When to intubate in hypoxemic respiratory failure? An international survey of clinicians: Analysis Plan

2024-01-09

**Christopher Yarnell and colleagues**

Contents

[1 Project overview 3](#_Toc155604954)

[1.1 Background 3](#_Toc155604955)

[1.2 Objective 3](#_Toc155604956)

[1.3 Methods 3](#_Toc155604957)

[1.4 Relevance 3](#_Toc155604958)

[2 Analysis plan overview 4](#_Toc155604959)

[2.1 Pre-registration 4](#_Toc155604960)

[3 Data 4](#_Toc155604961)

[3.1 Survey variables 4](#_Toc155604962)

[3.1.1 Table 1: Survey variables 5](#_Toc155604963)

[3.2 Questionnaire formatting and composition 5](#_Toc155604964)

[4 Descriptive analyses 6](#_Toc155604965)

[4.1 Recruitment approach 6](#_Toc155604966)

[4.2 Respondent characteristics 7](#_Toc155604967)

[4.3 Response characteristics 7](#_Toc155604968)

[4.4 Table 1 7](#_Toc155604969)

[5 Primary analysis: Bayesian multilevel proportional odds model 7](#_Toc155604970)

[5.1 Data processing 8](#_Toc155604971)

[5.2 Predictors 9](#_Toc155604972)

[5.3 Clustering 9](#_Toc155604973)

[5.4 Outcome 9](#_Toc155604974)

[5.5 Prior distributions 9](#_Toc155604975)

[5.6 Outputs 10](#_Toc155604976)

[6 Secondary analyses 10](#_Toc155604977)

[6.1 Comparing with observed data 10](#_Toc155604978)

[6.2 Requests for more information 10](#_Toc155604979)

[7 Sensitivity analyses 11](#_Toc155604980)

[8 Computation and reporting 11](#_Toc155604981)

[9 Additional analyses 11](#_Toc155604982)

[10 Survey screenshots 12](#_Toc155604983)

[10.1 Preamble 12](#_Toc155604984)

[10.2 Footnote 12](#_Toc155604985)

[10.3 Demographics page 13](#_Toc155604986)

[10.4 Clinical Question page 14](#_Toc155604987)

[11 References 15](#_Toc155604988)

# Project overview

## Background

The decision for invasive ventilation in patients with hypoxemic respiratory failure is influenced by many factors including the underlying diagnosis, work of breathing, oxygenation, level of consciousness, blood pressure, and the trajectory of abnormalities.1–3 However, the specific degrees and patterns of derangement in physiologic states that prompt invasive ventilation are unknown.4

## Objective

Survey clinicians involved in the decision for initiating invasive ventilation to quantify the influence of different clinical factors on the decision to intubate and provide invasive ventilation.

Participants: Self-identified clinicians who are involved in intubation decisions as part of their practice, including physicians (attending and trainee), respiratory therapists, nurses, and others.

## Methods

We will broadly distribute an online survey through specialty societies, academic groups, and personal email. The survey will be anonymous and participation will imply consent. We will gather basic demographic information including the specialty, role, country, and duration of clinical practice. Then we will present the respondents with several scenarios. For each scenario we will ask whether they recommend intubation and invasive ventilation at that time. We will also ask if any further information would be helpful. Respondents will see different scenarios, with each of the survey variables being randomly permuted between two or more values. The survey will be analyzed using parametric multilevel Bayesian models to investigate intubation recommendations across physiologic scenarios.

## Relevance

This survey will help describe the relationship between clinical variables and the decision to initiate invasive ventilation in hypoxemic respiratory failure.

# Analysis plan overview

This document records the analysis plan for this project to facilitate pre-registration and provide transparency regarding our analytic plan and objectives. We proceed by describing the Descriptive Analyses, the Primary Analysis, Secondary Analyses, and Sensitivity Analyses. A draft plan was circulated to collaborators on December 8, 2023. We also held a virtual meeting open to all collaborators on January 5, 2024. This plan includes feedback sent by email and from that meeting.

## Pre-registration

The survey opened on September 27, 2023 and will close on January 10, 2024. This analysis plan was finalized prior to the locking and analysis of the final database. However, we did inspect some aspects of the data prior to forming the final analysis plan. This mainly included inspecting the demographic characteristics of respondents in order to ensure that the Shiny/REDCap survey interface was appropriately capturing responses, and to assess survey uptake in different geographic regions. We did also inspect preliminary scenario and response data to ensure that the survey scenario randomization was working appropriately, and that the responses were being coded appropriately across multiple languages.

