Journal of Clinical Epidemiology 147 (2022) 11−20

Journal of Clinical Epidemiology

ORIGINAL ARTICLE

Year-to-year variation in attack rates could result in underpowered respiratory syncytial virus vaccine efficacy trials

Marie-Noëlle Billard[a,*], Joanne Wildenbeest[a], Louis J. Bont[a], Harish Nair[b], John P. McCracken[c,d], Katrien Oude Rengerink[e], for the RESCEU Investigators

[a]*Department of Pediatric Infectious Diseases and Immunology, Wilhelmina Children's Hospital, University Medical Center Utrecht, The Netherlands*
[b]*Centre for Global Health, Usher Institute, University of Edinburgh, Edinburgh, Scotland, UK*
[c]*Global Health Institute and Department of Epidemiology and Biostatistics, College of Public Health, University of Georgia, Athens, GA, USA*
[d]*Center for Health Studies, Universidad del Valle de Guatemala, Guatemala*
[e]*Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht, The Netherlands*

Accepted 16 February 2022; Published online 23 February 2022

Abstract

Objectives: Year-to-year variation in respiratory viruses may result in lower attack rates than expected. We aimed to illustrate the impact of year-to-year variation in attack rates on the likelihood of demonstrating vaccine efficacy (VE).

Study Design and Setting: We considered an individually randomized maternal vaccine trial against respiratory syncytial virus (RSV)-associated hospitalizations. For 10 RSV-associated hospitalizations per 1,000 infants, sample size to have 80% power for true VE of 50% and 70% was 9,846 and 4,424 participants. We reported power to show VE for varying attack rates, selected to reflect realistic year-to-year variation using observational studies. Eight scenarios including varying number of countries and seasons were developed to assess the influence of these trial parameters.

Results: Including up to three seasons decreased the width of the interquartile range for power. Including more seasons concentrated statistical power closer to 80%. Least powered trials had higher statistical power with more seasons. In all scenarios, at least half of the trials had $<80\%$ power. For three-season trials, increasing the sample size by 10% reduced the percentage of underpowered trials to less than one-quarter of trials.

Conclusion: Year-to-year variation in RSV attack rates should be accounted for during trial design. Mitigation strategies include recruiting over more seasons, or adaptive trial designs.

Keywords: RSV; Seasonality; Statistical power; Attack rate; Incidence; Sample size

1. Introduction

Phase III vaccine trials are of critical importance to provide pivotal evidence to decide whether the benefits of the vaccine outweigh its risks before they are made commercially available [1]. Planning and executing these large multicenter trials requires considerable time and resources [1,2]. Trials failing to show efficacy despite a clinically relevant effect may result in disregarding potentially effective vaccine candidates for lack of evidence.

Why is planning phase III vaccine trials such a challenge? To demonstrate the efficacy of an intervention, enough cases must be observed during the trial to show a statistically significant difference between the intervention and nonintervention groups [1]. Vaccine trials include healthy subjects that may later be exposed to the pathogen without knowing how many will be infected. This makes vaccine trials vulnerable to overestimating disease incidence [3].

Vaccine trials may face additional challenges due to year-to-year variation. Attack rates are often regarded as a stable parameter throughout the study for a given

Conflict of interest statement: Declaration of Competing Interests: All authors declare no support from any organization for the submitted work and no financial relationships with any organizations that might have an interest in the submitted work in the previous 3 years.

Author Contributions: B.MN, OR.K and B.LJ conceptualized the study. B.MN, OR.K, M.JP, *N*.H collected the data. B.MN and OR.K drafted the manuscript and conducted the analyses. All authors contributed to the interpretation of the results and revised the manuscript.

* Corresponding author. Pediatric Infectious Diseases, RSV Research Group, Universitair Medisch Centrum Utrecht, P.O. Box 85090, 3508 AB, Utrecht, The Netherlands. Tel.: +31 8 87 55 40 03; fax: +31 8 87 55 53 50.

E-mail address: m.n.billard@umcutrecht.nl (M.-N. Billard).

https://doi.org/10.1016/j.jclinepi.2022.02.003

What is new?

Key findings

- In the maternal RSV VE trial scenarios, including more seasons limited the deviations from the targeted statistical power due to year-to-year variation in attack rates. In all scenarios, at least half of the trials had <80% power. Increasing sample size by 10% reduced the number of trials having <80% power from half to a quarter in scenarios that included three seasons.

What this study adds to what is known?

- The study illustrated and quantified how year-to-year variation can decrease the ability of trials to demonstrate VE. Possible avenues for limiting and monitoring the potential impact like including more seasons, increasing the target sample size and adaptive trial designs were described.

What are the implications, what should change now?

- Year-to-year variation in attack rates should be taken into account at trial design. The parameters of the scenarios were selected to isolate the effect of year-to-year variation in each site. Future studies should not only take into account potential impact of year-to-year variations in attack rates, but also other factors to help guide sites selection.

geographic area [4,5]. For seasonal pathogens and particularly respiratory pathogens, the expectation of stable incidence over several seasons is unlikely.

Unexpectedly low attack rates have been suggested as explanation for an unsuccessful vaccine efficacy (VE) trial against respiratory syncytial virus (RSV) in older adults in 2016 [6]. In addition, early models showed that the unusual RSV circulation that has been reported since the start of the coronavirus disease 2019 (COVID-19) pandemic may impact upcoming seasons [7−11]. Yet, we are not aware of research assessing the impact of year-to-year variation in attack rates on statistical power. Articles discussing attack rates and sample size present either qualitative discussions on factors to consider when planning trials [1,12,13] or methods for estimating sample size in particular situations like for complex endpoints [14−16].

RSV causes winter epidemics in temperate and tropical regions. Particularly mild RSV seasons tend to follow severe seasons, suggesting that the build-up of susceptible individuals influences intensity [17]. However, the factors involved in how often this occurs are not yet well understood. As RSV is a major cause of acute lower respiratory tract infection in young infants worldwide [18], maternal immunization strategies have been strongly supported [19,20]. No vaccine against RSV is currently licensed, but several products are at advanced stages of clinical development [21].

