medRxiv preprint doi: https://doi.org/10.1101/2024.05.24.24307680; this version posted May 24, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Role of human challenge trials (HCTs) in drug development for respiratory syncytial virus (RSV)

Prosenjit Kundu, PhD[1], Mark Quinn, MPH[2], Jared C. Christensen, PhD[1], Elaine Thomas, PhD[2], Sima S. Toussi, MD[3], Negar Niki Alami, MD[3], Anindita Banerjee, PhD[1*]

[1]Pfizer Inc., Cambridge, MA, USA
[2]USA Pfizer Ltd, London, UK
[3]Pfizer Inc., New York, NY

Abstract

Human challenge trials (HCTs) are powerful in establishing early proof-of-concept for experimental drugs and understanding disease pathogenesis. In this study, we conducted a comprehensive assessment of HCTs to understand the viral load dynamics and symptom score kinetics of respiratory syncytial virus (RSV) to facilitate drug development for RSV. We estimated viral load (VL) and symptom related measures in placebo and relative reduction (RR) of the measures in the experimental drug using a systematic search on double-blind, placebo-controlled RSV HCTs from clinical databases. We used random-effects meta-analysis, except for time to mean peak VL and time to mean peak total symptom score (TSS), where descriptive statistics were summarized. Number of studies varied across measures, from 4 (151 subjects in total) to 7 (247 subjects in total). Overall, treatment reductions of 54% (95% CI: 32% – 76%, I^2 = 91%) and 76% (95% CI: 68% – 85%, I^2 = 19%) are observed in mean VL area under curve (AUC) and mean TSS AUC, respectively. In placebo, mean time to mean peak VL/TSS was 6.99/7.09 days (95% CI: 6.24/6.06 – 7.74/8.12 days). In comparison to other viruses in HCTs, peak VL occurs earlier for influenza (1.7 days) and at similar time for SARS-nCoV-2 (7 days), whereas peak TSS occurs earlier for influenza (2.9 days) and later for SARS-nCoV-2 (7.9 days). Our findings will inform researchers on disease course and VL kinetics, critical data needed for designing RSV treatment studies and understanding implications in clinical practice.

*Correspondence to: Anindita Banerjee, Pfizer Inc., 1 Portland St, Cambridge, MA-02139, US, anindita.banerjee@pfizer.com

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2024.05.24.24307680; this version posted May 24, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Introduction:

Respiratory Syncytial Virus (RSV) remains a major global health concern, particularly among vulnerable populations, including infants, the elderly, and individuals with underlying medical conditions[1,2]. In the United States and globally, RSV infection continues to cause severe respiratory illnesses in high-risk populations, often requiring hospitalization and posing a significant burden on healthcare systems[3-7]. In response to this ongoing public health challenge, numerous randomized clinical trials have been conducted to assess the efficacy of various anti-viral treatments for RSV as a proof of mechanism. However, the landscape of RSV treatment is complex and remains an unmet medical need, with a wide array of interventions being tested, and outcomes that vary across studies.

Human challenge trials (HCTs), also known as controlled human infection studies, involve intentionally exposing participants to a pathogen to test the effectiveness of treatments or vaccines[8]. For drug development it is important to assess proof of mechanism and assess signal of efficacy early in development. A human challenge model allows us to understand the effect of the drug early during development and establish proof of mechanism.[9] By running these standard human challenge studies, it is feasible to gauge some sign of efficacy before running larger late phase studies and de-risk later development. For RSV several companies have studied investigational drugs using the available human challenge model.

In this manuscript, we provide the first comprehensive assessment of placebo measures and treatment efficacy reported in RSV human challenge studies. By conducting meta-analysis[10,11] a widely popular statistical approach for combining studies, we seek to synthesize the available evidence, elucidate RSV VL kinetics and characterize the effectiveness of various interventions for RSV based on data from the HCTs. The main objectives of this meta-analysis include summarizing the collective findings of RSV human challenge studies from published literature or/and online clinical databases, identifying potential sources of heterogeneity across studies, such as patient populations, intervention types, and study designs, assessing overall placebo measures and treatment responses for outcomes from human challenge studies. These outcomes include viral load (VL) area under the curve (AUC), total symptom score (TSS) AUC, peak VL, time to peak VL and time to peak TSS. The estimated overall measures in placebo from this meta-analysis would inform the design of future RSV clinical trials. Further, we compare the placebo measures for some of the outcomes with those from published literature for influenza and SARS-nCoV-2[12-14], to understand the clinical profile of RSV.

Results:

Study Characteristics

Following the search strategy from PRISMA guidelines[15], seven HCTs[16-22] were identified to be included in the meta-analysis (Figure 1). The search started with a total of 428 RSV studies from clinical databases and literature websites with key words “respiratory syncytial virus”, “RSV”, “healthy volunteers”, “healthy patients”, “healthy cohort, “healthy participants”, “healthy adults”. The search strategy was limited to English language papers only and publication from 1990 to

medRxiv preprint doi: https://doi.org/10.1101/2024.05.24.24307680; this version posted May 24, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

present, with any pharmaceutical product as the intervention. The websites ClinicalTrials.gov and EudraCT were searched with the key words "respiratory syncytial virus", "RSV", "human challenge", "viral challenge". Out of the 428 studies, 11 studies were from HCTs. We excluded 4 studies, one due to unavailability of data on the outcomes considered in our analyses, one due to study design and objective (pre-exposure prophylaxis treatment) as our study aims at treatments after inoculation to challenge virus, one due to assay used other than RT-qPCR as our study looked at VL measurements assessed from RT-qPCR assay and one due to vaccine as the investigational treatment. The exclusion criterion led to 7 studies included in our analyses. These studies covered 7 different interventions in development, two of which are currently being evaluated in Phase 2 or Phase 2/3 trials, EDP-938 and Sisunatovir (RV521), respectively. The study characteristics, including demographics, are reported in Table 1. The number of subjects varied from 45 to 115 with a total of 544 subjects (including multiple cohorts from each studies), average age across the studies ranged from 23 to 28 years, proportion of females varied from 25 to 38% except for one study from Johnson & Johnson[21] which recruited male subjects only, the year of recruitment ranged from 2012 to 2018 and all were UK-based studies.

