

Supplementary Online Content

Radonovich JE, Simberkoff MS, Bessesen MT, et al. Effect of N95 respirators vs medical masks for preventing laboratory-confirmed influenza among outpatient health care personnel: a randomized clinical trial. *JAMA*. doi:10.1001/jama.2019.11645

Supplement 2. Statistical analysis plan

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

1. Analysis timeline and procedures

A pre-specified analysis plan for the primary manuscript of ResPECT was initially approved by all study principal and study-site lead investigators on April 2016 and updated in May 2017 to reflect the addition of the Laboratory Detected Respiratory Infection (LDRI) outcome to the analysis. At the time of this revision (a) all data collection is complete, (b) all laboratory specimen samples have been tested for the primary and secondary outcomes, and (c) the database housing all of the ResPECT data contained no information about which clinics were assigned to which arm of the study. Once the analysis plan is updated and changes submitted to the site IRB's, the data coordinating center will release labels that identify separate arms of the study to the ResPECT statisticians who will use those codes to implement the analysis as described in this document.

2. General outline of analysis framework

The ResPECT study was a cluster-randomized trial that used constrained randomization (i.e. matching) to ensure balance across arms. The analysis described in this document is an unmatched analysis, i.e. the analysis does not explicitly account for the matching. This has been described as an appropriate approach to analyzing data arising from a matched design. [1]

The final analysis of ResPECT outcome data will consist of intention-to-treat (ITT) and per-protocol analyses (PP) for each of the five study outcomes defined below. For each analysis, the investigators will fit and report results from both adjusted and unadjusted models. Unadjusted models will be analyzed at the cluster-level, and will only include a main effect estimate for the mask and the cluster-level random effects to account for repeated measures of related clusters across multiple seasons. Adjusted models will be analyzed at the individual-level and will include individual-level covariates and random effects to account for repeated measures of the same individual across seasons.

2.a Intention-to-treat analysis

The ITT analysis will include all of the ResPECT participants who were randomized— i.e., those assigned a mask based on their clinic affiliation. Their data will be included according to their treatment assignment, regardless of their adherence to protocol, subsequent withdrawal, failure to provide requested data/samples, or loss to follow-up. This analysis is intended to capture a more realistic outcome of intervention by acknowledging that noncompliance and protocol deviations are an unavoidable part of clinical practice.

In this study, any person who was eligible according to the baseline survey will be included in the ITT analysis. The outcomes for many participants will be missing, particularly those who withdrew during the course of the study. This missingness could conceivably be (a) related to outcome/illness status if individuals were more likely to quit the study because they became sick, or (b) related to the assigned intervention if those assigned one mask over another were more likely to withdraw from study participation. We will assess possible relationships between self-reported reasons for withdrawal and measured variables. Approaches for imputing missing data are addressed below.

2.b Per-protocol analysis

Any participant who completed at least eight weeks of study participation will be included in the per protocol analysis. This strategy will include some participants who only had one blood draw or who are missing reliable serological data due to timing of or lack of information on vaccination (see Participant flow for ResPECT study analysis approaches showing ITT and per protocol cohorts and Decision Algorithm for serological influenza outcome adjudication below). These inclusion/exclusion criteria were decided on by the study PIs (March 2016).

The reasons for missing participant blood samples include loss to follow-up with or without formal withdrawal/deactivation, sample loss due to handling/labeling error, or insufficient sample volume. Since the serologic definition of influenza seroconversion is a 4-fold increase in titer, unpaired serology cannot be assigned an influenza seroconversion status and must be imputed. Missing serologic data will not exclude the patient from the PCR-laboratory assessment. Hence, if an individual is missing a second blood draw but had lab-confirmed influenza by PCR, then this individual will be considered to have had a lab-confirmed influenza outcome. This may create non-random missingness, but it was decided by PIs that since this would not impact many study participants the risk of bias to the overall study was very low.

