are considerable ethical and logistical constraints of individually randomizing healthy subjects in an epidemic environment. However, randomization has a major advantage in that an observed difference between the randomized groups can be attributed to the intervention, providing a valid assessment of the vaccine effect. This is especially important because the epidemic is not homogeneous over space and time8,9 with few Ebola cases occurring in the tail of the epidemic and infection rates that vary substantially both within and between communities. As a result, there was a legitimate concern about the likelihood of unequal exposure to Ebola infection between randomized groups. In this setting, study design requires careful implementation of a randomized allocation scheme, to minimize the possibility of differential exposure between vaccinated and controls for any reason. Unfortunately, this is not straightforward when the number of independent randomization units is small and the knowledge of potential risk stratification factors is limited. With differential exposure, a primary analysis based on a permutation method for inference-making may be further required to ensure statistical validity.

With vaccine supplies initially limited, a design in which vaccination was given to some individuals immediately, but after a delay in the remaining individuals when additional supplies were available, would allow for a controlled evaluation of vaccine efficacy. From an ethical perspective, this delay period in an epidemic setting should be minimized for the individual at risk. However, the length of the delay must also be sufficiently long to have reasonable statistical power to evaluate the effect of an efficacious vaccine. With large-scale vaccination, the operational time needed to offer the vaccine to all subjects in the community causes a natural delay, and is further impacted by the timing of the availability of additional vaccine supplies. In the proposed trial, unvaccinated participants would be offered vaccine after a delay, but only in the event that efficacy was proven.

**Potential field vaccine efficacy designs**

The primary objective of the phase 3 study is to evaluate the effectiveness of a heterologous prime-boost regimen in preventing laboratory-confirmed Ebola. We assume that safety and immunogenicity results would be available in the targeted study region among different age groups in an accelerated manner, potentially in an initial part of the phase 3 protocol.2,10

Due to the operational and ethical considerations described earlier, alternative designs to the individually randomized controlled trial were considered and qualitatively assessed with regard to operational feasibility and statistical validity. The stepped wedge design at the cluster-level has been highlighted in the literature as the leading alternative to individual randomization in this setting.5 Under such designs, different clusters are sequentially rolled over to the vaccine at different periods in time, where the time-point of roll-over may or may not be randomized. Events collected prior to vaccination are used as a control. The stepped wedge offers the ethical advantage of conducting a study where all participants eventually receive the vaccine in the context of a vaccine roll-out campaign, which cannot immediately vaccinate all participants due to logistical constraints. This design may also have the potential to employ smaller sample sizes because the vaccine effect can be estimated utilizing both between and within-cluster variability.11,12 An unbiased evaluation of vaccine efficacy using statistical methodology that controls for time-dependent changes within clusters (e.g. via re-randomization or permutation based inference) is possible.13 As an alternative to the stepped wedge design, a more traditional cluster randomized controlled trial with clusters of small households, close contacts around an infected Ebola case or large geographical areas such as an entire village could potentially serve as an alternative design, requiring a larger sample size but alleviating some of the operational and ethical hurdles with individual randomization. The potential benefits of using a stepped wedge design described above have been challenged by Kotza *et al*14 who show that the advantages of the stepped wedge design can also be achieved using a classic cluster-randomized trial. They also note that the stepped wedge design may result in a longer duration of follow up, which was highly undesirable in this setting. Further, Bellan et al13 indicated that the statistical power to establish vaccine efficacy can be significantly diminished with the stepped wedge when using permutation-based inference relative to the individual randomization since permutation is carried out at the cluster level and the number of clusters may be small. For these reasons, the stepped wedge was not further pursued. With a vaccine availability of up to 400,000 doses and the ethical and operational hurdles against use of an individual randomization, a cluster-randomized controlled trial was selected.

The remainder of this paper evaluates the feasibility and statistical power of such an approach. Specifically, clusters would be assigned randomly in a 1:1 ratio to either immediate or no vaccination (control) whereby vaccination would be offered to the control group after effectiveness was established. Administrative division of the counties would have been used to identify the clusters. For example Sierra Leone is divided into districts, districts divided in chiefdoms and chiefdoms in sections with the latter sections being selected as the clusters. Since risk of infection is likely to be more similar for subjects from the same geographical region compared to the population as a whole, balancing in exposure across study arms was proposed via stratification of region. To that end, the chiefdoms or equivalent administrative divisions would have been selected as the stratification factor. The full design is illustrated in Figure 1.