Assessment of Risk of Bias

We reported the assessment of risk of bias for each of the HCT's in Supplementary Figure 1. Overall, the biases are low for most of the domains as the studies are well-controlled, quarantine and are run at a single site by the same vendor.

VL specific outcome measures

Mean placebo VL AUC

This analysis included 7 studies ranging in size from 12 to 30 participants, with a total of 128 participants. Individually, the placebo mean VL AUC dose varied from 432.80 to 790.20 $\log_{10}$PFUe.hr/mL (Figure 2). The I^2 statistic was used to assess the heterogeneity across studies and was equal to 82% ($p < 0.01$) indicating considerable degree of heterogeneity (Figure 2). Under a random effects model, the placebo mean VL AUC was 605.98 $\log_{10}$PFUe.hr/mL (95% CI: 491.60 – 720.37 $\log_{10}$PFUe.hr/mL) (Figure 2).

Relative reduction (RR) of mean VLAUC in treatment group compared to placebo

This analysis involved 7 studies ranging in size from 20 to 55 participants and included a total of 247 participants. Individually, the RR of mean VL AUC in treatment group compared to placebo, varied from 1% to 88% (Figure 3). The I^2 statistic was used to assess the heterogeneity across studies and was equal to 91% ($p < 0.01$) indicating considerable degree heterogeneity (Figure 3). Under a random effects model, the RR of mean VL AUC was 54% (95% CI: 32% – 76%) compared to placebo (Figure 3).

Mean placebo VL at peak

This analysis included 7 studies ranging in size from 12 to 30 participants, with a total of 128 participants. Individually, the mean placebo VL at peak varied from 4.7 to 6.24 $\log_{10}$PFUe/mL

medRxiv preprint doi: https://doi.org/10.1101/2024.05.24.24307680; this version posted May 24, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

(Figure 4). The I^2 statistic was used to assess the heterogeneity across studies and was equal to 55% (p = 0.04) indicating a moderate degree of heterogeneity (Figure 4). Under a random effects model, the mean placebo VL at peak was 5.38 $\log_{10}$PFUe/mL (95% CI: 4.94 – 5.81 $\log_{10}$PFUe/mL) (Figure 4).

Time (in days) to mean placebo peak VL

Since most of the studies reported time to mean peak VL rather than mean time to peak VL for placebo subjects, we summarized the descriptive statistics for time to mean peak VL which are evaluated assuming each study serves as a random subject. This analysis involved 6 studies ranging in size from 12 to 30 participants, a total of 124 participants were included. Individually, the time to mean placebo peak VL from inoculation ranged from 6.2 to 7.75 days. The mean and median time to mean placebo peak VL were 6.99 days (95% CI: 6.24 – 7.74 days) and 6.74 days (95% CI: 6.2 – 8 days), respectively. In comparison to other viral diseases, time to mean peak VL occurs much earlier for influenza (~ 2 days)[13,14] and at a similar time in disease course for SARS-nCoV-2 (~7.3 days)[12].

Symptom specific outcome measures

Mean placebo total symptom score area under the curve (TSS AUC)

This analysis involved 4 studies ranging from 12 to 30 participants, 80 participants in total were included. Individually, the mean placebo TSS AUC varied from 269.40 to 606.90 (Supplementary Figure 2). The I^2 statistic was used to assess the heterogeneity across studies and was equal to 48% (p = 0.12) indicating a moderate degree of heterogeneity (Supplementary Figure 2). Under a random-effects model, the placebo mean TSS AUC was 404.96 (95% CI: 272.84 - 537.08) (Supplementary Figure 2).

Relative Reduction (RR) of mean total symptom score area under the curve (TSS AUC)

This analysis involved 4 studies ranging in size from 20 to 55 participants, a total of 151 participants were included. Individually, the RR of mean TSS AUC in the treatment group compared to placebo varied from 29% to 82% (Figure 5). The I^2 statistic was used to assess the heterogeneity across studies and was equal to 19% (p = 0.30) indicating low heterogeneity across the studies (Figure 5). As a result, both common effect and random effects model produced the same estimates of the RR of mean TSS AUC in the treatment group compared to placebo which was 76% (95% CI: 68% - 85%) (Figure 5).

Time (in days) to mean placebo peak total symptom score (TSS)

We report the descriptive statistics as most of the studies reported time to mean peak TSS instead of mean time to peak TSS for placebo subjects. The statistics were evaluated in an analogous way to those described above for VL. This analysis involved 5 studies ranging from 12 to 30 participants, 107 participants in total were included. Individually, time to mean placebo peak TSS from inoculation varied from 6 to 8.04 days. The mean and median time to mean placebo peak TSS were 7.09 days (95% CI: 6.06 – 8.12 days) and 6.87 days (95% CI: 6 – 8.04 days) respectively.

medRxiv preprint doi: https://doi.org/10.1101/2024.05.24.24307680; this version posted May 24, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

In comparison to other viral diseases, peak TSS occurs earlier for influenza (~3 days)[13,14] and later in disease course for SARS-nCoV-2 (~9.4 days)[12].