2.c Handling of missing data via imputation methods

There will be substantial missing data in the outcome (lab-confirmed influenza) and other covariates. The missing data will be imputed using standard multiple imputation techniques, creating imputed datasets with no missing values for each analysis. Each of these datasets will be analyzed using the regression models described below. The results from all of the analyses will be pooled using standard multiple imputation techniques for combining estimates across imputed datasets.[3]

2. d Process for determining participant membership in ITT and per protocol cohorts

Participants signed informed consent. Those who failed to meet inclusion criteria or did not complete screening were excluded. Those who met the inclusion criteria were randomly assigned to a mask group and formulate the ITT cohort. The 'per protocol' cohort will not include those who withdrew before participating (i.e., those who do not fill out any daily or weekly surveys), or discontinued the intervention (withdraw with less than 8 weeks of participation). The 'per protocol' cohort will include those who completed at least 8 weeks of study. The investigators define, for each participant, the amount of time that they participated as the difference between the clinic activation date and latest of either the automatically-generated time-stamp of the last completed daily or weekly survey or the collection date of the last swab, with a maximum of 12 weeks. Those who participated for at least 8 weeks (56 days) according to this calculation will be included in the 'per protocol' cohort. For analyses using person time, the investigators will use the latest of the following; the last survey completed date or collection date from a swab collection.

Decision Algorithm for serological influenza outcome adjudication:

This decision algorithm documents the process for which ResPECT participants will be determined to have had laboratory-confirmed influenza based on serological testing only. The possible outcomes are: laboratory confirmed influenza confirmed by serology (LCI-S) and no laboratory confirmed influenza event confirmed by serology (no LCI-S). In some cases, outcomes (either LCI-S or no LCI-S) will be imputed. The algorithm to classify and/or impute these outcomes is as follows:

Step 1: Determine Study Completion

Determine if participants have completed the study (and thus in the 'per protocol' cohort) or if they have not and thus are in the ITT cohort

Step 2: Determine serological influenza outcome for those in the 'per protocol' cohort

2a. For those individuals in the 'per protocol' cohort who have two serological samples, collected at the beginning and end of the season according to protocol, and who experience a four-fold rise in influenza hemagglutination inhibition antibody (HAI) titer to exactly 0 strains, classify the serological influenza outcome as no LCI-S.

2b. For those individuals in the 'per protocol' cohort who have two serological samples, collected at the beginning and end of the season according to protocol, and who experience a four-fold rise in influenza HAI antibody titer to one or more strains, classify the serological influenza outcome as LCI-S.

2c. For those individuals in the 'per protocol' cohort who do not have two serological samples, collected at the beginning and end of the season according to protocol or who are missing vaccination info or were vaccinated during the study, impute the serological influenza outcome as LCI-S. Missing LCI-S status will be imputed using standard multiple imputation techniques, creating multiple imputed datasets with no missing values for each analysis.

Step 3: Impute the LCI-S outcome for the ITT cohort

Some members of the ITT cohort did not complete all weeks of the study and may be missing a serological outcome for the same reasons mentioned above. For these individuals, the serological influenza outcome must be imputed. Missing LCI-S status will be imputed using standard multiple imputation techniques, creating multiple imputed datasets with no missing values for each analysis.

2.d Model and variable selection

This data is from a cluster-randomized clinical trial. The investigators anticipate that the constrained randomization will ensure balance across important covariates. The clinics were pair-matched by the following characteristics:

- Study site

- Clinic size

- Clinic type (ED/Urgent care, Primary Care, Outpatient, Enhanced)

- Enhanced PPE (whether HCWs wore enhanced PPE during patient procedures, e.g. in dental and dialysis clinics)

- Patient population (Pediatric, Adult, or mixed)

Because these variables were matched on, the investigators will not adjust for any of them in the multivariable regression models. However, cluster-level random intercepts as well as additional participant-level covariates will be added to the model to adjust for possible residual confounding that is not controlled for by the cluster-randomized design. These covariates will be individual-level variables including:

- Age,

- Gender,

- Race (White, Black or African American, Asian, Native Hawaiian or Pacific Islander, American Indian or Alaskan Native) and ethnicity (Hispanic or Latino) [4]

- Number of household members under 5 (this has been noted as a strong risk factor for influenza [5]),

- Categorical occupation risk level (low, medium, or high),

- Binary season-specific flu vaccination status (was or was not vaccinated),

- Proportion of daily surveys where an individual reported exposure to someone with respiratory symptoms, and

- Individual-level (self-reported) measures of mask and hand hygiene compliance.

The investigators will attempt to include all of the above-listed variables in the analysis. No variable selection will be performed to optimize the goodness of fit of the model [6]. No Type I error rate adjustments will be made. Variables will be left out only if they contribute to instability in model estimation: e.g. collinearity (identified by variance inflation factors) or insufficient data to impute covariate status. In the model design stage, the investigators identified a full set of covariates that would satisfy the sample size recommendation [7] that the investigators have no more than $m/15$ parameters in our model, where $m = \min(n_1, n_2)$ and n_1 and n_2 are the numbers in each of the response variable categories. Based on preliminary estimates of the total number of influenza outcomes expected, the investigators aimed to keep the number of estimated parameters below 25.

