



**HAL**  
open science

## Ventilator-free days as an outcome: approaches on sample size estimation from a systematic review

Laurent Renard-Triche

► **To cite this version:**

Laurent Renard-Triche. Ventilator-free days as an outcome: approaches on sample size estimation from a systematic review. Human health and pathology. 2022. dumas-03769284

**HAL Id: dumas-03769284**

**<https://dumas.ccsd.cnrs.fr/dumas-03769284v1>**

Submitted on 5 Sep 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNIVERSITE CLERMONT AUVERGNE  
UFR DE MEDECINE ET DES PROFESSIONS PARAMEDICALES

THESE D'EXERCICE

Pour le

DIPLOME D'ETAT DE DOCTEUR EN MEDECINE

Par

RENARD-TRICHE Laurent, Daniel, Jean

Présentée et soutenue publiquement le 07 septembre 2022

**LES VENTILATOR-FREE DAYS COMME CRITERE DE JUGEMENT  
PRINCIPAL : APPROCHES POUR L'ESTIMATION DU NOMBRE  
DE SUJETS NECESSAIRES A PARTIR D'UNE REVUE SYSTEMATIQUE.**

Directeur de thèse :

Monsieur PEREIRA Bruno, Docteur en Biostatistiques, Direction de la Recherche Clinique et de l'Innovation, CHU Clermont-Ferrand

Co-directeur de thèse :

Monsieur JABAUDON Matthieu, Professeur en Anesthésiologie-Réanimation et Médecine Péri-Opératoire, UFR de Médecine et des Professions paramédicales, Clermont-Ferrand

Président du jury :

Monsieur FUTIER Emmanuel, Professeur en Anesthésiologie-Réanimation et Médecine Péri-Opératoire, UFR de Médecine et des Professions paramédicales, Clermont-Ferrand

Membres du jury :

Madame GARRIDO Myriam, Chargée de recherche, UMR EPIA, INRAE Clermont-Auvergne-Rhône-Alpes / LMBP, Université Clermont Auvergne  
Madame MOLLEVI Caroline, Docteure en Biostatistiques, Département de l'Information Médicale, CHU Montpellier



UNIVERSITE CLERMONT AUVERGNE  
UFR DE MEDECINE ET DES PROFESSIONS PARAMEDICALES

THESE D'EXERCICE

Pour le

DIPLOME D'ETAT DE DOCTEUR EN MEDECINE

Par

RENARD-TRICHE Laurent, Daniel, Jean

Présentée et soutenue publiquement le 07 septembre 2022

**LES VENTILATOR-FREE DAYS COMME CRITERE DE JUGEMENT  
PRINCIPAL : APPROCHES POUR L'ESTIMATION DU NOMBRE  
DE SUJETS NECESSAIRES A PARTIR D'UNE REVUE SYSTEMATIQUE.**

Directeur de thèse :

Monsieur PEREIRA Bruno, Docteur en Biostatistiques, Direction de la Recherche Clinique et de l'Innovation, CHU Clermont-Ferrand

Co-directeur de thèse :

Monsieur JABAUDON Matthieu, Professeur en Anesthésiologie-Réanimation et Médecine Péri-Opératoire, UFR de Médecine et des Professions paramédicales, Clermont-Ferrand

Président du jury :

Monsieur FUTIER Emmanuel, Professeur en Anesthésiologie-Réanimation et Médecine Péri-Opératoire, UFR de Médecine et des Professions paramédicales, Clermont-Ferrand

Membres du jury :

Madame GARRIDO Myriam, Chargée de recherche, UMR EPIA, INRAE Clermont-Auvergne-Rhône-Alpes / LMBP, Université Clermont Auvergne  
Madame MOLLEVI Caroline, Docteure en Biostatistiques, Département de l'Information Médicale, CHU Montpellier



## **UNIVERSITE CLERMONT AUVERGNE**

---

PRESIDENTS HONORAIRES  
UNIVERSITE D'AUVERGNE

: **JOYON** Louis  
: **DOLY** Michel  
: **TURPIN** Dominique  
: **VEYRE** Annie  
: **DULBECCO** Philippe  
: **ESCHALIER** Alain

