An international factorial vignette-based survey of intubation decisions in acute hypoxemic respiratory failure

Yarnell, Christopher J, MD PhD; ORCID 0000-0001-5657-9398
Paranthaman, Arviy RRT;
Reardon, Peter MD; ORCID 0000-0003-2118-1229
Angriman, Federico MD MPH PhD; ORCID 0000-0003-0971-386X
Bassi, Thiago; ORCID 0000-0001-5570-5057
Bellani, Giacomo; ORCID 0000-0002-3089-205X
Brochard, Laurent;
De Grooth, Harm Jan; ORCID 0000-0002-7499-076X
Dragoi, Laura; ORCID 0000-0003-2201-8116
Goligher, Ewan C, MD PhD;
Lewis, Kimberley;
Li, Baoli;
Kareemi, Hashim; ORCID 0000-0001-8599-1576
Tirupakuzhi Vijayaraghavan, Bharath Kumar; ORCID 0000-0002-1801-0667
Mehta, Sangeeta; ORCID 0000-0002-7073-4769
Mellado-Artigas, Ricard; ORCID 0000-0003-2815-6819
Moore, Julie; ORCID 0000-0002-7786-4347
Morris, Idunn MBChB(Hons) FRCA FCICM;
Roman-Sarita, Georgiana RRT;
Pham, Tai MD, PhD; ORCID 0000-0002-4373-0711
Sereeyotin, Jariya; ORCID 0000-0002-2880-8242
Tomlinson, George;
Wozniak, Hannah MD; ORCID 0000-0001-5404-4621
Yoshida, Takeshi MD PhD; ORCID 0000-0002-4695-5377
Fowler, Rob;
for the Canadian Critical Care Trials Group

2024-04-16

Abstract: 267 words

Text: 3,004 words (including headers)

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.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

Affiliations

- Department of Critical Care Medicine, Scarborough Health Network, Toronto, Ontario, Canada (CJY, PR, AP)
- Scarborough Health Network Research Institute, Toronto, Ontario, Canada (CJY)
- Keenan Research Centre for Biomedical Research, Li Ka Shing Knowledge Institute, St Michael’s Hospital, Unity Health Toronto, Canada (LB)
- Interdepartmental Division of Critical Care Medicine, University of Toronto, Toronto, Canada (CJY, FA, RAF, LB, SM, HW, ECG, TB)
- Department of Medicine, Division of Respirology, University Health Network
- Department of Medicine, Sinai Health System, Toronto, Canada (SM)
- Department of Medicine, University of Toronto, Toronto, Canada (RAF)
- Department of Respiratory Therapy, University Health Network, Toronto, Canada (GRS)
- Department of Medicine, University Health Network, Toronto, Canada (GT)
- Institute for Clinical Evaluative Sciences, Toronto, Canada (RAF)
- Institute of Health Policy, Management and Evaluation, University of Toronto, Canada (RAF, GT)
- Sunnybrook Health Sciences Centre, Toronto, Canada (FA, RAF)
- Division of Critical Care, Department of Medicine, McMaster University, Hamilton, Canada (KL)
- Department of Health Research, Methods, Evidence, and Impact, McMaster University, Hamilton, Canada (KL)
• Service de Medecine Intensive-Réanimation, Assistance Publique–Hôpitaux de Paris, Hôpital de Bicêtre, DMU CORREVE, FHU SEPSIS, Groupe de Recherche CARMAS, Le Kremlin-Bicêtre, France (TP)

• Universite Paris-Saclay, Université de Versailles Saint-Quentin-en-Yvelines, Université Paris-Sud, Inserm U1018, Equipe d’Epidemiologie Respiratoire Intégrative, Centre d’Epidémiologie et de Santé des Populations, Villejuif, France (TP)

• Surgical ICU, Department of Anesthesiology, Hospital Clínic de Barcelona, Barcelona (Spain) [RMA]

• Lawrence S. Bloomberg Faculty of Nursing, University of Toronto, Toronto, Canada [JM]

• Trent-Fleming School of Nursing, Trent University, Peterborough, Ontario [JM]

• Department of Nursing, Mount Sinai Hospital, Toronto, Canada [JM, BL]

• Nepean Clinical School, University of Sydney, New South Wales, Australia (IM)

• Department of Intensive Care Medicine, Nepean Hospital, Sydney, Australia (IM)

• Department of Physiology, University of Toronto, Toronto, Canada (IM, ECG)

• Intensive Care Center, UMC Utrecht, Utrecht, The Netherlands (HJdG)

• Department of Anesthesiology, Pharmacology, Intensive Care and Emergency Medicine, Geneva University Hospitals, Geneva, Switzerland (HW)

• Centre for Medical Sciences (CISMed), University of Trento, 38122, Trento, Italy (GB)

• Anaesthesia and Intensive Care, Santa Chiara Hospital, APSS, Trento, Italy (GB)

• Department of Critical Care Medicine, Apollo Hospitals, Chennai, India and Centre for Inflammation Research, University of Edinburgh, U.K. (BKTV)
• Department of Anesthesiology, Division of Critical Care Medicine, King Chulalongkorn Memorial Hospital and Faculty of Medicine, Chulalongkorn University, Bangkok, Thailand (JS)

• Department of Emergency Medicine, University of British Columbia, Vancouver, Canada (HK)

• The Department of Anesthesiology and Intensive Care Medicine, Osaka University Graduate School of Medicine, Suita, Japan (TY)
Contributions

- Concept: CJY, RF, GT, KL, SM
- Design: CJY, RF, GT, KL, SM, PR, AP, IM, FA, HJdG
- Data analysis and interpretation: All
- Data acquisition: All
- Drafting, revising for important intellectual content, final approval, and agreement to be accountable: All

Corresponding author
Christopher Yarnell
Department of Critical Care Medicine, Scarborough Health Network
3050 Lawrence Ave E, Toronto, Ontario, Canada, M1P 2V5
cyarnell@shn.ca

Declarations of interest
Dr Brochard’s laboratory received grants from Medtronic, Stimit, Vitalaire, equipment from Philips, Sentec, Fisher Paykel, Cerebra Health. Dr Ricard Mellado-Artigas has received lecturing fees from Medtronic, MSD and Fisher and Paykel. Dr. Bellani receives lecturing fees from Draeger Medical and consultancy fees from Flowmeter.