We selected a maternal vaccine trial against RSV to illustrate the challenge of year-to-year variation in attack rates for efficacy trials. The first objective was to estimate the impact of year-to-year variation in attack rates on the likelihood of having a successful phase III maternal RSV vaccine trial, using realistic variations in attack rates for consecutive seasons and vaccine parameters. The second was to quantify the implications for sample size.

2. Materials and methods

2.1. Illustrative trial

For illustrative purposes, we selected a maternal RSV VE trial. The endpoint was to show VE against laboratory-confirmed RSV-associated hospitalizations (RSVH) for acute lower respiratory tract infections in healthy 0- to 5-month-old infants. We considered a controlled individually randomized trial with a 1:1 ratio between vaccinated and unvaccinated pregnant women. We calculated the total sample size required to have 80% power for an arbitrarily chosen attack rate of 10 RSVH per 1,000 infants and assumed this would be the attack rate for the duration of the trial. Type I error was set to 5% and VE was 50% (VE50%) or 70% (VE70%), the lower bounds of acceptable and preferred efficacy for RSV maternal vaccines as per World Health Organisation recommendations [22]. Efficacy was measured over one season as the protection from maternal passive antibodies is expected to persist 3−6 months after birth [23,24]. The total sample size was 9,846 participants for true VE50% and 4,424 participants for true VE70%.

2.2. Scenarios

To assess the influence of (a) the duration of the trial (number of season), (b) the choice of sites included, and (c) the number of sites included, we developed eight scenarios (Table 1). The scenarios corresponded to different illustrative trial designs that included one to three seasons and two to six sites, with varying selection of sites.

2.3. Reported outcome

Statistical power is defined as the probability of obtaining a statistically significant effect if the vaccine is efficacious. The likelihood of having a successful trial was evaluated as the statistical power to show VE. We calculated statistical power of the illustrative trial when RSVH attack rates deviated from 10/1,000 infants (while keeping assumed true VE and sample size constant). For each

Table 1. Scenarios used to assess the influence of duration of the trial, and the number and selection of countries; and corresponding number of combinations of annual attack rates used in each scenario

Scenario	Number of RSV seasons	Number of countries	Selected countries	Number of attack rates
Two region scenarios in Kenya and Bolivia, varying number of seasons				
Scenario 1 (S1)	1	2	Kenya and Bolivia	36
Scenario 2 (S2)	2	2	Kenya and Bolivia	25
Scenario 3 (S3)	3	2	Kenya and Bolivia	16
Four region scenarios in Kenya, Bolivia, United States, and Spain, varying number of seasons				
Scenario 4 (S4)	2	4	Kenya, Bolivia, United States, and Spain	300
Scenario 5 (S5)	3	4	Kenya, Bolivia, United States, and Spain	96
Four region scenarios in Kenya, Bolivia, Guatemala, and Germany, varying number of seasons				
Scenario 6 (S6)	2	4	Kenya, Bolivia, Guatemala, and Germany	500
Scenario 7 (S7)	3	4	Kenya, Bolivia, Guatemala, and Germany	192
Six region scenarios in Kenya, Bolivia, United States, Spain, Guatemala, and Germany with 2 seasons				
Scenario 8 (S8)	2	6	Kenya, Bolivia, United States, Spain, Guatemala, and Germany	3,600

Abbreviations: RSV, respiratory syncytial virus.

scenario, a set of attack rates was calculated to represent realistic relative year-to-year variation. For that, we used observational studies conducted over several seasons. Power to show VE was calculated for each attack rates in the set. We reported the interquartile range (IQR) for power and the proportion of underpowered and severely underpowered trials, respectively defined as trials having $<$ 80% and $<$70% power per scenario.

2.4. Selected observational studies

We searched PubMed using the terms "RSV" or "Respiratory Syncytial Virus" and "rates" or "incidence" or "burden." Eligible studies had to report rates of RSVH in infants aged 0−5 months (preferred) or 0−11 months (more often reported), use laboratory-confirmed RSV cases, and include at least three RSV seasons (pre-COVID-19). We selected six studies in different countries to represent the six sites of the scenarios (Table 1). The selection was based on the following prioritization criteria: longest observation period for the country, annual attack rates per age group readily available in publication, and age group 0−5 months reported. Characteristics of the selected studies are available in Table 2 [18,25−29]. For simplicity, the rest of the manuscript refers to the studies by the name of the country, although they may not be representative of the entire country's seasonality.

Kenya and Bolivia were selected for two-country scenarios (S1−S3) because they had the longest observation periods. Four-country scenarios (S4−S7) included Kenya, Bolivia, and either Germany and Guatemala or the United States and Spain, respectively representing countries with substantial or limited year-to-year variation in RSV attack rates. The degree of year-to-year variation was defined as the fold-change between the highest and the lowest RSVH attack rates, with substantial variations corresponding to over two-fold changes. All six countries were included in the final scenario (S8).

2.5. Scenarios attack rates

Average RSVH attack rates in the selected studies ranged from 2.31/1,000 infants to 36.8/1,000 (Supplementary Fig. 1). This was largely due to methodological differences. The lowest attack rates were found in studies reporting RSVH among infants aged 0−11 months (vs. 0−5 months), using more severe clinical case definitions [severe acute respiratory infection (ARI)/clinical pneumonia vs ARI] and less sensitive testing methods (direct fluorescent antibody and enzyme-linked immunosorbent assay vs. reverse-transcriptase polymerase chain reaction).