The analysis plan itself is largely unchanged from the initially proposed analysis plan in May 2023, where we envisioned a Bayesian multilevel proportional odds model using both respondent and scenario data as predictors, with clustering by individual to account for repeated measures. However, the “pre-registration” of our plan would have been more robust if we had registered the analysis plan prior to opening the survey. This is a limitation of our work.

# Data

## Survey variables

We based initial drafts of the survey on a prior systematic review of the criteria for invasive ventilation used in randomized trials1, observational studies of physiologic thresholds for invasive ventilation,2,3 and our clinical experience in caring for patients with respiratory failure. We identified several key physiologic variables that likely influence the decision for invasive ventilation. For each variable, we settled on a range of potential values (Table 1). The initial set of variables was reduced and adjusted based on preliminary feedback from coauthors and clinical colleagues.

### Table 1: Survey variables

|  |  |
| --- | --- |
| **Variable** | **Possible values** |
| Diagnosis | 1. Community-acquired pneumonia 2. COVID pneumonia 3. Influenza pneumonia 4. Pancreatitis 5. Sepsis |
| Age | 20 to 70 years |
| Frailty | 1. Independent and fit (CFS 1-2) 2. Independent with well-controlled medical problems (CFS 3) 3. Assistance for shopping and heavy housework (CFS 5) |
| Peripheral oxygen saturation | 85% to 97% |
| Inspired oxygen fraction | 0.4 to 1.0 in increments of 0.1 |
| Oxygen device | 1. High-flow nasal oxygen 2. Non-invasive ventilation |
| Respiratory rate | 20 to 40 breaths per minute |
| Work of breathing | 1. No use of neck muscles, no abdominal paradox 2. Use of neck muscles, no abdominal paradox 3. Use of neck muscles, abdominal paradox |
| Norepinephrine use | 1. No 2. Yes |
| Level of consciousness | 1. Alert and obeying 2. Drowsy but obeying 3. Drowsy and not obeying |
| Duration of abnormalities | 1. 10 minutes 2. 30 minutes 3. 1 hour 4. 2 hours 5. 4 hours |

Caption: This table shows the survey variables and the possible values each variable can take. Not all combinations of values across variables are common or permitted (for example, when age is less than 40, frailty is set to one of “Independent” options). COPD = chronic obstructive pulmonary disease, CHF = congestive heart failure, COVID = coronavirus-19, CFS = clinical frailty scale

## Questionnaire formatting and composition

The decision to initiate invasive ventilation is influenced by many factors, and qualitative research shows that the interactions between these factors are important.4 For this reason, we chose a survey design that permitted evaluation of all possible interactions between relevant variables. We used the same stem (“Would you recommend intubation and invasive ventilation?”) and response options for every question, varying only the scenario variables displayed between the stem and the response options.

We displayed the variables and values for that particular scenario in a table. These were randomly selected from the potential values for each question, meaning that every participant sees different scenarios in a different order. All code for the stochastic stem generation is available in a public repository. Every scenario has the same format.

Beneath the scenario, we displayed a slider used to answer the question. Respondents select one of five options (“Definite no”, “Probable no”, “Uncertain”, “Probable yes”, “Definite yes”) using the slider. After the slider, there is one further optional question “Optional: any additional information that would help you decide?” When respondents press “Next”, their responses are logged (in an anonymous fashion) in a secure online REDCap database housed at the University of Toronto.

We have included screenshots of the survey in an appendix.

This survey format results in a large number of responses to the same overall question with permutation of the observed values of important clinical variables. With sufficient sample size and an appropriate analytic plan, this will allow for full investigation of the interactions between clinical variables. The survey was coded in R using *shiny*.5,6 All code will be made available in a publicly accessible repository.

# Descriptive analyses

## Recruitment approach

We will describe our recruitment approach in detail, including the months in which notifications were sent out from different societies or groups to potential respondents. Our recruitment methodology favoured maximizing the number and diversity of respondents over careful tracking of the denominator in the response rate, so we will not be able to report our response rate.

## Respondent characteristics

We will report the total number of respondents. We will also report the number and percentage for each specialty, role, region (or sub-region, depending on numbers), and language used to respond to the survey. Country names, codes, sub-regions, and regions will be defined using the ISO 3166 list of country codes (<https://www.iso.org/iso-3166-country-codes.html>) which itself uses information from the United Nations Division of Statistics. We will also report the median and interquartile range of duration in practice.