With incidences initially based on historical data and likely to reduce over time during the course of the epidemic, the optimal timing to perform the primary analysis was difficult to pre-specify. As such, a strategy to utilize real-time monitoring of Ebola incidence in subjects randomized to clusters in the control group was proposed to adaptively determine the time-point for primary analysis. As such, this adaptive decision rule aimed to determine efficacy as soon as possible (and minimize follow-up time on the control group) while maintaining the statistical power in the setting of a vaccine with at least 65% efficacy.

*Proposed Statistical Methods*

The primary analysis would compare all Ebola cases occurring after randomization in the immediate vaccination clusters with those that would be observed in the same time span in the control group. This time span is defined as the period that commences after prime vaccination until the last cluster reaches the minimum follow-up-period, and is adaptively determined by the study design. To have a consistent starting point for the unvaccinated control group, this time-window for analysis would be initiated for all clusters once the prime vaccination had been given to all subjects randomized to the immediate vaccination clusters within that chiefdom. While ascertainment of laboratory confirmed cases of Ebola from both study arms would commence when the first subject in the first cluster receives vaccination, Ebola cases observed in a cluster prior to randomization would be discarded from the primary analysis (see Figure 1). The analysis follows an intent-to-treat principle so that Ebola events are counted based on randomized treatment assignment rather than on whether the subject actually received vaccine. This is a study under field conditions to estimate vaccine effectiveness rather than vaccine efficacy with the interpretation of the overall vaccine effect, which combines both direct and indirect effects of vaccination (e.g. herd immunity).15

Two statistical methods were considered and evaluated for the primary analysis: the conditional Poisson test and a permutation (re-randomization) test. The conditional Poisson test is commonly used in vaccine development to monitor rare safety and/or disease events.16,17 This primary analysis model assumes that at the end of follow-up, the numbers of Ebola cases within each cluster are binomially distributed with parameters ($n\_{Vi}$, $π\_{Vi}$) and ($n\_{Cj}$, $π\_{Cj}$) in cluster $i$ the immediate vaccination arm V and cluster $j $ of the control arm C. When the cluster size *n* is large and/or the probability of an Ebola case at the end of follow-up small, these can be approximated with Poisson distributions with rate parameters $λ\_{Vi}=n\_{Vi}π\_{Vi}$ and $λ\_{Cj}=n\_{Cj}π\_{Cj}$. Assuming clusters are independent, the total number of Ebola cases in the immediate vaccination and control arms respectively, $V=\sum\_{i ϵ V arm}^{}V\_{i}$ and $C=\sum\_{j ϵ C arm}^{}C\_{j}$ are also approximated by Poisson distributions with rate parameters $λ\_{V}=\sum\_{i ϵ V arm}^{}n\_{Vi}π\_{Vi}$ and $λ\_{C}=\sum\_{j ϵ C arm}^{}n\_{Cj}π\_{Cj}$. Then, conditional on the total number of Ebola cases *T* = *V* + *C* observed, the number of vaccine cases *V* is binomially distributed with parameters *T* and $p=λ\_{V}/(λ\_{V}+λ\_{C}) $. The vaccine effectiveness (VE) is then expressed according to the relative risk i.e. as 1-RR whereby the relative risk RR equals $p/(\frac{n\_{V}}{n\_{C}}-p)$. This reduces to RR=$ p/(1-p)$ assuming the sample sizes are approximately equal between vaccinated and control groups. The primary null hypothesis that VE = 0 is equivalent to the simple binomial test with *p* = 0.5 under the null versus *p* < 0.5 under the alternative. The null hypothesis is rejected at the one-sided 2.5% significance level when the probability to observe *V* or less cases out of *T* is less than 0.025; using exact binomial inference with parameters *T* and *p* = 0.5.

The permutation (or re-randomization) test was evaluated as an alternative statistical method, anticipating that the type I error rate may be increased when using the conditional Poisson test due to differential exposure between randomized groups. To implement the permutation test, the treatment assignments of the clusters were randomly permuted 5000 times. Within each permutation, the vaccine efficacy was calculated as previously described, using the expression of relative risk. The null hypothesis was rejected at the one-sided 2.5% significance level when the fraction of permutations with a vaccine efficacy at least as large as observed was less than 0.025.