The following variables were considered but not included in the analysis for the final model. Justification is provided.

Follow-up variables such as contact with household members with flu: noisy, lacking flu confirmation, and too reliant on self-reporting biases.

Cumulative study-based vaccination status (i.e. ever vaccinated, never vaccinated): would be collinear with seasonal vaccine status.

Absence from work: not directly related to outcome, chose to include average number of hours worked instead.

Dummy variables of clinic types: while these encode important questions, they aren't the main purpose of the central study, and were characteristics that were matched on.

Size of household: for parsimony, the investigators will include number of household members under 5 instead.

Clinic size: was used in matching for randomization.

Comorbid conditions: hard to justify including some and not others, of secondary relevance to the main outcome.

Average number of hours worked per week defined each season for each individual: there was a minimum number of hours worked defined in inclusion criteria, so this range will not be substantial.

Smoking status: secondary relevance to main outcome.

2.e Pre-specified exploratory analyses

In addition to the pre-specified analyses of primary and secondary outcomes, the investigators will run several pre-specified exploratory analyses to assess the impact of vaccine coverage and protocol compliance with the study outcomes.

Using the models described in Sections 3 and 4 below, the investigators will consider adding additional covariates to the models from the primary and secondary analyses. Specifically, the investigators will examine the impact of covariates specific to a particular cluster-season including:

Vaccine coverage among participants in the cluster

Hand-hygiene compliance rate

Measure of how often any HCW in the clinic wore any mask, MM or N95

Proportion of clinic HCW enrolled in study and size of clinic

Additionally, the investigators will assess interaction terms considering the following variables:

Interaction of cluster-level mask compliance with mask group

Interaction of individual-level vaccination status with mask group

Finally, the investigators will investigate combinations of cluster-level, seasonal, individual-level and cluster-seasonal random effects to capture different possible correlation structures of the data. The magnitude of each variance component will dictate whether they are included in the final model.

3. Analysis plan for primary outcome: laboratory-confirmed influenza

3.a Outcome definitions

A dichotomous variable will indicate whether or not a participant had an episode of laboratory-confirmed influenza during a single influenza season. As specified in the protocol, individuals who have a PCR-confirmed influenza infection collected within seven days of symptom onset or who have a 4-fold rise in antibody titer will be considered as a positive case. As described above, the investigators will implement a per-protocol analysis and an ITT analysis.

3.b Planned descriptive analysis

The descriptive analysis will focus on aggregated participant numbers across the groups specified in "respect outcome tables.xlsx" (January, 2016, revised April 2016). The tables are as follows: 1) demographics, comprised of a breakout across treatment arms of characteristics including age, race, gender, occupation, clinic characteristics, vaccination status, and comorbid conditions, 2) Adjudication, where tallies of ResPECT participants are broken down into categories depending on their eligibility for the ITT and PP analyses and influenza adjudication outcome by year, 3) Nasopharyngeal swab lab results, where participants are broken out by year and mask type across the possible influenza and non-influenza viruses tested during the study, and 4) Summary results of lab-confirmed influenza, lab-confirmed non-influenza, ARI, LCRI, LDRI and ILI across intervention arms only.

3.c Planned Primary Analysis

The investigators will use an individual-level logistic regression model to estimate the difference in influenza infection between the N95 and medical mask groups. Let Y_{ijs} be an indicator of whether subject i in cluster j developed laboratory-confirmed influenza in season s , and $MASK_{js}$ is an indicator of which mask the clinic was assigned to in season s (0 if medical mask and 1 if N95). Then the investigators will fit a version of this model

$$\text{logit}[\Pr(Y_{ijs}=1 | MASK_{js})] = \beta_0 + \beta_1 * MASK_{js} + \sum_k (\theta_k * X_{k,ijs}) + \alpha_j + \alpha_i$$

where the α_j are the cluster-level random intercepts, the α_i are the individual-level random intercepts (both assumed to be normally distributed), and the X_{k} refer to the individual-level covariates listed in Section 2.d. Unadjusted analyses will drop individual-level covariates and random intercepts, but will retain the cluster-level random effects.