Prediction Interval for a future HCT and Reference Interval for an individual recruited in a future HCT

We report the prediction interval and reference range for placebo VL AUC, placebo VL at peak and placebo TSS AUC. From Supplementary Table 1, the 95% prediction interval for mean placebo VL AUC, mean placebo VL at peak and mean placebo TSS AUC in a future HCT are (218.97 – 993.0 $\log_{10}$PFUe.hr/mL), (4.13 – 6.63 $\log_{10}$PFUe/mL) and (0 – 892.11), respectively. From Supplementary Table 2, based on a frequentist approach, the 95% reference interval or the "normal" range where 95% of the measurements from individuals recruited in a future HCT would lie are (0 – 1295.54 $\log_{10}$PFUe.hr/mL), (1.58 – 9.18 $\log_{10}$PFUe/mL) and (0 – 1288.25) for placebo VL AUC, placebo VL at peak and placebo TSS AUC. We have also provided the plots including prediction intervals at study-level and reference interval at an individual-level (Supplementary Figure 3, Supplementary Figure 4, Supplementary Figure 5).

Discussion:

We conducted the first of its kind systematic review and meta-analysis of RSV outcomes from human challenge trials. We provided overall estimates for placebo group and relative reduction in the treatment compared to placebo for the key measures including amount of viral load (VL) and total symptom score (TSS) based on 10-symptom diary card, both evaluated during the window of quarantine period. Further, where traditional meta-analysis was not possible, we used descriptive statistics to provide overall estimates for some of the key time measures such as time to peak VL and TSS. We compared these estimates with influenza and SARS-nCoV-2 since these are anti-viral diseases that have approved treatments. Comparing the viral dynamics for RSV disease course to other viral diseases, it appears that the time to mean peak VL is similar to SARS-nCoV-2 but occurs much later compared to influenza. Moreover, comparing the symptom dynamics of these viral diseases with RSV, peak TSS occurs later for SARS-nCoV-2, but early on for influenza. Additionally, it appears that the mean time to reach the peak VL and the mean time to reach the peak TSS coincide for RSV. However, there is a delay of ~ 2 days for SARS-nCoV-2 and ~1 day for influenza in reaching the peak TSS compared to peak VL. Additionally, we provided 95% prediction interval for mean placebo VL AUC, mean placebo VL at peak and mean placebo TSS AUC in a new HCT and 95% reference interval for placebo VL AUC, placebo VL at peak and placebo TSS AUC at an individual-level.

Generally, RSV challenge studies assess for RSV positivity from day 2 to day 5 post-inoculation prior to randomizing participants to study treatment. However, typically by day 5 it is assumed that all participants may be RSV positive and are therefore randomized. The analysis for all the efficacy endpoints in challenge studies is designed for RSV positive randomized patients, however, a small proportion of the subjects in the challenge studies may become positive after randomization and therefore don't fully represent the clinical treatment situation. Imbalance in the viral load and symptom severity, prior infection timing across the underlying populations from challenge studies and varying mechanism of action of drugs across the challenge studies could be a potential source of heterogeneity, in addition to residual confounding due to unmeasured confounders.

medRxiv preprint doi: https://doi.org/10.1101/2024.05.24.24307680; this version posted May 24, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Some of the potential limitations of our study can be attributed to the drawbacks of HCTs[9]. For example, study sample size and number of studies are usually low due to ethical concerns and safety risks. Available RSV HCT data evaluate treatment response in healthy adults with known baseline RSV level seropositivity by controlled challenge strain that causes upper respiratory tract infection limiting the evaluation of treatment response for lower respiratory tract infection. Findings from challenge trials may not be directly applicable to the general population, as participants are young, healthy, White and/or Caucasian adults, and their response to the pathogen may differ from other demographic groups and those with co-morbidities or risk factors for severe RSV disease. Results from challenge trials may be specific to the target population and the strain of the pathogen used (the challenge virus in our review is RSV-A Memphis-37b), potentially limiting the generalizability of the findings to other strains and other populations. Future studies and research are needed to understand and assess the disease mechanisms for individuals in placebo and treatment groups for other strains such as RSV-B and should include the ethnic/racial populations that are under-represented in the current HCTs, and the young and older individuals with comorbidities that bare the greatest burden of disease. Other limitations of our study can be attributed to inconsistent availability of outcomes. For example, not all the study outcomes are reported in all seven studies (only 4 studies report mean TSS AUC) restricting the power for those outcomes. Studies mostly reported time to mean peak VL/TSS instead of mean time to peak VL/TSS which restricted us from performing the traditional meta-analysis.

Despite these limitations, our comprehensive assessment of the RSV HCTs provides detailed understanding of the VL and symptom score kinetics that will enhance understanding of RSV disease pathogenesis. Meta-analysis and descriptive statistics of the key measures allowed us to evaluate the overall placebo group characteristics and relative reduction by combining information across studies, thus increasing the overall sample size, and addressing the limitation of low sample size from a single HCT. Reference interval for some of the key measures will guide the clinician to determine if the patient's measurement lies within the normal range, and thus validating the patient's inclusion in the HCT. Our study mainly focused on HCTs that used RSV-A Memphis-37b as the challenge virus in the experimental stage and therefore the results from our study are applicable to RSV-A which is the most prevalent strain These results can be used to contextualize effects to other potential experimental drugs when tested on the same strain. Overall, findings from our analysis can help researchers and clinicians in significantly accelerating the development of RSV treatment interventions.