Note that both methods rely on observed Ebola cases only without knowledge of the actual number of subjects within each cluster, and therefore requires an assumption that the number of subjects within each cluster is distributed equally across randomization groups. This assumption provides an advantage given that the actual number of subjects in the control arm would be unknown and based on the most recent population census data available for e.g. Sierra Leone from 2004.

 *Adaptive decision rule for primary endpoint evaluation*

For the range of potential overall study incidences anticipated in the decline of the outbreak, a minimum follow-up time between 12 and 20 weeks was determined based on simulation results as described above. The follow-up time actually utilized in the study was proposed to be defined based on a statistical decision rule at a specific point in time (e.g., 4 weeks after all subjects in the immediate vaccination arm are vaccinated), by evaluating the incidence in the control group. Observed incidences in the control group that are lower than anticipated at this interim point in time would result in a decision to increase the minimum follow-up time of the study to longer than 12 weeks, while observed incidences similar or larger than anticipated would result in a decision to maintain the minimum follow-up at 12 weeks. For each simulation, Bayesian posterior probabilities that the true incidence is greater than 2.5/100,000 person-months in the control arm were generated at various study times to identify criteria to guide decisions regarding the appropriate follow-up time. Assuming 12-weeks would be needed to vaccinate all immediate vaccination clusters, the incidence in the control clusters was planned to be evaluated at one interim time point (determined by the simulation results) by an independent data monitoring committee. This committee would recommend to either maintain follow-up after the last prime vaccination in the immediate-vaccination clusters to 12 weeks or to extend follow-up time up to 20 weeks after the last prime vaccination, depending on whether there is or is not sufficient posterior probability that the true incidence is at least 2.5/100,000 person months (Figure 2). We note that there is FDA guidance on adaptively choosing a test statistic if this choice is made on blinded data.18 While adaptations made based on unblinded data can in theory inflate the type I error rate, our simulations reveal that this is not a practical problem in our setting where the overall number of events is small.

**Details of the simulation study**

The study was designed to establish superiority of the prime/boost regimen (i.e., null hypothesis of VE≤0 versus the alternative that VE>0) with approximately 90% power using a 2.5% one-sided significance level when the true VE is at least 65%. To that end, statistical power and type I error rate were evaluated via simulations under varying assumptions including adaptation rules related to early evaluation so that subjects not randomized to vaccination could receive vaccination as soon as the vaccine demonstrated efficacy.

The control incidence rates used in the simulations were based on actual and projected incidences for the region in which projected incidences were estimated using real-time dynamic transmission modeling8. The transmission model was continuously fitted toweekly reported confirmed and probable cases, combining data from the WHO reports and the daily Situation Reports from the Sierra Leone Ministry of Health. The fitted model was used to forecast the number of cases over the subsequent four weeks (so-called projections). Weekly updates of the case data and model-based estimates and projections were published in real time.19 With a fixed sample size, a target study region and associated projected incidences, the optimal follow-up time to maintain study power was determined via simulations. Operational characteristics were evaluated for follow-up times of 12 to 20-weeks after prime vaccination of the last immediate vaccination cluster under an expedited roll-out of prime vaccination to all immediate vaccination clusters of 12-weeks (6 to 7 clusters per week).

With a vaccine availability of 400,000, 160 clusters of sample size 5000 (the average size of the ‘sections’) were simulated, of which 80 were to be assigned immediate vaccination and 80 were to be assigned to control (no vaccination). The intra-cluster correlation has a significant impact on study power and sample size when using sample size calculations for cluster randomized trials. 20-22 Given the lack of available data to calculate and derive the intra-cluster correlation, the heterogeneity in control incidence was instead modeled using a mixture of 3 control incidence rates on average, with the large majority of clusters (70%) having a relatively small incidence, a smaller percentage of clusters (20%) having a relatively moderate incidence, and an even smaller percentage of clusters (10%) having a relatively larger incidence. To add to the heterogeneity in the control incidence, we further assumed that the individual cluster incidences within each of the ‘small’, ‘moderate’ and ‘large’ cluster groupings varied uniformly up to +/-20% of the overall mean incidence rate for that cluster. The incidences evaluated through simulation are provided in Table 1 and detail the mean incidence per month across all clusters, as well as within each cluster grouping (i.e., small, moderate, high). Within each cluster, an event time for Ebola was generated for each subject using the appropriate exponential distribution defined by the incidence of that cluster whereby the incidence for an immediate vaccination cluster was multiplied by one minus the vaccine efficacy. The number of observed events within each cluster was counted as those events that occurred within the defined follow-up periods (as defined above). Various vaccine efficacy scenarios ranging from 50% (specifically requested by health authority agencies) up to 80%, thereby taking into account various scenarios of noncompliance; with an uptake of the vaccine regimen less than 100% were assessed.