## Response characteristics

We will report the total number of responses, responses per respondent, and the distribution of responses by respondent characteristics. We will also report the number of respondents with 10 responses (the maximum).

The reason to report the number of respondents with 10 responses is that at times of high survey traffic, some users were disconnected from the survey prior to completing the survey. The survey itself saved responses as they were entered. Therefore, a respondent who answered only two scenarios and then was disconnected may have attempted to complete the survey again. However, we think it is unlikely that someone who completed ten scenarios and received the “Thank-you” completion message would then access the survey again. This means that the subset of respondents who answered ten scenarios is highly likely to be composed of unique respondents.

In addition to reporting the response data with respect to the main question, we will also report the number and percentage of respondents requesting additional information and the number and percentage who filled in responses in the free-text field for additional information.

## Table 1

Rows: Region / sub region, role, specialty, duration in practice (categorized))

Columns: Number of respondents, number of responses, responses per respondent, response (5-level), additional information requested

# Primary analysis: Bayesian multilevel proportional odds model

To analyze the ordinal intubation recommendation outcome, we will use a Bayesian multilevel proportional odds model.

*Why Bayesian?* Because Bayesian prior distributions allow us to encode skepticism that any one factor will be deterministic of the recommendation, and facilitate the multilevel structure of the model. Results from a Bayesian analysis can be framed in terms of probability which can be more intuitive for clinicians than a frequentist null-hypothesis-testing formulation.

*Why proportional odds model?* A proportional odds model uses all of the information contained in an ordinal outcome (it does not combine categories) but it assumes the same odds govern every transition from a lower to a higher outcome level (the proportional odds assumption). The benefit of this model is that it provides a single measure (odds ratio) of the influence of each parameter on whether intubation is recommended. Even if the proportional odds assumption does not hold, the practical implications of violating this assumption are usually negligible.7,8

*Why multilevel?* Two reasons. First, the same respondent may (should) answer multiple scenarios. Clustering at the individual level is a standard approach when data has this structure, which is known as a “repeated measures” structure.9 Second, multilevel clustering is helpful for variables with many categories, such as country, because it allows for the model to quantitatively find the best balance between every country having a unique intercept and every country being the same, based on how many responses there were from that country. For countries with few respondents, the random intercept will likely be close to the mean and have a large uncertainty. For countries with many responses, the opposite will be true.

## Data processing

We will categorize continuous variables in order to allow for non-linearity while also maintaining interpretable model results. We will categorize variables as follows:

* 1. Respondent: specialty, role, duration in practice (0-5 years, 6-10 years, 11-20 years, more than 20 years)
  2. Patient baseline information: age (20-40, 40-50, 50-60, 60-70), frailty, diagnosis
  3. Patient clinical status: peripheral oxygen saturation (< 89, 89-91, 92-94, 95-97), inspired oxygen fraction (categorical), oxygen device, respiratory rate (<21,21-24, 25-27, 28-32, >32), breathing pattern, norepinephrine use, level of consciousness, duration in current state

## Predictors

We plan to use all respondent and scenario data as predictors, and to include interactions between all scenario variables.

*Why include so many interactions?* Because clinical experience suggests that interactions may be important, and we will likely have a large number of data points to statistically power the analysis for interactions.However, introducing interactions dramatically increases the number of coefficients (from 47 parameters to 501 parameters!) so we will choose a prior distribution that guards against overfitting (see prior distribution section).

## Clustering

We will cluster responses at the level of respondent, country, and region. We will use the cluster variances to compute median odds ratios and compare variability at each level.

## Outcome

The outcome will be the ordinal intubation recommendation: “Definite no”, “Probable no”, “Uncertain”, “Probable yes”, “Definite yes.”

## Prior distributions

For all fixed effects (including interactions and intercept) we will use a horseshoe prior, degrees of freedom = 3, ratio of non-zero parameters = 0.15. (4 intercept parameters, 43 main effect parameters, 454 interaction parameters, totals 501 parameters; 0.15 ratio of non-zero to zero parameters means a prior belief that there are ~ 65 non-zero parameters)10,11.

*Why use a horseshoe prior distribution?* It encodes our prior knowledge as “we think at most X% of these coefficients will have odds ratios different from 0”, and we can choose X (we set it at 15%). This prevents model overfitting. On one hand, it is true that all of the variables from the survey are important in the clinical decision we are considering. On the other hand, it is not clear that the **interactions** are important. We hypothesize that the model will pull out the main effects and find very few interactions to have odds ratios different from zero.