Methods:

Search Strategy/Study Eligibility

The search parameters for studies to be included in this meta-analysis were as follows; randomized controlled trials (RCT) that featured an intervention of a small molecule antiviral investigational product for the indication respiratory syncytial virus (RSV). These studies had to be conducted under the human viral challenge model, with healthy volunteers, and had to have a total number of inoculated participants greater than ten. Additionally, we only included studies that had viral load AUC as the primary endpoint and measured at least one of the following: peak viral load, time to peak viral load, peak total symptom score, total symptom score AUC, and time to peak

medRxiv preprint doi: https://doi.org/10.1101/2024.05.24.24307680; this version posted May 24, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

symptom score. The search was conducted using Biosis Previews, Embase, Ovid MEDLINE, PubMed, ClinicalTrials.gov and EudraCT from data inception to August 2023. The search strategy we used was provided by a cross-functional team of researchers with expertise in conducting systematic reviews and meta-analyses on RSV research (all authors).

Assessment of Risk of Bias

Bias from each of the studies are assessed independently by four reviewers based on Cochrane's domain-based criterion. The risk of bias is evaluated using the visualization tool robvis[23].

Statistical Analysis

We considered seven different outcome measures, broadly classified into VL specific measures and symptom specific measures. VL specific measures included (1) mean placebo VL AUC where the AUC is calculated from first dose until end of quarantine period (window length of ~12-13 days for the studies included), (2) relative reduction (RR) of mean VL AUC in treatment group compared to placebo from the same study, (3) mean placebo VL at peak, (4) time (in days) to mean placebo peak VL. Symptom specific measures included (5) mean placebo TSS AUC where the window length is similar to mean placebo VL AUC in (1), (6) RR of mean TSS AUC in treatment group compared to placebo from the same study, and (7) time (in days) to mean placebo peak TSS. Both the time variables (4) and (7) are measured from inoculation. We derived the estimate of RR, (2) and (6) from the corresponding treatment and placebo estimates reported in the studies. For deriving the time variables (4) and (7), we assumed a uniform distribution for the time (in days) from RSV inoculation to first dose assignment, where the range varied from the day of inoculation to the day of first dose assignment. This assumption is applicable for the studies which did not report data on time variables (4) and (7) from inoculation but from assignment of first dose. In studies that reported multiple doses for treatment regimens, we chose the cohort with the highest dose for evaluating RR measures. Further, most (n=5) of the studies report the unit of VL AUC in log10 PFUe.hr/mL (log10 PFUe.hr/mL = log (base 10) plaque-forming unit equivalents per milliliter times hour), while some (n=2) report in log10 copies.hr/mL. For our analysis we assumed equivalence between these units as confirmed as appropriate by the hVIVO. hVIVO is a contract research organization specializing in infectious disease vaccines and therapeutics using human challenge clinical trials[24].

We constructed a database for each of the study outcome measures. The databases included the corresponding outcome measure and associated SD for all the studies passing the inclusion/exclusion criterion (see Figure 1). Details of evaluation of mean and SD from the literature is provided in Supplementary Methods. Frequentist random-effects model (using restricted maximum likelihood estimation)[25] was implemented for each of the study outcomes (except for time to mean peak VL and time to mean peak TSS) using the R package "meta" [26], where the constructed database associated with the respective study outcome served as an input for the function "metamean" from the "meta" package. Forest plots were generated using the function "forest.meta". We evaluate the prediction interval of a new study and reference interval for an individual recruited in HCT using the Rshiny app: RIMeta[27]. A 95% reference interval for a particular measure is defined to be the interval in which we would expect 95% of the individual's measurements from a future HCT to lie.

medRxiv preprint doi: https://doi.org/10.1101/2024.05.24.24307680; this version posted May 24, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Most of the studies reported time to mean peak VL/TSS instead of mean time to peak VL/TSS. Mean time to peak VL/TSS is assessed by first calculating the day of peak VL/TSS for each subject and then taking the mean over days. In this case, each study would have provided a mean estimate over subjects and therefore would have fallen into the framework of traditional meta-analysis. However, on the other hand, time to mean peak VL/TSS is assessed by first calculating the mean VL/TS curve over subjects at each day and then looking at the day where the peak is achieved. Each study, therefore, provided a single observation for the time (in days) to peak VL/TSS. We treated each study as a subject and used descriptive statistics (instead of traditional meta-analysis) such as mean and median to provide overall estimates for time to mean peak VL/TSS accounting for study-level uncertainty. 95% CI for median is calculated non-parametrically from quantiles[28], while for mean, we calculate it from Student's t-distribution.

Acknowledgement

We thank Dr. Haitao Chu, Senior Director and Statistics Lead at Pfizer Inc., for providing insightful comments to the paper.

References:

1. Li Y, Wang X, Blau DM, et al. Global, regional, and national disease burden estimates of acute lower respiratory infections due to respiratory syncytial virus in children younger than 5 years in 2019: a systematic analysis. *Lancet*. May 28 2022;399(10340):2047-2064. doi:10.1016/S0140-6736(22)00478-0
2. Nguyen-Van-Tam JS, O'Leary M, Martin ET, et al. Burden of respiratory syncytial virus infection in older and high-risk adults: a systematic review and meta-analysis of the evidence from developed countries. *Eur Respir Rev*. Dec 31 2022;31(166)doi:10.1183/16000617.0105-2022
3. Hall CB, Weinberg GA, Iwane MK, et al. The burden of respiratory syncytial virus infection in young children. *N Engl J Med*. Feb 5 2009;360(6):588-98. doi:10.1056/NEJMoa0804877
4. Langley JM, Bianco V, Domachowske JB, et al. Incidence of Respiratory Syncytial Virus Lower Respiratory Tract Infections During the First 2 Years of Life: A Prospective Study Across Diverse Global Settings. *J Infect Dis*. Aug 26 2022;226(3):374-385. doi:10.1093/infdis/jiac227
5. McLaughlin JM, Khan F, Begier E, Swerdlow DL, Jodar L, Falsey AR. Rates of Medically Attended RSV Among US Adults: A Systematic Review and Meta-analysis. *Open Forum Infect Dis*. Jul 2022;9(7):ofac300. doi:10.1093/ofid/ofac300
6. McLaughlin JM, Khan F, Schmitt HJ, et al. Respiratory Syncytial Virus-Associated Hospitalization Rates among US Infants: A Systematic Review and Meta-Analysis. *J Infect Dis*. Mar 15 2022;225(6):1100-1111. doi:10.1093/infdis/jiaa752

medRxiv preprint doi: https://doi.org/10.1101/2024.05.24.24307680; this version posted May 24, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