The operational characteristics were furthermore assessed under 2 different randomization schemes. The first scheme (‘controlled random allocation’) generates similar between-cluster heterogeneity between randomized groups thereby mimicking matching or stratifying on underlying Ebola exposure (visualized by the bracket in Figure 1) and resulting in 70% small, 20% moderate and 10% large incidence clusters within each randomized group. Alternatively, we assessed a ‘simple random allocation’, mimicking lack of additional heterogeneity control with a potential to result in differential exposure between randomized groups, by randomly assigning the 70% small, 20% moderate and 10% large incidence clusters across the 160 clusters.

**Simulation study results**

Table 1 provides simulation results assessing the type I error under VE = 0, indicating considerably increased type I error rates when using the conditional Poisson test unless randomization is controlled with appropriate stratification or matching to increase likelihood of balance in Ebola exposure across groups. Alternatively, re-randomization based inference controls the type I error rate regardless of the randomization scheme. In the field, effectively matching clusters was deemed unlikely due to the stochastic nature of the Ebola epidemic and the lack of known prognostic factors. As such, a permutation based testing approach is highly recommended for analysis and inference related to vaccine efficacy.

Table 2 provides simulation results assessing statistical power and indicate that with a fixed follow-up duration of either 12, 16 or 20 weeks after the last prime vaccination for a (0,28) day prime-boost schedule, the study is adequately powered to assess vaccine efficacy of 65% or more when the incidence is at least 1.25/100,000 person-months throughout the study in the targeted study area. This incidence (1.25/100,000 person-months) was considered as a threshold incidence to initiate an effectiveness study based on sufficient statistical power. This rate corresponds to observing 10 cases every month in the target study area. Not unexpectedly, there is a slight reduction in power when using the permutation based testing approach relative to the conditional Poisson test but the statistical power remains at or above approximately 80% at the threshold incidence.

With an incidence of 2.5/100,000 person-months or larger, a follow-up duration of 12 weeks after the last prime vaccination is sufficient while for an incidence of 1.25/100,000 person-months, a follow-up duration of 20 weeks is considered necessary for sufficient statistical power. This observation formed the basis for deriving the decision rule for the timing of the primary analysis as explained in the statistical method section and Figure 2. The performance of this decision rule (type I error, power and the probability of stopping at either 12 weeks or 20 weeks) was assessed and results are summarized in Table 3. The results indicate that power is maintained at approximately 80% or higher with 65% vaccine efficacy when the Ebola incidence is as low as 1.25/100,000 person-months. However, when Ebola incidences increase, power increases and the chance of evaluating VE earlier increases. There is 65% probability to evaluate VE at 12 weeks when the incidence is 2.5/100,000 person-months.

Additional assessments were evaluated to prematurely terminate the effectiveness study if there was sufficient evidence that the incidence was too low such that the feasibility to establish vaccine effectiveness was unlikely (i.e. ‘operational futility’. For example, with 12 weeks needed to vaccinate (prime) all immediate vaccination clusters, it was evaluated via modeling and simulation whether operational futility could be concluded 8 weeks after the first cluster has been vaccinated (this corresponds to 2/3 of the immediate vaccination completed). The results indicated (data not presented) a non-trivial power reduction coupled with a >5% probability of stopping the study when the true average incidence in the control group is 1.25 per 100,000 person-months, where without such a futility assessment there was adequate statistical power to show vaccine efficacy. Further, declaring futility in this manner was not predictive of later study success or failure, a highly undesirable feature, and hence no operational futility assessment was considered for inclusion in the protocol.