The random effects will be modeled as normal distributions in the lop-odds space with mean 0. The standard deviation of the random effects will have a prior half-normal distribution with standard deviation 0.512,13.

## Outputs

From the primary analysis, we will report odds ratios for all fixed effect coefficients (forest plot), median odds ratios for random effects,14 and a geographic forest plot (map + forest plot) for the country-level random intercepts.

# Secondary analyses

## Comparing with observed data

The goal of this secondary analysis is to compare survey responses to observed data. For this secondary analysis, we will use the MIMIC-IV cohort from the study “Do Thresholds for Invasive Ventilation in Hypoxemic Respiratory Failure Exist?” (PMID 36150166) 2. This is a cohort of patients who are in the intensive care unit receiving oxygen by high-flow nasal cannula, non-invasive ventilation, or non-rebreather mask.

For each of the following thresholds, we will identify whether each patient ever met that threshold while receiving oxygen via non-invasive ventilation or high-flow nasal cannula. If a patient ever met that threshold, then we will record the scenario covariates from the survey as observed in the database at the time of first meeting that threshold while on high-flow nasal cannula or non-invasive ventilation. We will then input those coefficients into the primary analysis model to output a distribution of intubation recommendation responses. We will combine all the response distributions across all patients observed to have met the threshold and plot the results graphically.

For thresholds, we will use a threshold similar to the FLORALI randomized controlled trial threshold, and thresholds of SF < 90, SF < 110, ROX < 4, and ROX < 5.15 The FLORALI threshold will require one of either hemodynamic, neurologic, or respiratory compromise, defined as follows: hemodynamic compromise – use of vasopressors. Neurologic compromise – GCS < 12. Respiratory compromise – two of: respiratory rate 40 or more, pH < 7.35, SpO2 < 90 on FiO2 0.80 or higher, or lack of improvement in work of breathing.

## Requests for more information

Respondents had the opportunity to indicate if they would find further information helpful. There were five prespecified types of information: chest X-ray, blood gas, NIV parameters (if patient was on NIV), esophageal manometry, and more observation time. We will perform a Bayesian multilevel logistic regression with the same structure as the primary analysis for each binary outcome of whether or not that type of information was requested.

# Sensitivity analyses

We will perform two sensitivity analyses using only respondents who answered 10 scenarios (because these are very unlikely to include duplicate respondents). We will repeat the primary analysis and repeat the secondary analysis focused on additional information.

We will also perform a sensitivity analysis where we analyze data from each scenario diagnosis separately. This is equivalent to enforcing an interaction between diagnosis and every other variable in the model. We will inspect the coefficients for similarity or differences across the five models that result.

# Computation and reporting

We will use R and the brms package for all Bayesian analyses.5,16 We will summarize posterior distributions with 95% credible intervals. We will report both effect estimates and, where helpful for understanding results, the probability of an odds ratio being greater than or less than 1. Where a region of practical equivalence is required, we will use the region from an odds ratio of 0.9 to odds ratio of 1/0.9 = 1.11.