Funding and role of funders in study
Survey dissemination costs were covered by an Interdepartmental Division of Critical Care Medicine Trainee Award to Dr Yarnell. Dr Fowler is the H. Barrie Fairley Professor of Critical Care at the University Health Network, Interdepartmental Division of Critical Care Medicine, University of Toronto. Funders had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; nor in the decision to submit the manuscript for publication. The opinions, results and conclusions reported in this paper are those of the authors and are independent from the funding sources. No endorsement by any of the funding agencies is intended or should be inferred.

Acknowledgements
We thank Dr Oleksa Rewa and Dr John Basmaji for comments on a draft of the manuscript. We thank the following for endorsement and/or dissemination: Australia New Zealand Intensive Care Society, Canadian Association of Critical Care Nurses, Canadian Association of Emergency Physicians, Canadian Critical Care Society, Canadian Society of Respiratory Therapists, the European Society of Intensive Care Medicine, Indian Registry of Intensive care and the Chennai branch of the Indian Society of Critical Care Medicine, Intensive Care Society (United Kingdom), Japanese Society of Respiratory Care, Pleural Pressure Working Group, Platform of Randomized Adaptive Clinical Trials in Critical Illness investigator group, Pragmatic Critical Care Research Group, REVA Network, Society for Critical Care Medicine, Societa’ Italiano di Anestesia.
Analgesia, Rianimazione e Terapia Intensiva (SIAARTI), la Société de Réanimation de Langue Française (SRLF), and the Thai Society of Critical Care Medicine (TSCCM).
Take-home message

In this international, interprofessional survey of 2,294 individuals from 74 countries on when to intubate patients with acute hypoxemic respiratory failure, low oxygen saturation, high inspired oxygen fraction, abnormal breathing patterns, and decreased level of consciousness were associated with recommending intubation. There was important variation among individuals and countries, suggesting an opportunity to improve care by studying the relationship between decision-making and outcomes.

Keywords

Acute Hypoxemic Respiratory Failure, Endotracheal Intubation, Healthcare Surveys, Multilevel Analyses
Abstract

**Purpose:** Intubation is a common procedure in acute hypoxemic respiratory failure (AHRF), with minimal evidence to guide decision-making. We conducted a survey of when to intubate patients with AHRF to measure the influence of clinical variables on intubation decision-making and quantify variability.

**Methods:** We developed an anonymous factorial vignette-based web survey to ask clinicians involved in the decision to intubate “Would you recommend intubation?” Respondents selected an ordinal recommendation from a 5-point scale ranging from “Definite no” to “Definite yes” for up to 10 randomly allocated vignettes. We disseminated the survey through clinical and academic societies, analyzed responses using Bayesian proportional odds modeling with clustering by individual, country, and region, and reported mean odds ratios (OR) with 95% credible intervals (CrI).

**Results:** Between September 2023 and January 2024, 2,294 respondents entered 17,235 vignette responses in 74 countries [most common: Canada (29%), USA (26%), France (9%), Japan (8%), and Thailand (5%)]. Respondents were attending physicians (63%), nurses (13%), trainee physicians (9%), respiratory therapists (9%), other (6%). Lower oxygen saturation, higher inspired oxygen fraction, non-invasive ventilation compared to high-flow, tachypnea, neck muscle use, abdominal paradox, drowsiness, and inability to obey were associated with increased odds of intubation; diagnosis, vasopressors, and duration of symptoms were not. Within a country the odds of recommending intubation changed between clinicians by an average factor of 2.60, while changing between countries within a region changed it by 1.56.

**Conclusion:** In this international, interprofessional survey of more than 2000 practicing clinicians, intubation for patients with AHRF was mostly decided based on oxygenation, breathing pattern, and consciousness, but there was important variation across individuals and countries.
Introduction

When to intubate people with acute hypoxemic respiratory failure (AHRF) is an important and common dilemma for clinicians, but optimal criteria to guide initiation are unknown(1). Uncertainty in this decision is important, because earlier intubation may unnecessarily expose patients to risks of periprocedural shock and cardiac arrest(2), laryngeal and dental trauma(3), infection(4), ventilator-induced lung injury(5), immobility and weakness(6), and delirium and post-traumatic stress disorder(7); later intubation may increase the risk of death due to physiologically dangerous intubating conditions(8), worsening of self-inflicted lung injury(9), or decompensated respiratory failure progressing to cardiac arrest(10).

In qualitative research and surveys, clinicians have highlighted the importance of physiologic variables such as degree of hypoxemia and breathing pattern in deciding when to intubate(11–13). However, there is variation in the use of invasive ventilation by location, time period, and patient race(14–18). No research yet investigates the relationships between physiological variables and intubation decisions, and the extent of variability in decision-making between individuals and around the world is unknown.

To address these uncertainties, we conducted an international, scenario-based electronic survey to quantify variability in practice and measure the influence of clinical variables on intubation decisions for patients with AHRF.