**Discussion/Conclusions**

Confronted with ethical imperatives and logistical hurdles in the most recent Ebola outbreak, the design of an effectiveness study that strives for an unbiased assessment of the effect of a novel, heterologous vaccine candidate regimen required pragmatic considerations in combination with innovative thinking. Due to the recent epidemic circumstances, field conditions, and statistical considerations, a population-based cluster-randomized controlled trial approach with geographically matched randomization between immediate vaccination versus an unvaccinated control group was favored over other design scenarios to assess the vaccine efficacy of a novel vaccine regimen.

Rapid onset of protection is desired in the group of individuals immediately exposed to an Ebola case. However the 2014-2015 Ebola outbreak lasted more than 20 months, with sporadic cases of Ebola continuing to occur in the tail of epidemic over a long period of time in spatially distant locations, potentially (re)-introducing the infection into communities without any recent Ebola cases. This highlights the importance of the durability of the vaccine induced immune response and wide spatial coverage. A large-scale, population-based approach using a vaccine regimen inducing long-term direct protection and herd immunity, may provide the best protection to prevent further Ebola transmission for vaccinated (and potentially unvaccinated) individuals in the affected region.

The simulation results demonstrate that a cluster randomized controlled trial implemented with re-randomization based inference can provide robust inference with less efficiency than an individual randomized trial (higher sample size), but with the advantage of greater operational ease and potentially wider community acceptance. If smaller cluster sizes, or ultimately individual randomization were preferred, one would still be confronted with the uncertain dynamics of the Ebola epidemic, the potentially low incidence, and the logistical field conditions (e.g. assuming 65% VE, an incidence of 2.5 cases per 100,000 per month, a recruitment period of 3 months would require to individually randomization 350,190 individuals in 2 groups). The framework above attempts to tackle these challenges and offers a reasonably efficient way of decision-making from a statistical perspective when the objective is to vaccinate a large number of people in an epidemic setting.

The test-negative case-control design has been considered as an efficient alternative to potentially demonstrate effectiveness to obtain post-marketing approval. 23-25 The test-negative design is a type of case–control study that has been applied to evaluate influenza vaccine efficacy and was planned to be included in our protocol as a sensitivity analysis. Under the test-negative design, the vaccination status (vaccinated or not) is compared between the confirmed Ebola-test positive and confirmed Ebola-test-negative subjects whereby all subjects tested were selected from the same case population. The same case population may be defined as subjects presenting with Ebola-like symptoms at an Ebola Treatment or Holding Center. In the evaluation of influenza vaccine efficacy, the comparison has shown to be valid if the vaccine itself has no effect on non-influenza infections causing influenza like symptoms, and if exposure to influenza is similar for the vaccinated and unvaccinated subjects.23-25 Generalizing to the entire population would require the additional assumption that vaccine efficacy does not vary by health care-seeking behavior.24 It remains to be assessed whether those assumptions are valid in the Ebola context, in a region where roll-out of the vaccine in a community at risk during a future outbreak would be embedded within a community-randomization framework.

In conclusion, by modeling the trial design under various scenarios with simulation assumptions based on the most currently available data was useful to guide the study design decisions. The study duration would have been adaptively monitored according to the control incidence with the goal of maintaining study power while minimizing trial duration and the potential delay to vaccinate those in the control group. The actual delay period depends on the time needed to vaccinate the entire community, as well as the availability of additional vaccine supplies and the incidence of Ebola. With higher incidences, follow-up times can be shortened or actual sample size reduced while maintaining study power and potentially meeting the minimally required operational time.

Finally, the design would allow vaccine effectiveness evaluation from different statistical perspectives (‘randomization-based’ or primary method, case-test negative design and other sensitivity analysis) and as such can provide a more robust assessment given the untestable statistical assumptions inherent to any design in this context.

1. Acknowledgements

We would like to acknowledge Stefan Thoelen (Janssen), Benoit Callendret (Janssen), Peter Smith (London School of Hygiene and Tropical Medicine), Helen Weiss (London School of Hygiene and Tropical Medicine), Niel Hens (University of Hasselt) and Stijn Vansteelandt (Ghent University) for their input in these discussions. We would like to thank Anton Camacho, Sebastian Funk, Conall Watson and John Edmunds (CMMID, London School of Hygiene and Tropical Medicine) for sharing case data and model-based projections in real-time, and for useful discussions. Finally, we would like to thank the 2 referees and associate editor for their constructive and insightful feedback which contributed to the final version of this manuscript.