PRESIDENTS HONORAIRES  
UNIVERSITE BLAISE PASCAL

: **CABANES** Pierre  
: **FONTAINE** Jacques  
: **BOUTIN** Christian  
: **MONTEIL** Jean-Marc  
: **ODOUARD** Albert  
: **LAVIGNOTTE** Nadine

PRESIDENT DE L'UNIVERSITE  
PREMIERE VICE-PRESIDENTE  
CHARGEE DU PILOTAGE ET DES MOYENS  
VICE PRESIDENTE CHARGEE DE LA FORMATION  
VICE-PRESIDENTE CHARGEE DE LA RECHERCHE  
DIRECTEUR GENERAL DES SERVICES

: **BERNARD** Mathias  
  
: **FOGLI** Anne  
: **PEYRARD** Françoise  
: **PREVOT** Vanessa  
: **PAQUIS** François



## **UFR DE MEDECINE ET DES PROFESSIONS PARAMEDICALES**

DOYENS HONORAIRES

: **DETEIX** Patrice  
: **CHAZAL** Jean

DOYEN  
RESPONSABLE ADMINISTRATIVE

: **CLAVELOU** Pierre  
: **ROBERT** Gaëlle

## LISTE DU PERSONNEL ENSEIGNANT

### **PROFESSEURS HONORAIRES :**

MM. BACIN Franck - BEGUE René-Jean - BEYTOUT Jean - BOIRE Jean-Yves -  
BOITEUX Jean-Paul - BOMMELAER Gilles - BOUCHER Daniel - BUSSIÈRE Jean-Louis -  
CANO Noël - CASSAGNES Jean - CATILINA Pierre - CHABANNES Jacques -  
CHAMOUX Alain - CHAZAL Jean - CHIPPONI Jacques - CHOLLET Philippe -  
CITRON Bernard - CLEMENT Gilles - COUDERT Jean - DASTUGUE Bernard -  
DAUPLAT Jacques - DECHELOTTE Pierre - DEMEOCQ François -  
DE RIBEROLLES Charles - DETEIX Patrice - ESCANDE Georges - Mme FONCK Yvette -  
M. GENTOU Claude - Mme GLANDDIER Phyllis - MM. IRTHUM Bernard -  
JACQUETIN Bernard - Mme LAVARENNE Jeanine - MM. LAVERAN Henri -  
LESOURD Bruno - LEVAI Jean-Paul - MAGE Gérard - MICHEL Jean-Luc -  
MONDIE Jean-Michel - PHILIPPE Pierre - PLAGNE Robert - PLANCHE Roger -  
PONSONAILLE Jean - REY Michel - Mme RIGAL Danièle - MM. ROZAN Raymond -  
SCHOEFFLER Pierre - SIROT Jacques - RIBAL Jean-Pierre - SOUTEYRAND Pierre -  
TANGUY Alain - TERVER Sylvain - THIEBLOT Philippe - TOURNILHAC Michel -  
VANNEUVILLE Guy - VIALLET Jean-François - Mme VEYRE Annie

### **PROFESSEURS EMERITES :**

MM. AUMAITRE Olivier - AVAN Paul - BAZIN Jean-Etienne - CAILLAUD Denis -  
DAPOIGNY Michel - DUBRAY Claude - ESCHALIER Alain - KEMENY Jean-Louis -  
LABBE André - Mme LAFEUILLE Hélène - MM. LEMERY Didier - LUSSON Jean-René -  
POULY Jean-Luc

## **PROFESSEURS DES UNIVERSITES - PRATICIENS HOSPITALIERS**

### **CLASSE EXCEPTIONNELLE**

M.	VAGO Philippe	Histologie, Embryologie, Cytogénétique
M.	DURIF Franck	Neurologie
M.	BOYER Louis	Radiologie et Imagerie Médicale Option Clinique
M.	CANIS Michel	Gynécologie-Obstétrique
Mme	PENAULT-LLORCA Frédérique	Anatomie et Cytologie Pathologiques
M.	BIGNON Yves Jean	Cancérologie option Biologique
M.	BOIRIE Yves	Nutrition Humaine
M.	CLAVELOU Pierre	Neurologie
M.	GILAIN Laurent	Oto-Rhino-Laryngologie
M.	LEMAIRE Jean-Jacques	Neurochirurgie
M.	CAMILLERI Lionel	Chirurgie Thoracique et Cardio-Vasculaire
M.	LLORCA Pierre-Michel	Psychiatrie d'Adultes
M.	PEZET Denis	Chirurgie Digestive
M.	SOUWEINE Bertrand	Réanimation Médicale
M.	BOISGARD Stéphane	Chirurgie Orthopédique et Traumatologie
Mme	DUCLOS Martine	Physiologie
M.	SCHMIDT Jeannot	Médecine d'Urgence