Methods

We designed and disseminated an anonymous electronic survey using established methods.(19,20) We obtained approval from the Scarborough Health Network research ethics
board (#ICU-23-009). Completion of the survey implied consent. We pre-registered the analysis plan prior to completion of data collection (21) and analytic code is publicly available. (cite doi)

Participants

We targeted a global population of clinicians who self-identified as “involved in the decision to initiate invasive ventilation for respiratory failure as part of their clinical practice.” Because we targeted a global population of diverse professions, we could not ascertain a denominator for calculating response rates. (22)

Development and piloting

We identified key physiologic variables hypothesized to influence intubation decisions using randomized trials (23), observational studies of thresholds for intubation (15, 24), research on mechanical ventilation weaning physiology (25–28), and our clinical experience with patients experiencing respiratory failure. We performed clinical sensibility testing with 8 clinicians who are involved in intubation decisions (2 respiratory therapists, 4 physicians, 2 nurses) and revised the survey based on that feedback. (19, 29)

Survey design

We designed a factorial vignette-based survey. (30–32) Variables were patient age, premorbid frailty (33), diagnosis, oxygen saturation, inspired oxygen fraction, oxygen device, respiratory rate, breathing pattern, norepinephrine use, level of consciousness, and duration in current state (eTable 1). Variables were randomly combined to generate vignettes for respondents. This design guarded against question order bias (34), allowed for detailed exploration of interactions (35), and minimized the impact of bias due to the perception of being either observed (Hawthorne effect) or evaluated (sentinel effect) (30). After discretizing continuous
variables (eg. respiratory rate) and introducing clinically relevant covariance (eg. removing
norepinephrine from vignettes with a diagnosis of COVID), there were 1.25 million potential
vignettes.

The survey began with a questionnaire to gather respondent demographics (eFigures 1 and 2)
including clinical role, primary clinical area of practice, duration in practice, and country of
practice. We then presented respondents with 10 randomly allocated tabular vignettes (eFigure
3).

For each vignette, respondents were asked “Would you recommend intubation and invasive
ventilation?” and selected an ordinal response from “Definite no”, “Probable no”, “Uncertain”,
“Probable yes”, to “Definite yes.” Respondents could also indicate if any of 6 additional
information types would be helpful for their decision (arterial blood gas, chest radiography,
positive end-expiratory pressure (PEEP), tidal volume measurement, esophageal manometry, or
more observation time, eFigure 4).

Translation and dissemination

The English survey was translated by fluent clinicians into Indonesian, Chinese (simplified),
French, German, Italian, Japanese, Portuguese, Spanish, and Thai. The survey was disseminated
through and/or endorsed by several societies and networks listed in the Acknowledgments.

Descriptive analyses

We reported the number of respondents according to characteristics. Regions, sub-regions, and
countries were defined using the ISO 3166 country codes. We reported the total, number
per respondent, the distribution of responses by respondent characteristics, and the number
requesting each additional information type.

Primary analysis

We analyzed the ordinal intubation recommendation outcome with a Bayesian multilevel
proportional odds model. The model clustered observations by region, country, and
individual to account for repeated measurements. Multilevel structure also allowed for
country-level odds ratios to be related to the number of respondents, by shrinking odds ratios
for countries with few respondents towards 1. This is a quantitative improvement on excluding
sparse categories or combining them into an “Other” category.

Data processing and predictors

We categorized continuous variables to allow non-linearity with interpretable coefficients. We
used all respondent and vignette data as predictors, including bivariate interactions between all
vignette variables. Introducing interactions increased the number of parameters from 47 to
501, therefore we chose prior distributions that prevent overfitting (see Supplement).

Outputs

From the primary analysis, we reported odds ratios for main effects, interactions, and
combinations of country- and region-level random intercepts. We quantified variability
between individuals, countries, and regions using the median odds ratio.

Secondary analyses

Requests for additional information

We performed Bayesian multilevel logistic regression, with the same structure as the primary
analysis, for whether each type of additional information was requested.
Comparing with observed data

We compared survey responses to observed data to address the common criticism of vignette-based surveys that their responses do not reflect real-world decision-making. To do so, we constructed a cohort of all observations of patients receiving oxygen via high-flow nasal cannula or non-invasive ventilation from the Medical Information Mart for Intensive Care IV (MIMIC-IV) database, similar to a prior study (15). For each observation, we used the model from the primary analysis to predict survey responses. Then, we compared the predicted survey responses to the observed outcome of intubation within 3, 8, or 24 hours (15, 43, 44). We reported discrimination and calibration of the predicted recommendations versus observed intubation rates (Supplement, section 10).

Sensitivity analyses

We repeated the primary analysis using only respondents who completed 10 scenarios. We did this because some users were prematurely disconnected during high survey traffic. If they reattempted the survey, their data appeared as two separate respondents. Given that it is unlikely that someone who completed 10 scenarios would reattempt the survey, the 10-response subset likely consists of unique respondents.

We also repeated the primary analysis for responses restricted to each scenario diagnosis. This amounted to forcing an interaction between diagnosis and every other variable in the model. In a final sensitivity analysis, we assessed the proportional odds assumption.
Computational details and reporting

The survey interface was coded in R and shiny(31,45), and stored responses in REDcap.(46,47) We used R and the brms package for all Bayesian analyses.(48,49) Bayesian models used 4 chains of 1000 iterations each, split between warm-up and sampling. We summarized posterior distributions with means and 95% credible intervals (CrI), used complete case analysis for modeling, and defined the region of practical equivalence as an odds ratio of 0.9 to 1.1.(50)

Results

Respondents

Between September 27, 2023 and January 11, 2024, 2,294 clinicians from 74 countries involved in the decision to intubate responded to the survey (Table 1). The top 5 countries were Canada (659, 29%), United States (597, 26%), France (199, 9%), Japan (187, 8%), and Thailand (121, 5%) (eFigure 5). Overall, 711 (31%) completed surveys in a non-English language, most commonly French (258, 11%), Japanese (174, 8%), and Spanish (107, 5%).