This project has received funding from the Innovative Medicines Initiative 2 Joint Undertaking under grant agreement No 115854. This Joint Undertaking receives support from the European Union’s Horizon 2020 research and innovation programme and the European Federation of Pharmaceutical Industries and Association.

1. declaration of conflict of interest

This project has received funding from the Innovative Medicines Initiative 2 Joint Undertaking under grant agreement No 115854. This Joint Undertaking receives support from the European Union’s Horizon 2020 research and innovation programme and the European Federation of Pharmaceutical Industries and Association.
Vandebosch, Mogg, Goeyvaerts, Truyers, Vangeneugden, Herrera-Taracena and Parys are employees of Janssen.

References

1. WHO Ebola Situation Report 22 July 2015, [http://apps.who.int/ebola/current-situation/ebola-situation-report-24-june-2015 (2015](http://apps.who.int/ebola/current-situation/ebola-situation-report-24-june-2015.%20%282015) Accessed 23 July 2015).
2. Johan Van Hoof. Presentation during Vaccines and Related Biological Products Advisory Committee Meeting May 12, 2015. Janssen Ebola Vaccine Program Update, [http://www.fda.gov/AdvisoryCommittees/CommitteesMeetingMaterials/BloodVaccinesandOtherBiologics/VaccinesandRelatedBiologicalProductsAdvisoryCommittee/ucm447997.htm (2015](http://www.fda.gov/AdvisoryCommittees/CommitteesMeetingMaterials/BloodVaccinesandOtherBiologics/VaccinesandRelatedBiologicalProductsAdvisoryCommittee/ucm447997.htm%20%282015), accessed 25 July, 2015)
3. Lipsitch M, Eyal N, Halloran ME et al. Vaccine Testing. Ebola and beyond. *Science* 2015*;* 348:46-48.
4. Cohen J, Kupferschmidt K. Ebola vaccine trials raise ethical issues. *Science* 2014; 346: 289–90.
5. Tully CM, Lambe T, Gilbert SC et al. Emergency Ebola response: a new approach to the rapid design and development of vaccines against emerging diseases. *Lancet Infect Dis* 2015; 15: 356-359.
6. Roca A, Afolabi MO, Saidu, Y et al. Ebola: A holistic approach is required to achieve effective management and control. *J Allergy Clin Immunol* 2015; 135: 856–867.
7. Adebamowo C, Bah-Sow O, Binka, F et al. Randomised controlled trials for Ebola: practical and ethical issues *Lancet* 2014; 384: 1423–1424.
8. Camacho A, Kucharski A, Aki-Sawyerr Y et al. Temporal Changes in Ebola Transmission in Sierra Leone and Implications for Control Requirements: a Real-time Modelling Study. *PLOS Currents Outbreaks* 2015; Feb 10. Edition 1. DOI: 10.1371/currents.outbreaks.406ae55e83ec0b5193e30856b9235ed2.
9. Merler S, Ajelli M, Fumanelli, L et al. Spatiotemporal spread of the 2014 outbreak of Ebola virus disease in Liberia and the effectiveness of non-pharmaceutical interventions: a computational modelling analysis. *Lancet Infect* *Dis* 2015; 15: 204-211.
10. Milligan ID, Gibani MM, Sewell R et al. Novel Adenovirus type 26 and MVA-vectored Ebola vaccines: Immunogenicity and Reactogenicity when used in Heterologous Prime-Boost Schedules. 2015. *Submitted*
11. Hussey MA, Hughes JP. Design and analysis of stepped wedge cluster randomized trials. *Contemp Clin Trials* 2007;28: 182–191.
12. Mdege ND, Man MS, Taylor (Nee Brown) CA et al. Systematic review of stepped wedge cluster randomized trials shows that design is particularly used to evaluate interventions during routine implementation. *J Clin Epidemiol* 2011; 64: 936–948
13. Bellan S, Pulliam JRC, Pearson CAB et al. Statistical power and validity of Ebola vaccine trials in Sierra Leone: a simulation study of trial design and analysis. *Lancet Infect Dis* 2015; 15: 703–10.
14. Kotza D, Spigta M, Arts ICW et al. Use of the stepped wedge design cannot be recommended: A critical appraisal and comparison with the classic cluster randomized controlled trial design. *J Clin Epidemiol* 2012; 65: 1249–1252
15. Halloran ME, Longini Jr IM, Struchiner CJ. Design and Analysis of Vaccine Studies. 1st ed. Springer-Verlag New York, 2010.
16. Heyse JF, Kuter BJ, Dallas MJ et al. Evaluating the safety of a rotavirus vaccine: the REST of the story. *Clin Trials* 2008; 5: 131-9.
17. Dragalin V, Fedorov V and  Cheuvart, B. Statistical approaches to establishing vaccine safety. *Statist Med* 2002; 21: 877–893.
18. FDA Draft Guidance on Adaptive Design Clinical Trials for Drugs and Biologics <http://www.fda.gov/downloads/Drugs/Guidances/ucm201790.pdf> (2010, 22 Sep 2015)
19. Centre for the Mathematical Modelling of Infectious Diseases. London School of Hygiene & Tropical Medicine. Visualisation and projections of the Ebola outbreak in West Africa, <http://ntncmch.github.io/ebola/> (Accessed 6 November 2015).
20. Campbell MJ. Cluster randomized trials in general (family) practice research. *Stat Methods Med Res* 2000; 9: 81-94.
21. Eldridge S, Ukoumunne OC, Carlin BC. The Intra-Cluster Correlation Coefficient in Cluster Randomized Trials: A Review of Definitions. *Int Stat Rev* 2009; 77: 378–394
22. Campbell MJ, Donner A and Klar. Developments in cluster randomized trials and Statistics in Medicine, *Statist Med* 2007; 26: 2-19.
23. De Serres G, Skowronski DM, Wu XW et al. 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: pii20585.
24. Jackson ML and Nelson JC. The test-negative design for estimating influenza vaccine effectiveness *Vaccine* 2013; 31:2165-2168.
25. FDA Briefing Document Vaccines and Related Biological Products Advisory Committee Meeting May 12, 2015 Licensure of Ebola Vaccines: Demonstration of Effectiveness, <http://www.fda.gov/downloads/AdvisoryCommittees/CommitteesMeetingMaterials/BloodVaccinesandOtherBiologics/VaccinesandRelatedBiologicalProductsAdvisoryCommittee/UCM445819.pdf> (2015, accessed 17 July 2015)