These differences were expected as more severe clinical case definitions and less sensitive testing methods would capture a smaller part of the true RSV incidence. Due to these differences, it was uncertain if RSV incidence was truly higher in the selected studies that measured generally high incidence. To avoid these studies weighing more on power estimations, the attack rates of the selected studies

Table 2. Characteristics of selected studies

Country	Authors	Country income level[a]	Study period	Age group	Source of cases	Case definition[b]
Kenya	Nokes et al. [27]	LMIC	2002–2007	0–11 months	Hospitals	Clinical pneumonia
Bolivia	Chavez et al. [25]	LMIC	2012–2017	0–11 months	Hospitals	SARI
United States	Hall et al. [26]	HIC	2000/2001 to 2004/2005	0–5 months	Hospitals	ARI
Spain	Vicente et al. [28]	HIC	1996/1997 to 1999/2000	0–5 months	Hospitals	ARI
Guatemala	McCracken et al. [18]	UMIC	2008–2013	0–5 months	Hospitals and outpatient clinics	ARI
Germany	Weigl et al. [29]	HIC	1996/1997 to 2000/2001	0–11 months	Hospitals	ARI

RSV testing[c]	Total number of cases	Denominator	Range of raw incidence (per 1,000)	Range of rescaled incidence (per 1,000)	Degree of year-to-year variation[d]
DFA	424	Population estimates	9.4–20.7	6.0–13.2	1.19
RT-PCR	-	Population estimates	1.2–3.2	5.4–14.4	1.70
RT-PCR	328	Population estimates	11.7–21.7	7.2–13.3	0.85
ELISA	235	Population estimates	16.9–50.1	4.6–13.6	1.96
RT-PCR	709	Population estimates	22.2–154.5	2.5–17.1	5.97
RT-PCR	230	Population estimates	8.5–27.2	5.2–16.7	2.21

[a] LMIC, low-middle income county; UMIC, upper middle in country; HIC, high income country; according to The World Bank 2022.
[b] ARI, acute respiratory infection; SARI, severe acute respiratory infection.
[c] DFA, direct fluorescent antibody test; ELISA, enzyme-linked immunosorbent assay; RT-PCR, reverse-transcriptase polymerase chain reaction.
[d] The degree of year-to-year variation in attack rates was defined as the fold change between the highest and the lowest annual attack rates reported in the study.

were rescaled to obtain common average RSVH attack rate of 10/1,000 infants (Supplementary Fig. 1). Annual attack rates were multiplied by $\frac{10/1{,}000}{study\ average\ attack\ rate}$ which preserved the relative variations (ratios) between annual attack rates and homogenized the order of magnitude of attack rates between studies.

The set of attack rates of each scenario corresponded to all possible combinations of annual attack rates for the seasons and countries included (Table 1). For example, S1 included one season in Kenya and Bolivia. RSVH attack rates were available for six seasons in these countries, corresponding to $6 \times 6 = 36$ different attack rates. The scenarios included 16–3,600 different overall attack rates, depending on the number of countries and seasons.

As there may be patterns in year-to-year variation in RSV incidence (eg, alternation in mild and severe seasons), multiple season scenarios included only consecutive seasons. However, because the observation periods of the selected studies started at different calendar years, the first season included could differ in calendar time between countries of the same scenario (Table 2). For example, seasons 2002 and 2003 in Kenya could be coupled with seasons 2012 and 2013 or 2013 and 2014 in Bolivia but not with seasons 2012 and 2014 (as those are not consecutive). We calculated the overall cumulative RSVH attack rates for each combination of annual attack rates. For simplicity, we assumed the number of participants to be equal between countries and seasons. The overall attack rates were used to calculate power to show VE. VE was assumed to be independent of attack rates and stable over time.

2.6. Evaluation of strategies to improve the likelihood of a successful trial

One strategy to improve the likelihood of a successful trial is to aim for a higher sample size than the minimal sample size requirement for average incidence. This could prevent the trial from being underpowered if the attack rates are lower than expected. To assess the effect of increasing sample size on the distribution of power, we calculated the power to show VE in the eight scenarios when sample size was increased by 10% and 20%.

2.7. Statistical analysis

Statistical power and sample size based on Fisher's exact test were done with the exact 2×2 package with R [30]. The scenario's sample size was calculated for average attack rate (10 RSVH per 1,000 infants) with the ss2 $\times$ 2 function. This function repeats statistical power calculations with varying sample size to determine the smallest sample size returning the desired power (80%). Statistical power was calculated for each cumulative attack rates in the scenarios for the previously determined sample size by repeating Fisher's tests and summing probability of rejection.

3. Results

3.1. Impact of year-to-year variation in attack rates

3.1.1. Relationship between statistical power and attack rates.

For a given sample size, the statistical power to show VE depends on the number of events observed during the trial. Sample size was calculated to have 80% power for 10 RSVH per 1,000 participants. Thus, power was >80% when attack rates were higher than 10/1,000 participants and <80% when incidence was lower than 10/1,000 participants (Fig. 1).

3.1.2. Number of seasons.

Extending trial duration while keeping sample size and VE constant concentrated statistical power around 80% (Fig. 2, Table 3). In two-country scenarios (S1−S3), the width of the IQR for power gradually decreased when up to three seasons were included for VE50% and VE70%. This decrease was also observed when a third season was added in four-country scenarios whether additional countries showed limited (S4−S5) or substantial (greater than two-fold change) year-to-year variation in attack rates (S6 vs S7) and for both VE values. In two-country scenarios, including more seasons increased the proportion of trials with <80% power from 50% in S1 to 56% in S3. Although the proportion of trials with <80% power was higher in S5 than S4 (85% and 69%), it was 67% in S6 and S7. On the contrary, the proportion of trials with <70% power decreased when more seasons were included. No trials had <70% power when three seasons were included.

3.1.3. Selection of countries.

To assess the influence of country selection, we compared scenarios including Kenya, Bolivia, and two countries showing either substantial (Guatemala and Germany, greater than two-fold change, S6 and S7) or limited (United States and Spain, S4 and S5) year-to-year variation in RSVH attack rates (Fig. 2, Table 3). The width of the IQR for power was larger with Guatemala and Germany than with Spain and the United States. This was observed in two-season (S6, S4) and three-season scenarios (S7, S5), and for VE50% and VE70%. The proportion of trials with <80% power was lower with Guatemala and Germany than with Spain and the United States, for two-season (67% in S6, 69%in S4) and three-season trials (73% in S7, 85% in S5). When two seasons were included, the proportion of trials with <70% power was 3% with Spain and the United States (S4) and 4% with Guatemala and Germany (S6).