Critical care medicine was the primary clinical area of practice for 1,825 (80%). Multiple primary areas were selected by 334 (15%), including anesthesia and emergency medicine. Most respondents were attending physicians (1437, 63%), and 507 (22%) respondents were nurses or respiratory therapists. Median duration of practice was 11 years (interquartile range [IQR] 5 to 20, eFigure 6).

Responses

We gathered 17,235 vignette responses from the 2,294 respondents. 1,372 (60%) answered 10 scenarios (IQR 4 to 10, eFigure 7 & 8). Common responses included “Probable yes” (4,567,
26%), “Uncertain” (4,024, 23%), and “Probable no” (4,190, 24%). The remainder (22%) were either “Definite yes” or “Definite no.” Requests for additional information ranged from 11,450 (66%) for arterial blood gas to 539 (3%) for esophageal manometry.

Primary analysis: Intubation recommendation

Several variables were associated with recommending intubation (Figure 1). Among clinician characteristics, a role of nurse or specialty of emergency medicine were associated with decreased odds of recommending intubation. Among patient characteristics, patient age, frailty, or diagnosis were not associated with recommending intubation. For oxygen saturation, inspired oxygen fraction, respiratory rate, breathing pattern, and level of consciousness, increasing derangement from normal was associated with increasing odds of recommending intubation (Figure 1). Only 2 of the 454 interaction term coefficients in the model had credible intervals not including 1, and in both cases this was likely artifactual due to limitations of an ordinal scale. (Supplement, section 8, eFigures 9 & 10).

Variation by individual, country, and region

We quantified variability between individuals, countries, and regions using the median odds ratio. The median odds ratio between individuals within the same country was 2.60 (CrI 2.48 to 2.73), indicating that randomly switching between individuals increased or decreased the odds of recommending intubation by an average factor of 2.6, after accounting for all other variables in the model. The median odds ratio between countries within the same region was 1.56 (CrI 1.35 to 1.85), and 1.17 (CrI 1.01 to 1.63) between regions.
The odds ratio for recommending intubation by country varied, but there was important uncertainty in these estimates for most countries (Figure 2). Only 6 of 74 countries had odds ratios with 95% credible intervals that did not include 1: Bangladesh (OR 0.39, CrI 0.20 to 0.74), Canada (OR 0.53, CrI 0.40 to 0.70), United States of America (OR 0.63, CrI 0.48 to 0.84), Japan (OR 1.37, CrI 1.02 to 1.86), Thailand (OR 1.48, CrI 1.03 to 2.08), and Argentina (OR 2.23, CrI 1.31 to 3.94).

**Secondary analysis: Additional information**

Factors associated with requesting additional information varied by information type (Figure 3). No respondent or scenario variables were associated with requesting tidal volume or esophageal manometry. Nurse or respiratory therapist role was associated with increased odds of requesting arterial blood gas and chest x-ray, and respiratory therapist role was also associated with requesting PEEP information. Patient state of drowsy and not obeying was associated with increased odds of requesting an arterial blood gas, and decreased odds of requesting PEEP information or more observation time.

**Secondary analysis: Comparison to observed data**

There were 18,263 observations from 826 patients in the MIMIC-IV cohort. Intubation occurred within 3, 8, and 24 hours following 745 (4%), 1,375 (8%), and 2,635 (14%) observations. The average predicted probabilities of each response across all cohort observations were: “Definite no” – 25%, “Probable no” – 41%, “Uncertain” – 20%, “Probable yes” – 11%, “Definite yes” – 3%. The probability of “Definite yes” recommendations was similar to the observed intubation rate within 3 hours, and the probability of “Definite yes” or “Probable yes” recommendations was similar to the probability of observed intubation within 24 hours (eFigure 11). Predicted
intubation recommendations of “Definite yes”, or “Definite yes and “Probable yes”, had similar
discrimination with respect to intubation at 3 hours with an area under receiver operating
curve of 0.71 (eFigures 12 & 13).

Sensitivity analyses

The sensitivity analyses limited to respondents with 10 responses (eFigures 14 & 15) or divided
by diagnosis had similar results to the primary analysis (eFigures 16-20). The proportional odds
assumption was supported for all variables except nurse respondent role (eFigures 21-24).

Discussion

In this survey of when to intubate patients with acute hypoxemic respiratory failure, we
gathered 17,235 vignette-based intubation recommendations from 2,294 clinicians in 74
countries. Oxygenation, breathing pattern, and level of consciousness had a strong influence on
intubation recommendations, while diagnosis, norepinephrine use, and duration of symptoms
did not. We also found important variation in recommendations across clinicians and countries.
The results persisted in multiple sensitivity analyses, and the survey responses had good
calibration and moderate discrimination with respect to external data. Our findings suggest
important elements of criteria to guide intubation, but also highlight uncertainty as to which
criteria may be optimal in a given scenario.

We found that oxygenation, breathing pattern, and level of consciousness were strongly
associated with intubation recommendations. These variables are highly relevant to the
decision to intubate, as evidenced by their inclusion as intubation criteria in clinical
trials(23,51–53) and by qualitative and survey data.(11–13) Our findings are notable for the
magnitude of the odds ratios and the robust dose-response relationship between physiological
derangement and probability of recommending intubation. For example, the probability of a
respondent answering “Definite yes” or “Probable yes” increased from 33% to 53% after
applying the odds ratio for use of neck muscles (2.32), and increased further to 76% after
applying the odds ratio for use of neck muscles and abdominal paradox (6.39). The odds ratios
relating to breathing pattern, inspired oxygen fraction, and level of consciousness were large
enough that extreme values of any of these variables likely determine a clinician’s
recommendation.