7. Zheng Z, Warren JL, Shapiro ED, Pitzer VE, Weinberger DM. Estimated incidence of respiratory hospitalizations attributable to RSV infections across age and socioeconomic groups. *Pneumonia (Nathan)*. Oct 25 2022;14(1):6. doi:10.1186/s41479-022-00098-x

8. Darton TC, Blohmke CJ, Moorthy VS, et al. Design, recruitment, and microbiological considerations in human challenge studies. *Lancet Infect Dis*. Jul 2015;15(7):840-51. doi:10.1016/S1473-3099(15)00068-7

9. Valdés E. Human Challenge Trials in Pandemics: Risks and Benefits. In: Valdés E, Lecaros JA, eds. *Handbook of Bioethical Decisions Volume I: Decisions at the Bench*. Springer International Publishing; 2023:373-384.

10. Cornell JE, Mulrow CD, Localio R, et al. Random-effects meta-analysis of inconsistent effects: a time for change. *Ann Intern Med*. Feb 18 2014;160(4):267-70. doi:10.7326/M13-2886

11. DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials*. Sep 1986;7(3):177-88. doi:10.1016/0197-2456(86)90046-2

12. Killingley B, Mann AJ, Kalinova M, et al. Safety, tolerability and viral kinetics during SARS-CoV-2 human challenge in young adults. *Nat Med*. May 2022;28(5):1031-1041. doi:10.1038/s41591-022-01780-9

13. Carrat F, Vergu E, Ferguson NM, et al. Time lines of infection and disease in human influenza: a review of volunteer challenge studies. *Am J Epidemiol*. Apr 1 2008;167(7):775-85. doi:10.1093/aje/kwm375

14. Bagga B, Woods CW, Veldman TH, et al. Comparing Influenza and Rsv Viral and Disease Dynamics in Experimentally Infected Adults Predicts Clinical Effectiveness of Rsv Antivirals. *Antiviral Therapy*. 2013;18(6):785-792. doi:10.3851/imp2629

15. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*. Mar 29 2021;372:n71. doi:10.1136/bmj.n71

16. Ahmad A, Eze K, Noulin N, et al. EDP-938, a Respiratory Syncytial Virus Inhibitor, in a Human Virus Challenge. *N Engl J Med*. Feb 17 2022;386(7):655-666. doi:10.1056/NEJMoa2108903

17. DeVincenzo J, Cass L, Murray A, et al. Safety and Antiviral Effects of Nebulized PC786 in a Respiratory Syncytial Virus Challenge Study. *J Infect Dis*. Jun 15 2022;225(12):2087-2096. doi:10.1093/infdis/jiaa716

18. DeVincenzo J, Tait D, Efthimiou J, et al. A Randomized, Placebo-Controlled, Respiratory Syncytial Virus Human Challenge Study of the Antiviral Efficacy, Safety, and Pharmacokinetics of RV521, an Inhibitor of the RSV-F Protein. *Antimicrob Agents Chemother*. Jan 27 2020;64(2)doi:10.1128/AAC.01884-19

19. DeVincenzo JP, McClure MW, Symons JA, et al. Activity of Oral ALS-008176 in a Respiratory Syncytial Virus Challenge Study. *N Engl J Med*. Nov 19 2015;373(21):2048-58. doi:10.1056/NEJMoa1413275

20. DeVincenzo JP, Whitley RJ, Mackman RL, et al. Oral GS-5806 activity in a respiratory syncytial virus challenge study. *N Engl J Med*. Aug 21 2014;371(8):711-22. doi:10.1056/NEJMoa1401184

21. Stevens M, Rusch S, DeVincenzo J, et al. Antiviral Activity of Oral JNJ-53718678 in Healthy Adult Volunteers Challenged With Respiratory Syncytial Virus: A Placebo-Controlled Study. *J Infect Dis*. Jul 24 2018;218(5):748-756. doi:10.1093/infdis/jiy227

medRxiv preprint doi: https://doi.org/10.1101/2024.05.24.24307680; this version posted May 24, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

22. Limited B. A Randomised, Phase 2a, Double-Blind, Placebo-Controlled Study to Evaluate the Safety and Antiviral Activity Against Respiratory Syncytial Virus Infection, and the Pharmacokinetics of Multiple Oral Doses of BTA-C585 in the Virus Challenge Model. 23 Sep 2018. 2015-004296-77. Accessed 15 Aug 2023. https://www.clinicaltrialsregister.eu/ctr-search/trial/2015-004296-77/results

23. McGuinness LA, Higgins JPT. Risk-of-bias VISualization (robvis): An R package and Shiny web app for visualizing risk-of-bias assessments. *Research Synthesis Methods*. 2020/04/26 2020;n/a(n/a)doi:10.1002/jrsm.1411