3.1.4. Number of countries.

In scenarios with two seasons, increasing the number of countries from two to six (S2, S4, S6, S8) had limited effect on the width of the IQR for statistical power (Fig. 2, Table 3). The proportion of trials with <80% power increased from 56% with two countries in S2 to 73% with six countries in S8. On the contrary, the proportion of trials with less than 70% power decreased from 8% to 1% when the number of countries increased.

3.2. Accounting for variation in sample size

We assessed the extent to which year-to-year variation can be accounted for by increasing the original sample size by 10% or 20%. This corresponded to 986−1,970 additional participants for VE50% and 444−886 additional participants for VE70% (Table 3). In all three-season scenarios (S3, S5, and S7), less than one-quarter of trials had <80% power when sample size was increased by 10% (Q1 > 80%). Conversely for S1 (one-season two-country scenario), this was achieved only by increasing sample size by 20%.

4. Discussion

This study aims to quantify the impact of year-to-year variation in RSVH attack rates on power to show VE using realistic parameters. We developed eight RSV maternal vaccine trial scenarios with different numbers of countries and seasons included to illustrate the impact of the selection on power. Power was mostly influenced by the number of seasons. Although including more seasons decreased the deviations from the targeted power by decreasing the width of the IQR for power (S1−S7), including more countries had no clear statistical benefit (S2, S4, S6, and S8). Including countries with substantial variations in RSVH attack rates had a small but negative effect, as it increased the width of the IQR for power (S4−S7). In all scenarios, at least half of the trials had <80% power and the proportion of trials with <80% power increased slightly with more seasons and more countries. Including more seasons centered power closer around 80% which decreased the proportion of severely underpowered trial (<70%). To reduce the proportion of underpowered trials <25%, sample size had to be increased by 10% for the three-season scenarios (S4, S5, and S7) and 20% for the one-season scenario (S1).

The scenarios included all possible combinations of RSVH attack rates for the chosen number of seasons and countries. Consequently, including more countries did not eliminate extreme scenarios with simultaneous mild or severe seasons in all countries. Thus, only extending the number of seasons improved the likelihood of achieving the targeted power. However, as shown in the literature, including more sites can spread the risk of low attack rates, low recruitment rates, and improve generalizability to facilitate licensing [31,32].

One apparent disadvantage of including more seasons and countries was the larger proportion of trials with < 80% power. This was due to the selected studies and the parameters of the scenarios. In most of the selected studies,

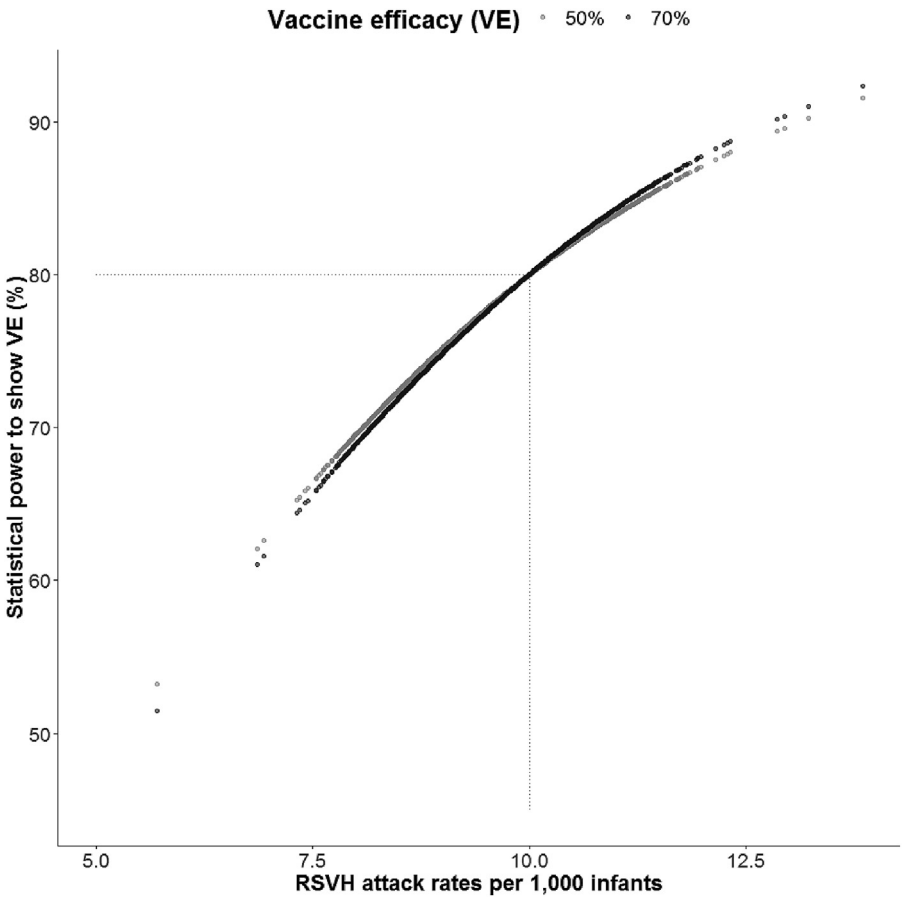

Fig. 1. Statistical power to show a statistically significant VE based on the expected number of events for all sets of RSVH attack rates included in scenarios S1−S8 combined, for VE = 50% (9,846 participants) and VE = 70% (4,424 participants). RSVH, respiratory syncytial virus-associated hospitalization; VE, vaccine efficacy.

attack rates were higher than average in the first and last seasons while the mildest seasons were recorded in the middle of the observation period (Supplementary Fig. 1). Because multiseason scenarios included only consecutive seasons, middle seasons with lower incidence were included in twice as many attack rate calculations than extremity seasons. Including more seasons or more countries amplified the effect and increased the proportion of trials having $<80\%$ power. This increase was not observed when the last and first season were assumed to be consecutive (data not shown). This underlines that average attack rates are sensitive to the observation period and should not be considered as stable values.