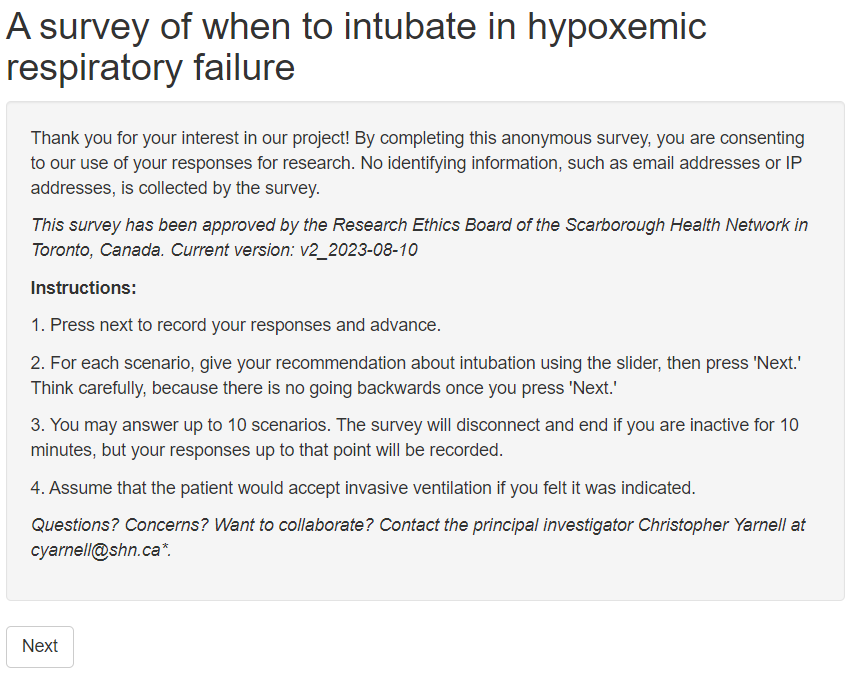
# Additional analyses

There are additional potential analyses from the data in this manuscript. These analyses may appear in the supplement of the main manuscript, or they may comprise secondary manuscripts / letters, depending on the feasibility of the analyses and the volume of the findings:

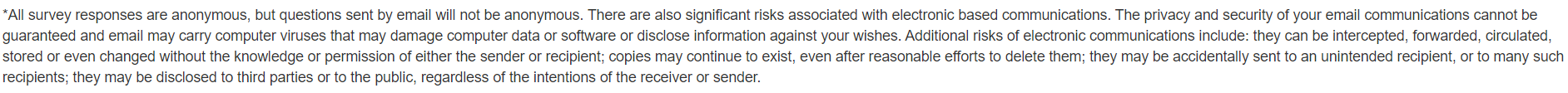
* Qualitative analysis of free-text responses
* Descriptive analysis and primary model restricted to respiratory therapist respondents
* Descriptive analysis and primary model restricted to nurse respondents

# Survey screenshots

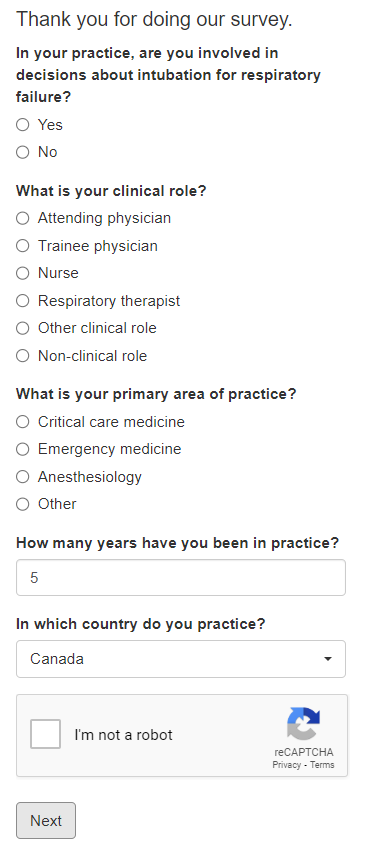
## Preamble



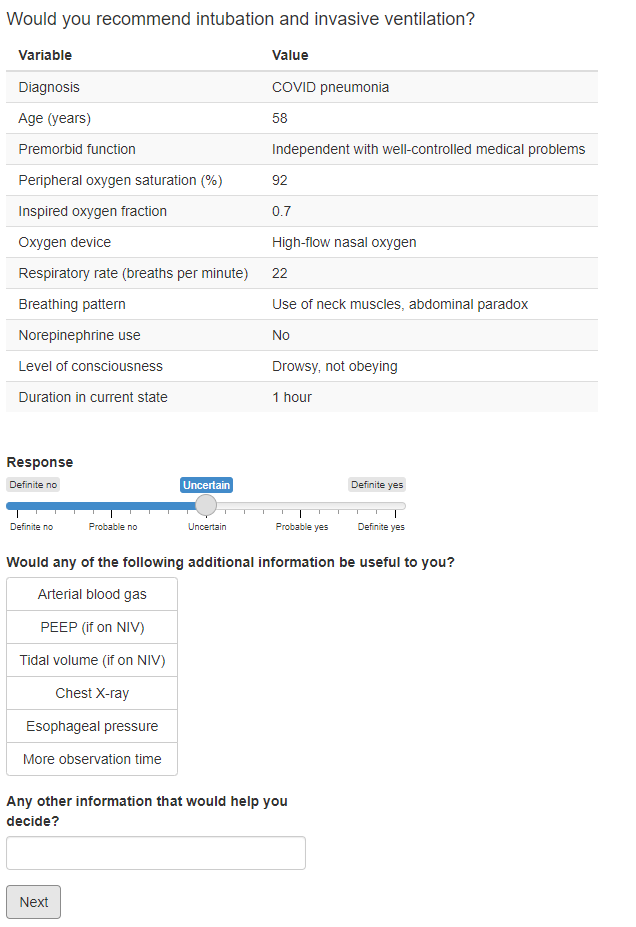
## Footnote



## Demographics page



## Clinical Question page



# References

1. Hakim R, Watanabe-Tejada L, Sukhal S, Tulaimat A. Acute respiratory failure in randomized trials of noninvasive respiratory support: A systematic review of definitions, patient characteristics, and criteria for intubation. *Journal of Critical Care*. 2020;57:141-147. doi:10.1016/j.jcrc.2020.02.018