Table 1: Type I Error (%) for Various Control Incidences and Minimum Follow-up Times using Controlled versus Simple Random Cluster Allocation to Immediate Vaccination versus Control

|  |  |  |
| --- | --- | --- |
| **Average Control Incidence (/100,000 person-months)** | **Within Cluster Control Incidences (/100,000 person-months)** | **Type I error rate** |
| **Low Incidence Clusters** | **Moderate****Incidence Clusters** | **High Incidence Clusters** | **Controlled Random Allocation using Conditional Poisson Analysis** | **Simple Random Allocation using Conditional Poisson Analysis** | **Simple Random Allocation using Permutation Test Analysis** |
| **Follow-up Time (weeks):** | **12w** | **16w** | **20w** | **12w** | **16w** | **20w** | **12w** | **16w** | **20w** |
| 5  | 2 | 10 | 16 | 2.2 | 2.1 | 2.3 | 7.5 | 8.5 | 9.8 | 2.0 | 2.3 | 2.4 |
| 2.5 | 1 | 5 | 8 | 1.7 | 1.9 | 1.9 | 4.9 | 5.1 | 5.7 | 1.9 | 2.4 | 2.3 |
| 1.25 | 0.5 | 2.5 | 4 | 1.5 | 2.0 | 1.7 | 2.7 | 3.3 | 4.1 | 1.6 | 1.6 | 1.8 |
| 0.75 | 0.25 | 1.5 | 2.75 | 1.8 | 1.8 | 2.1 | 2.5 | 2.7 | 3.3 | 1.7 | 1.6 | 1.8 |
| Based on 5000 simulations, with 80 clusters of size 5000 per study armAssumes 12 weeks to vaccinate (prime) all immediate vaccination clusters and (0, 28) day prime/boost scheduleControlled Random Allocation accounts for heterogeneity through matching on baseline Ebola incidenceSimple Random Allocation does not account for heterogeneity through matching on baseline Ebola incidenceLow, Moderate and High incidence clusters account for 70%, 20%, and 10% of total clusters, respectively12 w, 16 w 20 w corresponds to 12, 16 and 20 weeks minimum follow-up time, respectively |