As expected, the width of the IQR for power increased with countries showing substantial year-to-year variation in attack rates as more extreme attack rates were included. The effect was small and of similar magnitude with two (S4 and S5) or three seasons (S6 and S7). In field trials, the effect of including countries with important year-to-year variation could be counterbalanced by generally high RSVH attack rates. Particular patterns such as biannual seasonality can impact ability of trials to show VE. In regions with alternating mild and severe RSV seasons, like Finland, a three-season trial could include two mild or two severe seasons, depending on the first season of the trial [17,33]. This would substantially alter the number of expected events despite equal recruitment.

Strategies to limit the impact of year-to-year variation in attack rates on trial outcomes can be integrated in the trial design. First, the most influential factor in the scenarios was the number of seasons. For the same sample size, adding a second season eliminated extreme possibilities and a third season further centered power (S1−S3). This suggests that phase III trials should not be conducted over a single season without strong indications of a severe season coming or when the duration of the trial is of particular importance (eg, public health emergencies). Including more than two seasons should be considered because two mild RSV seasons can occur consecutively (Supplementary Fig. 1).

Second, most trials in the scenarios had $<80\%$ power, suggesting that a safety margin could be added to the minimal required sample size (or sample size should be calculated for a higher power than intended). Year-to-year variation can be of such amplitude that accounting for the lowest possible attack rate would likely result in large sample size increases. However, when combined with including

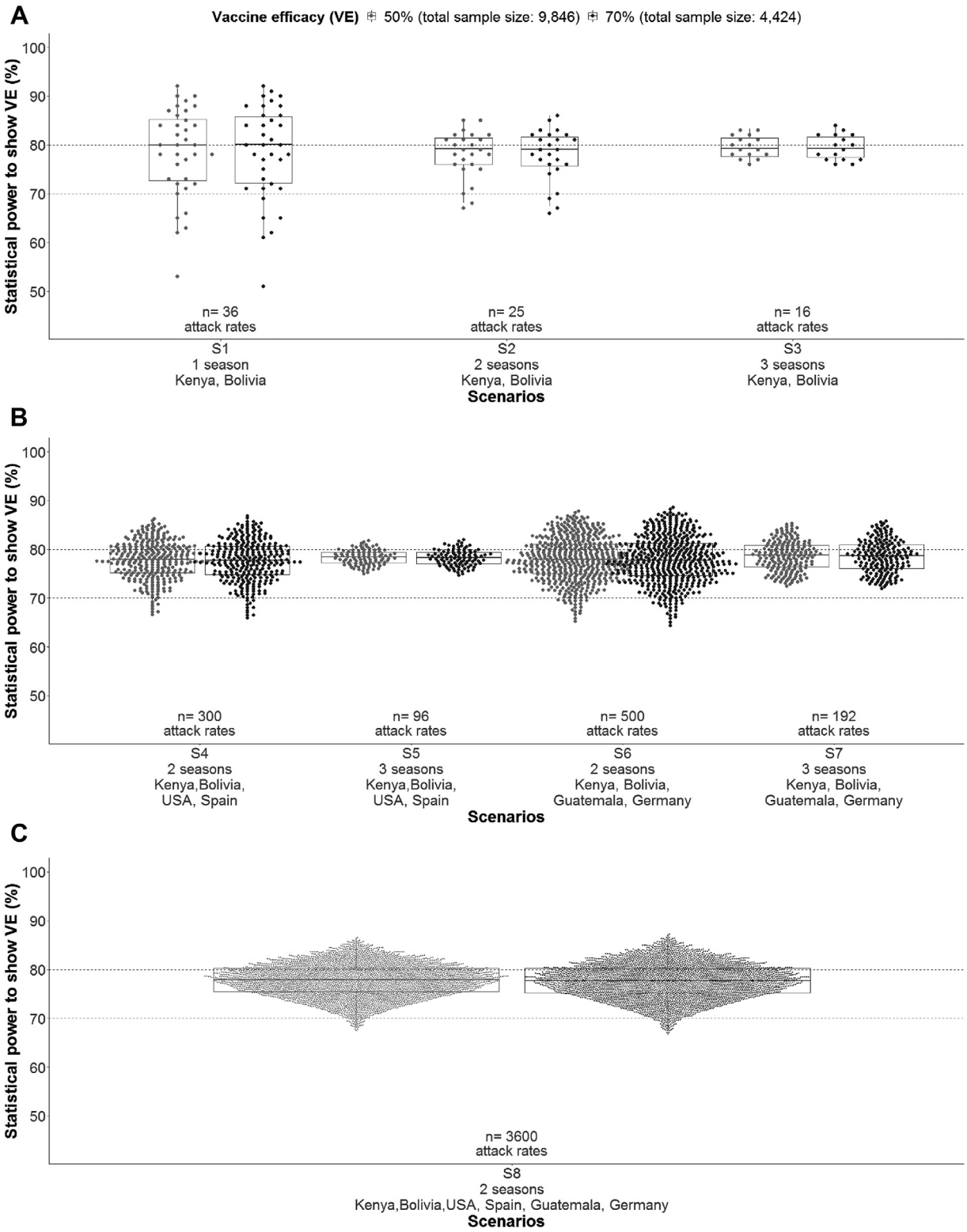

Fig. 2. Statistical power to show VE based on the expected number of events for the RSVH attack rates included in the set of scenarios S1–S8 and number of RSVH attack rates in each scenario, for VE = 50% with 9,846 participants and VE = 70% with 4,424 participants. (A) Scenarios S1–S3, including one to three seasons in Kenya and Bolivia. (B) Scenarios S4–S7, including two and three seasons in Kenya, Bolivia, and either the United States and Spain or Guatemala and Germany. (C) Scenario S8 including two seasons in Kenya, Bolivia, the United States, Spain, Guatemala, and Germany. RSVH, respiratory syncytial virus-associated hospitalization; VE, vaccine efficacy.