24. hVIVO. Accessed January, 30, 2024, 2024. https://hvivo.com

25. Langan D, Higgins JPT, Jackson D, et al. A comparison of heterogeneity variance estimators in simulated random-effects meta-analyses. *Res Synth Methods*. Mar 2019;10(1):83-98. doi:10.1002/jrsm.1316

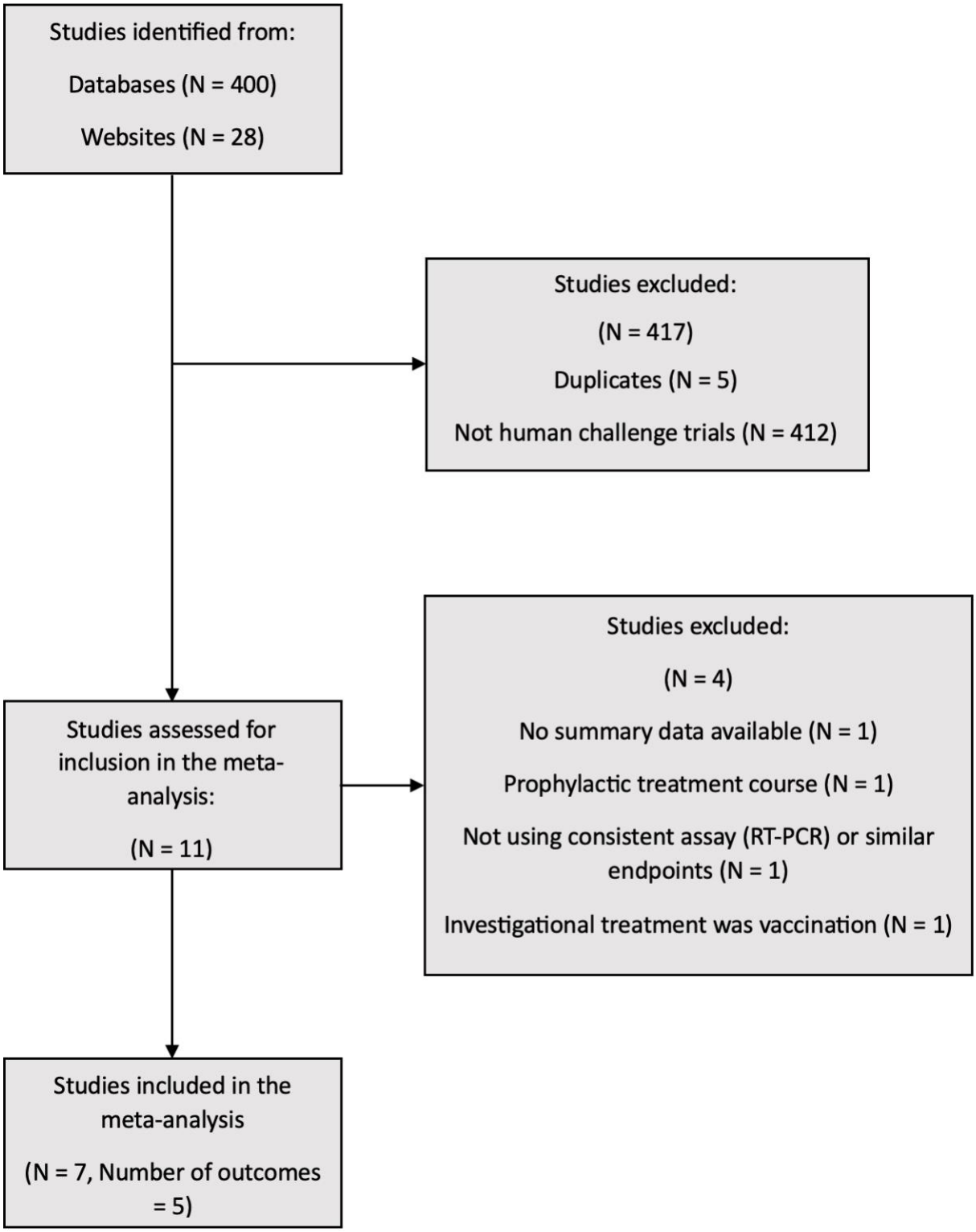
26. Balduzzi S, Rucker G, Schwarzer G. How to perform a meta-analysis with R: a practical tutorial. *Evid Based Ment Health*. Nov 2019;22(4):153-160. doi:10.1136/ebmental-2019-300117

27. Jiang Z, Cao W, Chu H, Bazerbachi F, Siegel L. RIMeta: An R shiny tool for estimating the reference interval from a meta-analysis. *Res Synth Methods*. May 2023;14(3):468-478. doi:10.1002/jrsm.1626

28. Conover WJ. *Practical Nonparametric Statistics*. vol 3rd Edition. Wiley; 1999:608.

medRxiv preprint doi: https://doi.org/10.1101/2024.05.24.24307680; this version posted May 24, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