Table : Power (%) for Various Control Incidences, Minimum Follow-up Times and True Effects of Vaccine Efficacy (VE)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Average Control Incidence (/100,000 person-months)** | **Within Cluster Control Incidences (/100,000 person-months)** | **65% VE using Conditional Poisson Analysis** | **65% VE using Permutation Test Analysis** | **80% VE using Permutation Test Analysis** |
| **Low Incidence Clusters** | **Moderate****Incidence Clusters** | **High Incidence Clusters** |
| **Follow-up Time (weeks):** | **12w** | **16w** | **20w** | **12w** | **16w** | **20w** | **12w** | **16w** | **20w** |
| 5  | 2 | 10 | 16 | 99.6 | 99.9 | >99.9 | 97.7 | 98.9 | 99.3 | 99.8 | >99.9 | >99.9 |
| 2.5 | 1 | 5 | 8 | 92.6 | 96.8 | 98.4 | 86.8 | 92.9 | 95.2 | 95.1 | 99.0 | 99.9 |
| 1.25 | 0.5 | 2.5 | 4 | 68.3 | 78.5 | 85.1 | 62.3 | 71.4 | 78.5 | 75.6 | 90.1 | 96.1 |
| 0.75 | 0.25 | 1.5 | 2.75 | 45.1 | 54.9 | 62.8 | 35.8 | 46.2 | 53.4 | 49.7 | 67.1 | 80.2 |
| Based on 5000 simulations, with 80 clusters of size 5000 per study armAssumes 12 weeks to vaccinate (prime) all immediate vaccination clusters and (0, 28) day prime/boost schedule with Simple Random Allocation that does not account for heterogeneity through matching on baseline Ebola incidenceLow, Moderate and High incidence clusters account for 70%, 20%, and 10% of total clusters, respectively12 w, 16 w 20 w corresponds to 12, 16 and 20 weeks minimum follow-up time, respectively |

Table Type I error and power with implemented adaptation rule

|  |  |  |  |
| --- | --- | --- | --- |
| **Average Control Incidence**  | **Probability define minimum follow-up time equal to** | **Type I** | **Power** |
| **(/100,000 person-months)** | **12 w** | **20 w** | **error** | **65% VE** | **80% VE** |
| 5 | > 99.9 | < 0.1 | 2.0 | 97.7 | 99.8 |
| 2.5 | 65.6 | 34.4 | 2.0 | 92.6 | 99.0 |
| 1.25 | 0.8 | 99.2 | 1.9 | 78.5 | 96.1 |
| 0.75 | < 0.1 | > 99.9 | 1.8 | 53.4 | 80.2 |
| Based on 5000 simulations, with 80 clusters of size 5000 per study arm using Permutation Test AnalysisAssumes 12 weeks to vaccinate (prime) all immediate vaccination clusters and (0, 28) day prime/boost schedule with Simple Random Allocation that does not account for heterogeneity through matching on baseline Ebola incidenceLow, Moderate and High incidence clusters account for 70%, 20%, and 10% of total clusters, respectively12 w and 20 w corresponds to 12 and 20 weeks minimum follow-up time, respectivelyVE=vaccine efficacy |

|  |
| --- |
| Figure : Graphical visualization of the key design elements for the evaluated cluster-randomized design |
|  |

Figure 2: Flexible follow-up monitoring rule on control incidence with an anticipated time of 12 weeks to vaccinate (prime) all immediate clusters and (0, 28) day prime/boost schedule

At 16

weeks in time

:

Observe



2

2

cases

in DV arm?

Commit to analysis at

24

weeks in ti

me (min

(12

weeks follow

-

up)

Commit to analysis at

32

weeks in time

(2

0

weeks follow

-

up)

**YES**

**NO**