Table 3. Median and interquartile range for statistical power in scenarios S1–S8 for the base scenario and sample sizes increased by 10% and 20%; and number of combinations of annual attack rates used in each scenario

Scenarios	Number of attack rates	Vaccine efficacy 50%			Vaccine efficacy 70%		
		Base scenario	10% increase	20% increase	Base scenario	10% increase	20%increase
Sample size		9,846	10,832	11,816	4,424	4,868	5,310
Two countries							
Scenario S1[a]	36	80.0 [72.6–85.2]	83.9 [77.1–88.5]	87.1 [81.0–91.2][b]	80.0 [72.2–85.8]	84.3 [76.9–89.3]	87.8 [81.1–92.0][b]
Scenario S2[a]	25	79.2 [76.0–81.5]	83.2 [80.3–85.2][b]	86.4 [83.8–88.3]	79.1 [75.7–81.6]	83.5 [80.3–85.8][b]	87.1 [84.2–89.0]
Scenario S3[a]	16	79.3 [77.6–81.5]	83.3 [81.8–85.2][b]	86.6 [85.2–88.3]	79.3 [77.4–81.6]	83.7 [82.0–85.8][b]	87.2 [85.7–89.0]
Four countries							
Scenario S4[a]	300	77.9 [75.2–80.5]	82.0 [79.5–84.4]	85.4 [83.2–87.5][b]	77.7 [74.8–80.6]	82.3 [79.5–84.9]	86.0 [83.5–88.2][b]
Scenario S5[a]	500	78.4 [77.2–79.4]	82.5 [81.4–83.4][b]	85.8 [84.8–86.6]	78.3 [77.0–79.4]	82.7 [81.6–83.8][b]	86.4 [85.3–87.3]
Scenario S6[a]	96	77.9 [75.0–81.1]	82.0 [79.3–84.9]	85.4 [83.0–87.9][b]	77.8 [74.6–81.2]	82.3 [79.3–85.4]	86.0 [83.3–88.7][b]
Scenario S7[a]	192	78.8 [76.4–80.8]	82.8 [80.6–84.7][b]	86.1 [84.2–87.8]	78.7 [76.1–80.9]	83.1 [80.7–85.1][b]	86.7 [84.6-88.5]
Six countries							
Scenario S8[a]	3,600	77.9 [75.5–80.2]	82.0 [79.8–84.1]	85.4 [83.4–87.3][b]	77.8 [75.2–80.3]	82.3 [79.8–84.5]	86.0 [83.8–88.0][b]

[a] S1: one season in Kenya and Bolivia; S2: two seasons in Kenya and Bolivia; S3: three seasons in Kenya and Bolivia; S4: two seasons in Kenya, Bolivia, United States, and Spain; S5: three seasons in Kenya, Bolivia, United States, and Spain; S6: two seasons in Kenya, Bolivia, Guatemala, and Germany; S7: three seasons in Kenya, Bolivia, Guatemala, and Germany; S8: two seasons in Kenya, Bolivia, United States, Spain, Guatemala, and Germany.

[b] Less than one-quarter of trials underpowered (Q1 $\geq$ 80% power).

more seasons, increasing sample size by 10% was sufficient to reduce the proportion of underpowered trials to $<25\%$. Intensifying recruitment when and where attack rates are high would maximize the benefit of including more participants but requires predicting RSV circulation weeks in advance.

Third, the statistical analysis plan should include dispositions to ensure enough cases are observed during the trial. For binary outcomes, demonstrating VE requires a fixed number of events depending on VE [34,35]. Comparing the number of observed events with the required number (''case-counting'') allows one to assess the likelihood of demonstrating VE without unblinding the trial. Alternatively, adaptive trial designs can be used as ''insurance'' against overestimating attack rates [36,37]. Adaptive protocols include decision-making rules to modify sample size or trial duration according to predefined interim data analyses. Event-driven designs are less dependent on accurately estimating incidence but are impractical for maternal vaccine trials due to the long delay between vaccination and occurrence of cases. Although event-driven designs will be less likely to be underpowered, an estimate of attack rates and duration of the trial is still needed to inform planning of the trial.

The strength of this study is the use of realistic relative year-to-year variation in RSVH attack rates. The six observational studies in different countries we selected reported laboratory-confirmed RSV cases and used testing methods with $>95\%$ specificity [38–40]. The main limitations result from developing the scenarios to ensure that power variation was solely due to year-to-year variation. By rescaling attack rates only the relative variation (ratios) between annual attack rates was kept which resolved systematic underestimation of true RSV incidence due to each study methodology. However, we cannot exclude that changes during the studies observation period might have influenced relative year-to-year variation. Rescaling also removed differences in RSVH attack rates between countries, thus true variation may be larger than accounted for in the scenarios. To assess the impact of rescaling, scenarios were run using the original attack rates, before rescaling (Supplementary Fig. 2). Without rescaling, including more seasons still had the largest effect on statistical power while the effect of including more countries was inconsistent. Second, VE was assumed equal between seasons and between countries. However, maternal VE depends on placental transfer of antibodies, which could vary with time or between countries. Vaccinating late in pregnancy can decrease newborns antibody titers [41,42]. Placental malaria and HIV infection were shown to reduce transfer of antibodies against malaria [43,44]. Third, scenarios assumed the same number of participants per country and per season. In field trials, few sites often contribute to a majority of the recruitment [18]. The shortfall of events resulting from low attack rates in a major site would be more difficult to compensate than in the scenarios. Overall, rescaling, assuming stable VE and equal recruitment limit the generalizability and direct applicability of the results as these parameters can independently impact the outcome of the trial. Also, all scenarios had a 1:1 randomization ratio. Unequal randomization ratios lead to higher sample size for the same power [45]. However, the objective was not to estimate sample size. Finally, all calculations were based on Fisher's exact test, known to be conservative [46].

Our results illustrate how year-to-year variation can decrease the ability of trials to achieve the primary endpoint if it is not considered at the design stage. We described possible avenues to limit and monitor the potential impact, including adding more seasons or using adaptive trial designs. Sensitivity to overestimating incidence due to normal year-to-year variation is a peculiarity of vaccine trials that is often overlooked in sample size discussions. This is particularly relevant for the RSV field as more phase III trials are expected in the near future, and RSV attack rates may be less predictable and show different variations due to the disruption caused by the COVID-19 pandemic. Of note, similar approaches can be applied to other seasonal diseases such as influenza or malaria.