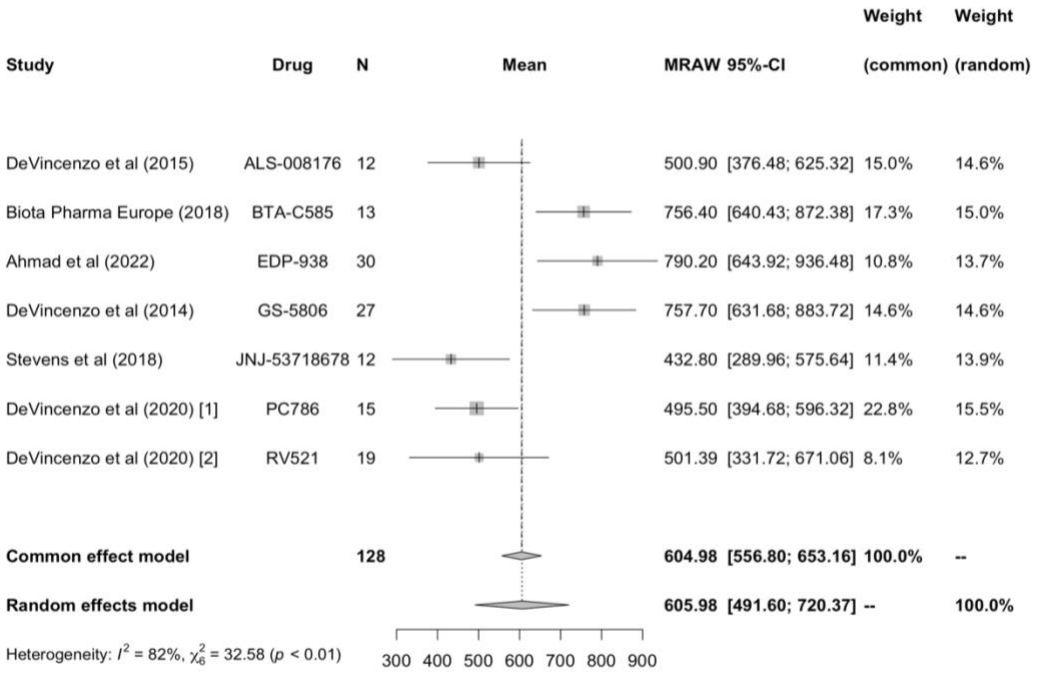
Figure 1: PRISMA Flow Diagram



Note: The meta-analysis is performed for 5 of the 7 study outcomes. Descriptive statistics is reported for the other 2 outcomes.

medRxiv preprint doi: https://doi.org/10.1101/2024.05.24.24307680; this version posted May 24, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

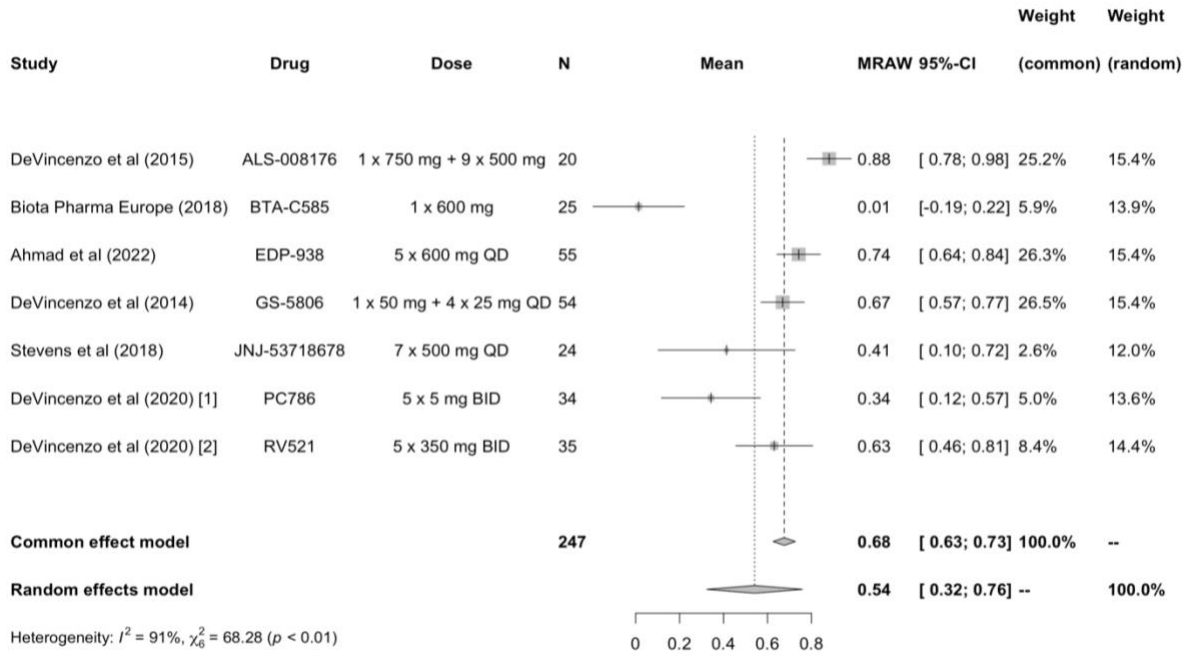
Figure 2: Forest plot for mean viral load area under the curve (VL AUC) in placebo.



Note: VL is assessed using RT-qPCR assay. Studies report the unit of VL in log10 copies/mL or log10 PFUe/mL, where log10 PFUe/mL = log (base 10) plaque-forming unit equivalents per milliliter. In our analysis, we assumed a conversion factor of 1 across the units and is confirmed from hVIVO. Unit of VL AUC is log10 PFUe.hr/mL.

medRxiv preprint doi: https://doi.org/10.1101/2024.05.24.24307680; this version posted May 24, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