A respondent role of nurse was associated with modestly decreased odds of recommending
intubation. Nurse respondents were also more likely to request arterial blood gases and chest
x-rays. Potential explanations could include that nurses are not the final decision-maker, that
nurses are less convinced of the benefits of invasive ventilation in a given scenario, or because
higher rates of burnout in ICU nurses than ICU physicians(54) may contribute to a sense of
therapeutic nihilism.

We identified several variables that were not associated with intubation recommendations,
including diagnosis, norepinephrine use, duration in current state, and almost all of the 454
interaction terms in the model. The findings for diagnosis, where not even COVID-19
pneumonia was associated with decreased odds, suggest that intubation decisions are based on
how a diagnosis affects oxygenation, breathing pattern, and level of consciousness, rather than
on the diagnosis itself. Our findings are also consistent with other studies showing that
norepinephrine use alone rarely determines intubation decision-making (16).
The lack of association between duration in current state or 99.5% of the interaction terms and intubation recommendations is somewhat surprising. Both clinical trajectory and the relationships between clinical variables have face validity as important components of intubation decision-making. Although one explanation could be that “duration in current state” does not meaningfully capture clinical trajectory, our findings are consistent with cohort studies of patients with acute hypoxemic respiratory failure that suggest those who receive invasive ventilation are often those who fail to improve, and remain in the same (hypoxemic/tachypneic) state over time, as opposed to those who worsen. Regarding interactions, our findings suggest that clinicians gravitate towards parsimonious, heuristic approaches which do not rely on interactions.

We found important variation in intubation recommendations between individual clinicians and countries. The median change in odds of recommending intubation between two individuals from the same country was 2.60, which was numerically similar to the change in odds between patients on inspired oxygen of 1.0 versus 0.7, patients with and without neck muscle use, or patients who are drowsy but obeying versus alert. These findings may partially explain prior work showing surprisingly low rates of intubation in ICU patients with AHRF who attain relatively severe degrees of physiologic derangement.

While variation between countries was less marked than variation between individuals, odds ratios by country ranged from 0.39 to 2.23. A case-based survey of 1,136 intensivists focused on COVID-19 pneumonia management also found large variation in the proportion of respondents recommending intubation, ranging from 0% in Australia-New Zealand to 23% in Asia. Our finding that clinicians in Canada and the USA were so much less likely to recommend intubation...
was striking, because health administrative data shows higher age-adjusted rates of invasive ventilation in both countries relative to the United Kingdom. Potential explanations for country-level variation could relate to interpretation of the survey question, vignettes, and responses; access to resources including monitoring, non-invasive oxygen supports, and ventilators; or cultural, religious, and social practices and expectations regarding life support.

The most important limitation of our work is that a vignette-based survey is not guaranteed to reflect decisions made in clinical practice. Our secondary analysis using MIMIC-IV addresses this limitation, but provides only modest assurance that the survey responses match true clinical behaviour.

Selection bias due to non-response is another important limitation. We were unable to measure the response rate, which is one conventional way to assess vulnerability to selection bias. However, response rates in other surveys concerning intubation do not rule out the possibility of non-response bias, with rates ranging from 11% (12) to 50% (61). Surveys in other areas of critical care medicine have had response rates as low as 0.8% (62). Further, to calculate response rates, these surveys generally limited their distribution to medical societies with known membership. Our survey included non-physician clinicians involved in the decision to intubate, and physicians from countries and networks where the denominator is unknown.

On balance, we felt that the benefits of a broader, larger sample more reflective of the global and team-based nature of intensive care medicine outweighed the potential harms of an unknown response rate.
Further limitations relate to survey design and modeling. The number of possible vignettes was much higher than the number of possible respondents, reflecting the multiplicative complexity of nuanced decision-making in medicine. This limited our ability to appreciate higher-order interactions. However, we obtained sufficient responses to evaluate bivariate interactions between any two variables. The proportional odds assumption implied by our modeling approach does not hold for all coefficients, based on our sensitivity analysis. However, proportional odds models can provide accurate treatment effect estimates even when there are deviations from the proportional odds assumption. (63–66) It was possible for respondents to complete the survey more than once, and so our primary analysis may be underestimating the individual-level variation. Reassuringly, our sensitivity analysis focused on those who fully completed the survey supported the primary analysis results.

**Conclusion**

In this survey of 2,294 clinicians from 74 countries, intubation recommendations for patients with AHRF were associated with oxygenation, breathing pattern, and level of consciousness, and varied across individuals and countries. These data can help justify and inform randomized trials of physiological thresholds to guide intubation decisions that may improve patient-important outcomes.
### Table 1