Acknowledgments

This work was supported by the Respiratory Syncytial virus Consortium in Europe (RESCEU), European Union. RESCEU has received funding from the Innovative Medicines Initiative, European Union (IMI) two Joint Undertaking under grant agreement No 116019. This Joint Undertaking receives support from the European Union's Horizon 2020, European Union research and innovation program and the European Federation of Pharmaceutical Industries and Associations, Belgium (EFPIA). The contents are the responsibility of the authors and neither IMI nor the European Union, EFPIA, or any associated partners are responsible for any use that may be made of the information contained therein.

Study Group Members: The RESCEU investigators are as follows: Marie-Noelle Billard, Katrien Oude Rengerink, Louis Bont, Joanne Wildenbeest (University Medical Centre Utrecht); Harish Nair, Harry Campbell (University of Edinburgh); Philippe Beutels (University of Antwerp); Peter Openshaw, (Imperial College London); Andrew Pollard (University of Oxford); Federico Martinon-Torres (Servicio Galego de Saude); Terho Heikkinen (University of Turku and Turku University Hospital); Adam Meijer (National Institute for Public Health and the Environment); Thea Kølsen Fischer (Statens Serum Institut); Maarten van den Berge (Academisch Ziekenhuis Groningen); Carlo Giaquinto (Fondazione PENTA for the Treatment and Care of Children with HIV-ONLUS); Jeroen Aerssens (Janssen); Michael Abram (AstraZeneca); Kena Swanson (Pfizer); Clarisse Demont, Scott Gallichan (Sanofi Pasteur); Veena Kumar (Novavax); Sonia Stoszek (GlaxoSmithKline); Eva Molero (Team-It Research).

References

[1] Friedman LM, Furberg CD, DeMets D. Fundamentals of Clinical Trials. 4th ed. New York, NY: Springer-Verlag; 2010. Available at: https://www.springer.com/gp/book/9781441915863. Accessed March 25, 2021.

[2] Fogel DB. Factors associated with clinical trials that fail and opportunities for improving the likelihood of success: a review. Contemp Clin Trials Commun 2018;11:156−64.

[3] Graham SM, Shah PS, Aesch ZC, Beyene J, Bayoumi AM. A systematic review of the quality of trials evaluating biomedical HIV prevention interventions shows that many lack power. HIV Clin Trials 2009;10(6):413−31.

[4] Griffin MP, Yuan Y, Takas T, Domachowske JB, Madhi SA, Manzoni P, et al. Single-dose nirsevimab for prevention of RSV in preterm infants. N Engl J Med 2020;383:415−25.

[5] Madhi SA, Polack FP, Piedra PA, Munoz FM, Trenholme AA, Simoes EA, et al. Vaccination of pregnant women with respiratory syncytial virus vaccine and protection of their infants. N Engl J Med 2020;383:426−39.

[6] Gaithersberg M. Novavax announces topline RSV F vaccine data from two clinical trials in older adults | Novavax Inc. - IR site. 2016. Available at: https://ir.novavax.com/news-releases/news-release-details/novavax-announces-topline-rsv-f-vaccine-data-two-clinical-trials. Accessed May 2, 2021.

[7] Kuitunen I, Artama M, Mäkelä L, Backman K, Heiskanen-Kosma T, Renko M. Effect of social distancing due to the COVID-19 pandemic on the incidence of viral respiratory tract infections in children in Finland during early 2020. Pediatr Infect Dis J 2020;39(12):e423−7.

[8] Tang JW, Bialasiewicz S, Dwyer DE, Dilcher M, Tellier R, Taylor J, et al. Where have all the viruses gone? Disappearance of seasonal respiratory viruses during the COVID-19 pandemic. J Med Virol 2021;93(7):4099−101.

[9] Varela FH, Scotta MC, Polese-Bonatto M, Sartor IT, Ferreira CF, Fernandes IR, et al. Absence of detection of RSV and influenza during the COVID-19 pandemic in a Brazilian cohort: Likely role of lower transmission in the community. J Glob Health 2021;11:05007.

[10] Zheng Z, Pitzer VE, Shapiro ED, Bont LJ, Weinberger DM. Intensity of reemergence of respiratory syncytial virus following the COVID-19 pandemic in the US. JAMA Netw Open 2021;4(12).

[11] Baker RE, Park SW, Yang W, Vecchi GA, Metcalf CJE, Grenfell BT. The impact of COVID-19 nonpharmaceutical interventions on the future dynamics of endemic infections. Proc Natl Acad Sci U S A 2020;117(48):30547−53.

[12] Wang X, Ji X. Sample size estimation in clinical research: from randomized controlled trials to observational studies. Chest 2020;158:S12−20.

[13] Noordzij M, Tripepi G, Dekker FW, Zoccali C, Tanck MW, Jager KJ. Sample size calculations: basic principles and common pitfalls. Nephrol Dial Transplant 2010;25(5):1388−93.

[14] Harden M, Friede T. Sample size calculation in multi-centre clinical trials. BMC Med Res Methodol 2018;18:156.

[15] Chan ISF, Bohidar NR. Exact power and sample size for vaccine efficacy studies. Commun Stat 1998;27(6):1305−22.

[16] Schulgen G, Olschewski M, Krane V, Wanner C, Ruf G, Schumacher M. Sample sizes for clinical trials with time-to-event endpoints and competing risks. Contemp Clin Trials 2005;26:386−96.

[17] Li Y, Reeves RM, Wang X, Bassat Q, Brooks WA, Cohen C, et al. Global patterns in monthly activity of influenza virus, respiratory syncytial virus, parainfluenza virus, and metapneumovirus: a systematic analysis. Lancet Glob Health 2019;7(8):e1031−45.

[18] Shi T, McAllister DA, O'Brien KL, Simoes EAF, Madhi SA, Gessner BD, et al. Global, regional, and national disease burden estimates of acute lower respiratory infections due to respiratory syncytial virus in young children in 2015: a systematic review and modelling study. Lancet 2017;390(10098):946−58.

[19] Higgins D, Trujillo C, Keech C. Advances in RSV vaccine research and development − a global agenda. Vaccine 2016;34(26):2870−5.