For each fitted model, the estimated odds ratio comparing the odds of infection for those HCPs wearing N95s compared to those HCPs wearing medical masks (i.e. $\exp(\beta_1)$) will be reported, with a 95% CI.

Our ITT and PP will use the same model equation (shown above) but will use different subsets of participants from the full cohort as described above.

3.d Planned Sensitivity Analysis

To account for the unavoidable additional uncertainty regarding the missing data from our primary outcome, the investigators will conduct a sensitivity analysis that randomly assigns binary outcomes to participants who did not complete the study. Specifically, the investigators will create a two-dimensional

grid on which the investigators vary the influenza attack rates in participants who dropped out of the study for both the medical mask (MM) and N95 arm, separately. The investigators will fix the MM dropout attack rate between half and twice the observed MM attack rate, based on complete data. The investigators will fix the N95 dropout attack rate between half and twice the observed N95 attack rate, based on complete data. By varying these two parameters across the grid, and for each combination, calculating the adjusted odds ratio (averaged across n=50 imputed datasets for each point on the grid), the investigators will observe the sensitivity of our results to values of the missing data.

Additionally, the investigators will compare rates reporting of symptomatic events in the two study arms. If the investigators detect a statistically significant difference in symptomatic reporting between arms, the investigators will include a covariate adjustment of person time in each model to account for the amount of person time under observation.

4. Analysis plan for secondary outcomes

4.a Definitions of secondary outcomes:

Acute Respiratory Illness (ARI): This outcome is the incidence of ARI as a clinical syndrome. ARI will be defined as the occurrence of signs or symptoms of respiratory infection, as defined by Table 2 in the published protocol [8] with or without laboratory confirmation.

Influenza-Like Illness (ILI): This outcome is the incidence of ILI as a clinical syndrome. ILI will be defined as temperature of 100°F [37.8°C] or greater plus cough and/or a sore throat, with or without laboratory confirmation.

Laboratory Confirmed Respiratory Illness (LCRI): This outcome is defined as a laboratory confirmed respiratory illness from any of the pathogens listed in Table 4 in the protocol. Laboratory confirmed respiratory illness is ARI combined with laboratory confirmation by RT-PCR of infection with any of the pathogens listed in Table 4 in an upper respiratory specimen swab after symptoms were reported and within seven days of the original symptomatic report (PP definition of LCRI and confirmed April, 2016; [8]). Events with multiple viruses detected will count as a single event of LCRI (April 2016). If a swab that tested positive but was not associated with a symptomatic event (i.e. was not collected between symptom onset and seven days after symptom onset) then the incident does not count as a LCRI event. If an individual seroconverts to influenza, had symptoms at some time during the study, and does not have a PCR-confirmed pathogen event already, then the investigators will assign them a single LCRI event (May 2016).

Laboratory-detected respiratory infection (LDRI): For a participant with or without symptoms, a laboratory-detected infection is defined as: 1) detection of a respiratory pathogen by PCR or other laboratory methods or 2) serological evidence of infection (e.g., seroconversion) with a respiratory pathogen during the study surveillance period(s). In a case where two or more pathogens are identified in the same specimen, each pathogen will be considered to represent a separate infection (e.g., 2 pathogens as 2 events, 3 pathogens as 3 events) for that study participant for that time-point. Sequential detection of the same pathogens by PCR or other laboratory method in swabs collected at least 21 days apart will be considered separate infections.

For all of these endpoints, an individual may experience any or all of the outcomes more than once during the course of the 12-week study. Within the same study ID, participants must report being symptom-free for at least seven days prior to the beginning of the second event (May, 2016), except for LDRI which has the longer 21-day window separating events. As in the primary endpoint section, the secondary outcomes analysis will also include a per-protocol and an ITT analysis. A general description of these approaches is provided above, with specific modifications discussed below.

4.b Planned secondary outcome ITT analysis

As in the primary outcome ITT, this analysis will include all of the randomized ResPECT participants regardless of withdrawal status, participation, or protocol adherence. Secondary outcomes will be characterized using a per-week rate of infection so that all participants may be included. The investigators will use a covariate-adjusted individual-level log-linear Poisson regression analysis with person time as an offset term as well as cluster-level and individual-level random intercepts. For the ITT analysis, the amount of person time will be fixed at 12 weeks for each participant, regardless of how much time they participated in the study. The investigators will include the same covariates as described in the primary outcome analysis section above in the Poisson regression model for the ITT and per-protocol analyses. Unadjusted models will include only the cluster-level random intercepts.