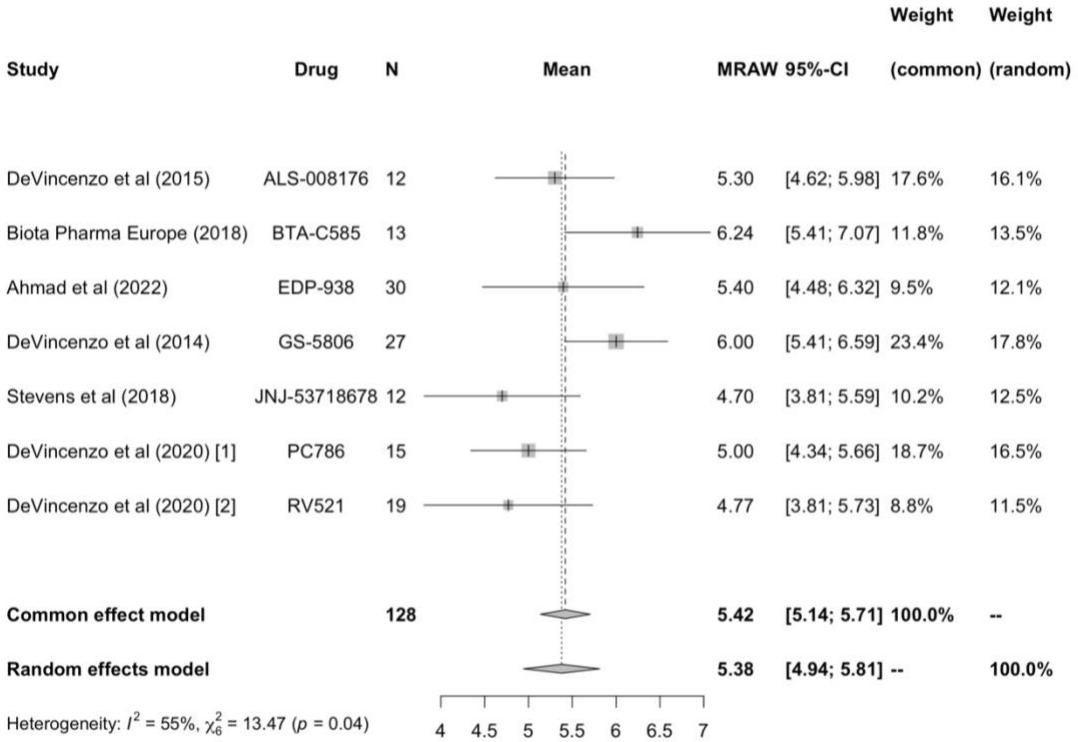
Figure 3: Forest plot for relative reduction (RR) of mean viral load area under the curve (VL AUC) in treatment compared to placebo.



Note: VL is assessed using RT-qPCR assay. Studies report the unit of VL in log10 copies/mL or log10 PFUe/mL, where log10 PFUe/mL = log (base 10) plaque-forming unit equivalents per milliliter. In our analysis, we assumed a conversion factor of 1 across the units and is confirmed from hVIVO. Unit of VL AUC is log10 PFUe.hr/mL.

medRxiv preprint doi: https://doi.org/10.1101/2024.05.24.24307680; this version posted May 24, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

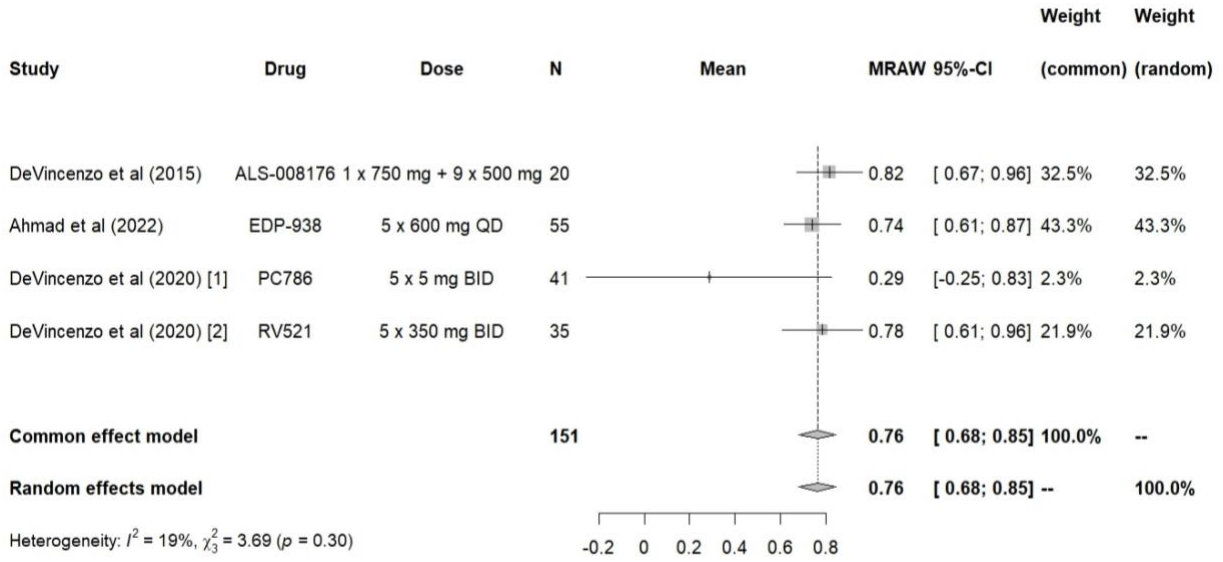
Figure 4: Forest plot for mean viral load at peak in placebo.



Note: VL is assessed using RT-qPCR assay. Studies report the unit of VL in log10 copies/mL or log10 PFUe/mL, where log10 PFUe/mL = log (base 10) plaque-forming unit equivalents per milliliter. In our analysis, we assumed a conversion factor of 1 across the units and is confirmed from hVIVO.

medRxiv preprint doi: https://doi.org/10.1101/2024.05.24.24307680; this version posted May 24, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Figure 5. Forest plot for relative reduction (RR) of mean total symptom score area under the curve (AUC) in treatment compared to placebo.



Note: The total symptom score is evaluated based on 10 clinical symptoms from symptom dairy card. The reported AUC from each study is based on a window length of approximately 12-13 days.