[20] Munoz FM. Respiratory syncytial virus in infants: is maternal vaccination a realistic strategy? Curr Opin Infect Dis 2015;28:221−4.

[21] PATH. RSV vaccine and mAb snapshot. 2021. Available at: https://www.path.org/resources/rsv-vaccine-and-mab-snapshot/. Accessed May 2, 2021.

[22] Organization WH. WHO preferred product characteristics for respiratory syncytial virus (RSV) vaccines. World Health Organization.

2017. Available at: https://apps.who.int/iris/handle/10665/258705. Accessed March 25, 2021.
[23] Madhi SA, Polack FP, Piedra PA, Munoz FM, Trenholme AA, Simoes EAF, et al. Respiratory syncytial virus vaccination during pregnancy and effects in infants. N Engl J Med 2020;383:426−39.
[24] Baral R, Li X, Willem L, Antillon M, Vilajeliu A, Jit M, et al. The impact of maternal RSV vaccine to protect infants in Gavi-supported countries: estimates from two models. Vaccine 2020;38(33): 5139−47.
[25] Chavez D, Gonzales-Armayo V, Mendoza E, Palekar R, Rivera R, Rodriguez A, et al. Estimation of influenza and respiratory syncytial virus hospitalizations using sentinel surveillance data—La Paz, Bolivia. 2012−2017. Influenza Other Respir Viruses 2019;13(5): 477−83.
[26] Hall CB, Weinberg GA, Iwane MK, Blumkin AK, Edwards KM, Staat MA, et al. The burden of respiratory syncytial virus infection in young children. N Engl J Med 2009;360:588−98.
[27] Nokes D, Ngama M, Bett A, Abwao J, Munywoki P, English M, et al. Incidence and severity of respiratory syncytial virus pneumonia in rural Kenyan children identified through hospital surveillance. Clin Infect Dis 2009;49(9):1341−9.
[28] Vicente D, Montes M, Cilla G, Perez-Yarza EG, Perez-Trallero E. Hospitalization for respiratory syncytial virus in the paediatric population in Spain. Epidemiol Infect 2003;131(2):867−72.
[29] Weigl JaI, Puppe W, Schmitt HJ. Seasonality of respiratory syncytial virus-positive hospitalizations in children in Kiel, Germany, over a 7-year period. Infection 2002;30(4):186−92.
[30] Fay MP. Confidence intervals that match Fisher's exact or Blaker's exact tests. Biostatistics 2010;11(2):373−4.
[31] Bosomprah S. Sample size calculation for multicentre efficacy trials of blood-stage malaria antigens. Malar J 2013;12(1):253.
[32] Shenoy P. Multi-regional clinical trials and global drug development. Perspect Clin Res 2016;7(2):62−7.
[33] Obando-Pacheco P, Justicia-Grande AJ, Rivero-Calle I, Rodriguez-Tenreiro C, Sly P, Ramilo O, et al. Respiratory syncytial virus seasonality: a global Overview. J Infect Dis 2018;217:1356−64.
[34] Schulz KF, Grimes DA. Sample size calculations in randomised trials: mandatory and mystical. Lancet 2005;365(9467): 1348−53.
[35] Stallard N, Whitehead J, Todd S, Whitehead A. Stopping rules for phase II studies. Br J Clin Pharmacol 2001;51(6):523−9.
[36] Pallmann P, Bedding AW, Choodari-Oskooei B, Dimairo M, Flight L, Hampson LV, et al. Adaptive designs in clinical trials: why use them, and how to run and report them. BMC Med 2018;16:29.
[37] Liu M, Li Q, Lin J, Lin Y, Hoffman E. Innovative trial designs and analyses for vaccine clinical development. Contemp Clin Trials 2021;100:106225.
[38] Gröndahl B, Puppe W, Weigl J, Schmitt H-J. Comparison of the BD Directigen Flu A+B Kit and the Abbott TestPack RSV with a multiplex RT-PCR ELISA for rapid detection of influenza viruses and respiratory syncytial virus. Clin Microbiol Infect 2005;11(10):848−50.
[39] Thomas EE, Book LE. Comparison of two rapid methods for detection of respiratory syncytial virus (RSV) (Testpack RSV and ortho RSV ELISA) with direct immunofluorescence and virus isolation for the diagnosis of pediatric RSV infection. J Clin Microbiol 1991;29(3):632−5.
[40] Voermans JJC, Seven-Deniz S, Fraaij PLA, van der Eijk AA, Koopmans MPG, Pas SD. Performance evaluation of a rapid molecular diagnostic, MultiCode based, sample-to-answer assay for the simultaneous detection of Influenza A, B and respiratory syncytial viruses. J Clin Virol 2016;85:65−70.
[41] Vaz-de-Lima LRA, Sato HK, Fernandes EG, Sato APS, Pawloski LC, Tondella ML, et al. Association between the timing of maternal vaccination and newborns' anti-pertussis toxin antibody levels. Vaccine 2019;37(36):5474−80.
[42] Abu Raya B, Srugo I, Kessel A, Peterman M, Bader D, Gonen R, et al. The effect of timing of maternal tetanus, diphtheria, and acellular pertussis (Tdap) immunization during pregnancy on newborn pertussis antibody levels − a prospective study. Vaccine 2014;32(44):5787−93.
[43] Dechavanne C, Cottrell G, Garcia A, Migot-Nabias F. Placental malaria: decreased transfer of maternal antibodies directed to plasmodium falciparum and impact on the incidence of febrile infections in infants. PLoS One 2015;10:e0145464.
[44] Alonso S, Vidal M, Ruiz-Olalla G, González R, Jairoce C, Manaca MN, et al. HIV infection and placental malaria reduce maternal transfer of multiple antimalarial antibodies in Mozambican women. J Infect 2021;82(4):45−57.
[45] Hey SP, Kimmelman J. The questionable use of unequal allocation in confirmatory trials. Neurology 2014;82:77−9.
[46] Crans GG, Shuster JJ. How conservative is Fisher's exact test? A quantitative evaluation of the two-sample comparative binomial trial. Stat Med 2008;27:3598−611.