2. Yarnell CJ, Johnson A, Dam T, Jonkman A, Liu K, Wunsch H, Brochard L, Celi LA, De Grooth HJ, Elbers P, Mehta S, Munshi L, Fowler RA, Sung L, Tomlinson G. Do Thresholds for Invasive Ventilation in Hypoxemic Respiratory Failure exist? A Cohort Study. *Am J Respir Crit Care Med*. 2023;207(3):271-282. doi:10.1164/rccm.202206-1092OC

3. Yarnell CJ, Angriman F, Ferreyro BL, Liu K, De Grooth HJ, Burry L, Munshi L, Mehta S, Celi L, Elbers P, Thoral P, Brochard L, Wunsch H, Fowler RA, Sung L, Tomlinson G. Oxygenation thresholds for invasive ventilation in hypoxemic respiratory failure: a target trial emulation in two cohorts. *Critical Care*. 2023;27(1):67. doi:10.1186/s13054-023-04307-x

4. Bauer PR, Kumbamu A, Wilson ME, Pannu JK, Egginton JS, Kashyap R, Gajic O. Timing of Intubation in Acute Respiratory Failure Associated With Sepsis: A Mixed Methods Study. *Mayo Clinic Proceedings*. 2017;92(10):1502-1510. doi:10.1016/j.mayocp.2017.07.001

5. R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing; 2020. https://www.r-project.org

6. Kaufman AR. Implementing novel, flexible, and powerful survey designs in R Shiny. *PLoS One*. 2020;15(4):e0232424. doi:10.1371/journal.pone.0232424

7. Harrell FEJr. *Regression Modeling Strategies - With Applications to Linear Models, Logistic Regressions, and Survival Analysis*. Vol 26. Springer New York; 2001.

8. Harrell F. Statistical Thinking - Assessing the Proportional Odds Assumption and Its Impact. Published March 9, 2022. Accessed January 8, 2024. https://www.fharrell.com/post/impactpo/

9. Fitzmaurice GM, Laird NM, Ware JH. *Applied Longitudinal Analysis*. John Wiley & Sons; 2012.

10. Piironen J, Vehtari A. Sparsity information and regularization in the horseshoe and other shrinkage priors. *Electronic Journal of Statistics*. 2017;11(2):5018-5051. doi:10.1214/17-EJS1337SI

11. Carvalho CM, Polson NG, Scott JG. Handling Sparsity via the Horseshoe. In: *Proceedings of the Twelth International Conference on Artificial Intelligence and Statistics*. PMLR; 2009:73-80. Accessed December 7, 2023. https://proceedings.mlr.press/v5/carvalho09a.html

12. Gelman A, Carlin JB, Stern HS, Dunson DB, Vehtari A, Rubin DB. *Bayesian Data Analysis, Third Edition*. CRC Press; 2013.

13. Gelman A. Prior distributions for variance parameters in hierarchical models (comment on article by Browne and Draper). *Bayesian Analysis*. 2006;1(3):515-534. doi:10.1214/06-BA117A

14. Austin PC, Merlo J. Intermediate and advanced topics in multilevel logistic regression analysis. *Statistics in Medicine*. 2017;36(20):3257-3277. doi:10.1002/sim.7336

15. Frat JP, Thille AW, Mercat A, Girault C, Ragot S, Perbet S, Prat G, Boulain T, Morawiec E, Cottereau A, Devaquet J, Nseir S, Razazi K, Mira JP, Argaud L, Chakarian JC, Ricard JD, Wittebole X, Chevalier S, Herbland A, Fartoukh M, Constantin JM, Tonnelier JM, Pierrot M, Mathonnet A, Béduneau G, Delétage-Métreau C, Richard JCM, Brochard L, Robert R. High-flow oxygen through nasal cannula in acute hypoxemic respiratory failure. *New England Journal of Medicine*. 2015;372(23):2185-2196. doi:10.1056/NEJMoa1503326

16. Bürkner PC. brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software*. 2017;80(1):1-28. doi:10.18637/jss.v080.i01