<table>
<thead>
<tr>
<th>Survey language</th>
<th>Total</th>
<th>Northern America</th>
<th>Europe</th>
<th>Asia</th>
<th>Oceania</th>
<th>Latin America and the Caribbean</th>
<th>Africa</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bahasa Indonesia</td>
<td>&lt; 5 (0%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>&lt; 5 (0%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>Chinese (simplified)</td>
<td>&lt; 5 (0%)</td>
<td>&lt; 5 (0%)</td>
<td>0 (0%)</td>
<td>&lt; 5 (1%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>English</td>
<td>1583 (69%)</td>
<td>1222 (97%)</td>
<td>113 (25%)</td>
<td>154 (37%)</td>
<td>72 (100%)</td>
<td>10 (15%)</td>
<td>12 (63%)</td>
</tr>
<tr>
<td>German</td>
<td>10 (0%)</td>
<td>0 (0%)</td>
<td>10 (2%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>French</td>
<td>258 (11%)</td>
<td>31 (2%)</td>
<td>218 (48%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>&lt; 5 (3%)</td>
<td>7 (37%)</td>
</tr>
<tr>
<td>Italian</td>
<td>58 (3%)</td>
<td>0 (0%)</td>
<td>58 (13%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>Japanese</td>
<td>174 (8%)</td>
<td>0 (0%)</td>
<td>174 (41%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>Portuguese</td>
<td>10 (0%)</td>
<td>0 (0%)</td>
<td>&lt; 5 (0%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>9 (13%)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>Spanish</td>
<td>107 (5%)</td>
<td>&lt; 5 (0%)</td>
<td>58 (13%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>47 (69%)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>Thai</td>
<td>88 (4%)</td>
<td>0 (0%)</td>
<td>88 (21%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>Specialty</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Critical care medicine</td>
<td>1825 (80%)</td>
<td>1051 (84%)</td>
<td>385 (84%)</td>
<td>246 (58%)</td>
<td>67 (93%)</td>
<td>61 (90%)</td>
<td>15 (79%)</td>
</tr>
<tr>
<td>Anesthesia</td>
<td>302 (13%)</td>
<td>80 (6%)</td>
<td>91 (20%)</td>
<td>116 (28%)</td>
<td>8 (11%)</td>
<td>&lt; 5 (3%)</td>
<td>5 (26%)</td>
</tr>
<tr>
<td>Emergency medicine</td>
<td>340 (15%)</td>
<td>192 (15%)</td>
<td>46 (10%)</td>
<td>89 (21%)</td>
<td>&lt; 5 (1%)</td>
<td>7 (10%)</td>
<td>5 (26%)</td>
</tr>
<tr>
<td>Other</td>
<td>202 (9%)</td>
<td>97 (8%)</td>
<td>24 (5%)</td>
<td>73 (17%)</td>
<td>0 (0%)</td>
<td>5 (7%)</td>
<td>&lt; 5 (16%)</td>
</tr>
<tr>
<td>Role</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nurse</td>
<td>310 (14%)</td>
<td>214 (17%)</td>
<td>58 (13%)</td>
<td>12 (3%)</td>
<td>24 (33%)</td>
<td>&lt; 5 (1%)</td>
<td>&lt; 5 (5%)</td>
</tr>
<tr>
<td>Attending physician</td>
<td>1437 (63%)</td>
<td>689 (55%)</td>
<td>331 (72%)</td>
<td>323 (77%)</td>
<td>32 (44%)</td>
<td>49 (72%)</td>
<td>13 (68%)</td>
</tr>
<tr>
<td>Other clinical role</td>
<td>146 (6%)</td>
<td>106 (8%)</td>
<td>13 (3%)</td>
<td>18 (4%)</td>
<td>&lt; 5 (4%)</td>
<td>6 (9%)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>Trainee physician</td>
<td>204 (9%)</td>
<td>89 (7%)</td>
<td>40 (9%)</td>
<td>56 (13%)</td>
<td>11 (15%)</td>
<td>&lt; 5 (6%)</td>
<td>&lt; 5 (21%)</td>
</tr>
<tr>
<td>Respiratory therapist</td>
<td>197 (9%)</td>
<td>158 (13%)</td>
<td>16 (3%)</td>
<td>12 (3%)</td>
<td>&lt; 5 (3%)</td>
<td>8 (12%)</td>
<td>&lt; 5 (5%)</td>
</tr>
<tr>
<td>Duration in practice</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0-5 years</td>
<td>662 (29%)</td>
<td>404 (32%)</td>
<td>111 (24%)</td>
<td>109 (26%)</td>
<td>13 (18%)</td>
<td>21 (31%)</td>
<td>&lt; 5 (21%)</td>
</tr>
<tr>
<td>6-10 years</td>
<td>440 (19%)</td>
<td>232 (18%)</td>
<td>103 (22%)</td>
<td>76 (18%)</td>
<td>18 (25%)</td>
<td>&lt; 5 (6%)</td>
<td>7 (37%)</td>
</tr>
<tr>
<td>11-20 years</td>
<td>486 (21%)</td>
<td>243 (19%)</td>
<td>113 (25%)</td>
<td>99 (24%)</td>
<td>12 (17%)</td>
<td>14 (21%)</td>
<td>5 (26%)</td>
</tr>
<tr>
<td>20-60 years</td>
<td>686 (30%)</td>
<td>364 (29%)</td>
<td>127 (28%)</td>
<td>136 (32%)</td>
<td>28 (39%)</td>
<td>29 (43%)</td>
<td>&lt; 5 (11%)</td>
</tr>
<tr>
<td>Missing</td>
<td>20 (1%)</td>
<td>13 (1%)</td>
<td>&lt; 5 (1%)</td>
<td>&lt; 5 (0%)</td>
<td>&lt; 5 (1%)</td>
<td>0 (0%)</td>
<td>&lt; 5 (5%)</td>
</tr>
<tr>
<td>Responses</td>
<td>Total</td>
<td>Northern America</td>
<td>Europe</td>
<td>Asia</td>
<td>Oceania</td>
<td>Latin America and the Caribbean</td>
<td>Africa</td>
</tr>
<tr>
<td>-----------</td>
<td>-------</td>
<td>------------------</td>
<td>--------</td>
<td>------</td>
<td>---------</td>
<td>---------------------------------</td>
<td>-------</td>
</tr>
<tr>
<td>Responses</td>
<td>17235</td>
<td>9168</td>
<td>3605</td>
<td>3159</td>
<td>606</td>
<td>542</td>
<td>155</td>
</tr>
<tr>
<td>Respondents</td>
<td>2294</td>
<td>1256</td>
<td>458</td>
<td>421</td>
<td>72</td>
<td>68</td>
<td>19</td>
</tr>
<tr>
<td>Mean number per respondent</td>
<td>7.5</td>
<td>7.3</td>
<td>7.9</td>
<td>7.5</td>
<td>8.4</td>
<td>8</td>
<td>8.2</td>
</tr>
<tr>
<td>Respondents with 10 responses</td>
<td>1372 (60%)</td>
<td>712 (57%)</td>
<td>294 (64%)</td>
<td>254 (60%)</td>
<td>54 (75%)</td>
<td>44 (65%)</td>
<td>14 (74%)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Recommendation</th>
<th>Total</th>
<th>Northern America</th>
<th>Europe</th>
<th>Asia</th>
<th>Oceania</th>
<th>Latin America and the Caribbean</th>
<th>Africa</th>
</tr>
</thead>
<tbody>
<tr>
<td>Definite yes</td>
<td>2825 (16%)</td>
<td>1120 (12%)</td>
<td>757 (21%)</td>
<td>667 (21%)</td>
<td>111 (18%)</td>
<td>139 (26%)</td>
<td>31 (20%)</td>
</tr>
<tr>
<td>Probable yes</td>
<td>4567 (26%)</td>
<td>2310 (25%)</td>
<td>962 (27%)</td>
<td>915 (29%)</td>
<td>185 (31%)</td>
<td>162 (30%)</td>
<td>33 (21%)</td>
</tr>
<tr>
<td>Uncertain</td>
<td>4024 (23%)</td>
<td>2119 (23%)</td>
<td>899 (25%)</td>
<td>699 (22%)</td>
<td>137 (23%)</td>
<td>132 (24%)</td>
<td>38 (25%)</td>
</tr>
<tr>
<td>Probable no</td>
<td>4190 (24%)</td>
<td>2533 (28%)</td>
<td>691 (19%)</td>
<td>711 (23%)</td>
<td>133 (22%)</td>
<td>84 (15%)</td>
<td>38 (25%)</td>
</tr>
<tr>
<td>Definite no</td>
<td>1629 (9%)</td>
<td>1086 (12%)</td>
<td>296 (8%)</td>
<td>167 (5%)</td>
<td>40 (7%)</td>
<td>25 (5%)</td>
<td>15 (10%)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Additional information requested</th>
<th>Total</th>
<th>Northern America</th>
<th>Europe</th>
<th>Asia</th>
<th>Oceania</th>
<th>Latin America and the Caribbean</th>
<th>Africa</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arterial blood gas</td>
<td>11450 (66%)</td>
<td>5821 (63%)</td>
<td>2253 (62%)</td>
<td>2474 (78%)</td>
<td>410 (68%)</td>
<td>379 (70%)</td>
<td>113 (73%)</td>
</tr>
<tr>
<td>Chest x-ray</td>
<td>8294 (48%)</td>
<td>4454 (49%)</td>
<td>1137 (32%)</td>
<td>1974 (62%)</td>
<td>383 (63%)</td>
<td>263 (49%)</td>
<td>83 (54%)</td>
</tr>
<tr>
<td>More observation time</td>
<td>6840 (40%)</td>
<td>3962 (43%)</td>
<td>1261 (35%)</td>
<td>1098 (35%)</td>
<td>279 (46%)</td>
<td>163 (30%)</td>
<td>77 (50%)</td>
</tr>
<tr>
<td>PEEP (if on NIV)</td>
<td>4353 (25%)</td>
<td>2101 (23%)</td>
<td>850 (24%)</td>
<td>1036 (33%)</td>
<td>211 (35%)</td>
<td>110 (20%)</td>
<td>45 (29%)</td>
</tr>
<tr>
<td>Tidal volume (if on NIV)</td>
<td>3625 (21%)</td>
<td>1671 (18%)</td>
<td>737 (20%)</td>
<td>908 (29%)</td>
<td>184 (30%)</td>
<td>105 (19%)</td>
<td>20 (13%)</td>
</tr>
<tr>
<td>Esophageal pressure</td>
<td>539 (3%)</td>
<td>84 (1%)</td>
<td>207 (6%)</td>
<td>203 (6%)</td>
<td>10 (2%)</td>
<td>31 (6%)</td>
<td>&lt;5 (3%)</td>
</tr>
</tbody>
</table>
Figures