For each fitted model, the estimated incidence rate ratio between the N95 and medical mask arm will be estimated and reported, with a 95% CI.

4.c Secondary outcome per-protocol analysis

Per-protocol analyses will use the same Poisson regression methods described for the secondary outcome ITT analyses. Additionally, the per-protocol analyses will include ResPECT study participants who completed at least 8 weeks (starting at the time of site activation) of the 12-week trial. All randomized participants will be included unless they withdrew, were administratively withdrawn, or deactivated before participating for at least 8 weeks.

Calculation of person-weeks for each participant will proceed as follows: for individuals who withdrew, completion date will be determined by the earliest withdrawal or deactivation date; in the event that these dates conflict, the earlier date will be used. For all other participants, active participation time will be calculated as the time between clinic activation and the latest of either the automatically-generated timestamp of the last completed daily or weekly survey or the collection date of the last swab, up to 12 weeks.

4.d Missing covariate data for secondary outcomes

The analysis approaches for our secondary outcomes will encounter instances of missing data, either in NP swab results or failure to report relevant information on self-reported forms. Areas in which these issues may require special handling are 1) missing swab collection dates, 2) missing swab results, and 3) incomplete symptomatic event reporting.

Missing swab collection dates are relevant for matching swab results to symptomatic event reports. Where this data is missing (often in the case of swabs collected using take-home kits, where participants self-collected the NP samples), the investigators will attempt to match swab results to symptomatic reporting events using the swab number or process of elimination (ie, only one event was reported and only 1 symptomatic swab was provided).

Missing swab results may occur due to practical considerations (running out of PCR plates), participant noncompliance, or handling errors. These results are truly missing, cannot be recovered, and therefore must be discarded. There are also a few instances (<30 out of >11,000, or <0.27%) in which results cannot be reliably matched to the correct individual due to barcode transcription errors. These will be discarded if there is any doubt about the correct assignment barcode. Since these errors did not arise in a systematic way and comprise a very small portion of the overall available and reliable swab samples, this decision should not affect the analysis outcome.

A few instances also exist in which participants provided a symptomatic swab but failed to complete a symptomatic event form. Since the participant provided no details to accompany the biological specimen, the investigators will not include these data in the analysis of ILI events (which require specific symptom reports). However, positive symptomatic swab data lacking specific symptom data will be included in the ARI and LCRI.

References

Collection of Race and Ethnicity Data in Clinical Trials.
<http://www.fda.gov/RegulatoryInformation/Guidances/ucm126340.htm> = archived; no longer available online. → U.S. Department of Health and Human Services: Food and Drug Administration (2005). Guidance for Industry: Collection of Race and Ethnicity Data in Clinical Trials. Rockville, MD: Food and Drug Administration.

Fisher LD, Dixon DO, Herson J, et al. (1990). Intention to treat in clinical trials. In: Peace KE, editor. Statistical issues in drug research and development. New York: Marcel Dekker. pp. 331–50. → Fisher LD, Dixon DO, Herson J, et al. (1990) Intention to treat in clinical trials. In: Peace KE (Ed.), Statistical issues in drug research and development (pp. 331-50). New York, NY: Marcel Dekker.

Hayes RJ, Moulton LH (2009) Cluster Randomised Trials. ISBN: 978–1-58488-816-1; 315 pages; Chapman and Hall/CRC. → Hayes RJ and Moulton LH (2009) Cluster Randomised Trials. Chapman and Hall/CRC.

ICH Expert Working Group (1998). ICH Harmonised Tripartite Guideline, Statistical Principles for Clinical Trials E9, Current Step 4 version dated 5 February 1998.
http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E9/Step4/E9_Guideline.pdf (last accessed 04 June 2017). → same as highlighted

PMID: 8134732 J Whitehead, Sample size calculations for ordered categorical data. Stat Med, 12:2257-2271, 1993.

PMID: 27255755 Radonovich LJ Jr, Bessesen MT, Cummings DA, et al. BMC Infect Dis; 2016(2)16:243.

PMID: 10347857 Schafer, Joseph L. "Multiple imputation: a primer." Statistical Methods in Medical Research 8.1 (1999): pp. 3-15.

PMID: 20067628 Williams CJ, Schweiger B, Diner G, et al. Seasonal influenza risk in hospital healthcare workers is more strongly associated with household than occupational exposures: results from a prospective cohort study in Berlin, Germany, 2006/07. BMC Infect Dis. 2010 Jan 12;10:8.