Figure 1: Posterior odds ratios for recommending intubation

Caption: This figure shows the posterior odds ratios across respondent, patient, and physiological variables for the strength of intubation recommendation (based on the ordinal response scale comprised of Definite no / Probable no / Uncertain / Probable Yes / Definite Yes). Odds ratios greater than 1 indicate variables associated with higher odds of recommending intubation, ie. a recommendation more towards the “Definite yes” end of the scale, whereas odds ratios less than 1 indicate variables associated with lower odds of recommending intubation, ie. a recommendation more towards the “Definite no” end of the scale.
### Posterior odds ratios for recommending intubation, by category of variable

#### Role
- **Nurse**: 0.65 (0.54 to 0.77)
- **Other clinical role**: 0.85 (0.65 to 1.02)
- **Respiratory therapist**: 0.99 (0.87 to 1.1)
- **Attending physician**: Reference
- **Trainee physician**: 1.11 (0.97 to 1.37)

#### Duration in practice
- **5 or fewer years**: Reference
- **6 to 10 years**: 1.03 (0.96 to 1.1)
- **11 to 20 years**: 0.96 (0.85 to 1.04)
- **21 or more years**: 0.97 (0.87 to 1.05)

#### Patient frailty
- **Independent**: 1.03 (0.94 to 1.18)
- **Independent with medical problems**: 1.03 (0.94 to 1.18)
- **Assistance for some IADLs**: 1.03 (0.91 to 1.2)

#### Oxygen saturation
- **95-97**: Reference
- **92-94**: 1.11 (0.97 to 1.35)
- **89-91**: 1.74 (1.43 to 2.1)
- **85-88**: 4.12 (3.33 to 5.09)

#### Oxygen device
- **High-flow nasal cannula**: Reference
- **Non-invasive ventilation**: 1.31 (1.09 to 1.54)

#### Breathing Pattern
- **Normal**: Reference
- **Neck muscle use, no abdominal paradox**: 2.32 (1.98 to 2.73)
- **Neck muscle use and abdominal paradox**: 6.39 (5.02 to 8.13)

#### Norepinephrine use
- **No**: Reference
- **Yes**: 1.12 (0.98 to 1.38)

#### Specialty
- **Emergency medicine**: 0.85 (0.71 to 1)
- **Other**: 1.0 (0.89 to 1.1)
- **Anesthesia**: 1 (0.91 to 1.1)
- **Critical care medicine**: 1.06 (0.97 to 1.25)

#### Patient age
- **20-39 years**: Reference
- **40-49 years**: 1.04 (0.95 to 1.22)
- **50-59 years**: 1.05 (0.96 to 1.24)
- **60-70 years**: 1.06 (0.97 to 1.3)

#### Diagnosis
- **COVID-19 pneumonia**: 0.88 (0.72 to 1.01)
- **Community-acquired pneumonia**: Reference
- **Influenza pneumonia**: 1.02 (0.93 to 1.14)
- **Sepsis**: 1.03 (0.93 to 1.18)
- **Pancreatitis**: 1.04 (0.94 to 1.22)

#### Inspired oxygen fraction
- **0.4**: Reference
- **0.5**: 1.17 (1.09 to 1.45)
- **0.6**: 1.71 (1.30 to 2.11)
- **0.7**: 2.77 (2.25 to 3.7)
- **0.8**: 4.11 (3.33 to 5.1)
- **0.9**: 6.38 (5.18 to 7.84)
- **1.0**: 8.98 (7.1 to 11.36)

#### Respiratory rate
- **15-20**: Reference
- **21-24**: 1.02 (0.92 to 1.15)
- **25-27**: 1.29 (1.02 to 1.6)
- **28-32**: 1.77 (1.48 to 2.11)
- **33-40**: 3.12 (2.58 to 3.78)

#### Level of consciousness
- **Alert and obeying**: Reference
- **Drowsy but obeying**: 2.49 (2.11 to 2.95)
- **Drowsy, not obeying**: 9.83 (7.95 to 12.32)

#### Duration in current state
- **10 minutes**: Reference
- **30 minutes**: 1.01 (0.92 to 1.14)
- **1 hour**: 1.07 (0.95 to 1.28)
- **2 hours**: 1.25 (1 to 1.55)
- **4 hours**: 1.13 (0.98 to 1.4)

---

Odds ratios less than 1 = less likely to recommend intubation
Odds ratios more than 1 = more likely to recommend intubation
Figure 2: International variation in odds of recommending intubation

Caption: This figure shows the odds ratios for recommending intubation associated with respondent countries. The mean odds ratio is shown both as a black dot in the peripheral ring, and as the colour of the country within the map. The 95% credible interval or uncertainty or the estimate is shown by the size of the light blue bar in the peripheral ring. The odds ratios correspond to the combined random intercepts for region and country from the multilevel Bayesian proportional odds model. This means that countries with few respondents have wide uncertainty and their posterior estimates have been shrunk towards an odds ratio of 1.
Figure 3: Posterior odds ratios for requesting additional information

**Arterial blood gas**

- Nurse (vs attending physician): 4.43 (3.18 to 6.28)
- Respiratory therapist (vs attending physician): 3.11 (2.1 to 4.6)
- Drowsy, not obeying (vs alert): 2.05 (1.53 to 2.81)
- Drowsy but obeying (vs alert): 1.75 (1.4 to 2.23)
- Sepsis and oxygen saturation 89 to 91*: 1.31 (1 to 1.76)
- Use of neck muscles, abdominal paradox, drowsy, not obeying*: 0.62 (0.45 to 0.87)

**PEEP value (if on NIV)**

- Respiratory therapist (vs attending physician): 3.95 (2.27 to 6.61)
- Drowsy, not obeying (vs alert): 0.63 (0.51 to 0.81)

**Chest x-ray**

- Respiratory therapist (vs attending physician): 5.4 (3.39 to 8.43)
- Nurse (vs attending physician): 3.16 (2.14 to 4.76)
- COVID-19 pneumonia: 0.76 (0.61 to 0.97)

**More observation time**

- Oxygen saturation 85 to 88 on FiO2 1.0*: 0.6 (0.4 to 0.95)
- Duration in current state: 1 hour: 0.58 (0.46 to 0.76)
- Drowsy and not obeying on FiO2 1.0*: 0.57 (0.4 to 0.83)
- Drowsy and not obeying: 0.4 (0.31 to 0.52)
- Duration in current state: 2 hours: 0.39 (0.32 to 0.49)
- Duration in current state: 4 hours: 0.29 (0.24 to 0.36)
Caption: This figure shows the posterior odds ratios associated with requesting each additional type of information. For clarity, we present only the posterior odds ratios with 95% credible intervals that do not include 1. For two types of additional information (tidal volume if on non-invasive ventilation, and esophageal manometry), there were no odds ratios that met this criterion. PEEP = positive end-expiratory pressure, NIV = non-invasive ventilation, COVID-19 = coronavirus 2019. *These odds ratios are interactions between the described variables.
Data sharing statement

The data used in this study was deidentified and is available through reasonable request to the author. We plan to share the deidentified data through the Health Data Nexus repository (https://healthdatanexus.ai/) organized by the Temerty Center for Artificial Intelligence Research and Education in Medicine at the University of Toronto, in Toronto, Canada.
References


29. Surveys [Internet]. ESICM. [cited 2024 Feb 12]. Available from: https://www.esicm.org/research/surveys/


44. Johnson A, Bulgarelli L, Pollard T, Horng S, Celi LA, Mark R. MIMIC-IV (version 0.4) [Internet]. PhysioNet; 2020. Available from: https://doi.org/10.13026/a3wn-hq05