M.	BERGER Marc	Hématologie
M.	GARCIER Jean-Marc	Anatomie – Radiologie et Imagerie Médicale
M.	SOUBRIER Martin	Rhumatologie
M.	ABERGEL Armando	Hépatologie
Mme	BARTHELEMY Isabelle	Chirurgie Maxillo-Faciale
M.	RUIVARD Marc	Médecine Interne

**1<sup>ère</sup> CLASSE**

M.	VERRELLE Pierre	Radiothérapie Option Clinique
M.	D'INCAN Michel	Dermatologie – Vénérologie
Mme	JALENQUES Isabelle	Psychiatrie d'Adultes
M.	GERBAUD Laurent	Epidémiologie, Economie de la Santé et Prévention
M.	TAUVERON Igor	Endocrinologie et Maladies Métaboliques
M.	MOM Thierry	Oto-Rhino-Laryngologie
M.	RICHARD Ruddy	Physiologie
M.	SAPIN-DEFOUR Vincent	Biochimie et Biologie Moléculaire
M.	BAY Jacques-Olivier	Cancérologie
M.	COUDEYRE Emmanuel	Médecine Physique et de Réadaptation
Mme	GODFRAIND Catherine	Anatomie et Cytologie Pathologiques
M.	LAURICHESSE Henri	Maladies Infectieuses et Tropicales
M.	TOURNILHAC Olivier	Hématologie
M.	CHIAMBARETTA Frédéric	Ophthalmologie
M.	FILAIRE Marc	Anatomie – Chirurgie Thoracique et Cardio-Vasculaire
M.	GALLOT Denis	Gynécologie-Obstétrique
M.	GUY Laurent	Urologie
M.	TRAORE Ousmane	Hygiène Hospitalière
M.	ANDRE Marc	Médecine Interne
M.	BONNET Richard	Bactériologie, Virologie
M.	CACHIN Florent	Biophysique et Médecine Nucléaire
M.	COSTES Frédéric	Physiologie
M.	FUTIER Emmanuel	Anesthésiologie-Réanimation
Mme	HENG Anne-Elisabeth	Néphrologie
M.	MOTREFF Pascal	Cardiologie
Mme	PICKERING Gisèle	Pharmacologie Clinique
M.	RABISCHONG Benoît	Gynécologie-Obstétrique
M.	CHABROT Pascal	Radiologie et Imagerie Médicale
M.	DESCAMPS Stéphane	Chirurgie Orthopédique et Traumatologique
Mme	HENQUELL Cécile	Bactériologie Virologie
M.	POMEL Christophe	Cancérologie – Chirurgie Générale

**2<sup>ème</sup> CLASSE**

Mme	CREVEAUX Isabelle	Biochimie et Biologie Moléculaire
M.	FAICT Thierry	Médecine Légale et Droit de la Santé
Mme	KANOLD LASTAWIECKA Justyna	Pédiatrie
M.	TCHIRKOV Andréï	Cytologie et Histologie
M.	CORNELIS François	Génétique

M.	LESENS Olivier	Maladies Infectieuses et Tropicales
M.	AUTHIER Nicolas	Pharmacologie Médicale
M.	BROUSSE Georges	Psychiatrie Adultes – Addictologie
M.	BUC Emmanuel	Chirurgie Digestive
M.	LAUTRETTE Alexandre	Néphrologie – Réanimation Médicale
Mme	BRUGNON Florence	Biologie et Médecine du Développement et de la Reproduction
M.	ESCHALIER Romain	Cardiologie
M.	MERLIN Etienne	Pédiatrie
Mme	TOURNADRE Anne	Rhumatologie
M.	DURANDO Xavier	Cancérologie
M.	DUTHEIL Frédéric	Médecine et Santé au Travail
Mme	FANTINI Maria Livia	Neurologie
M.	SAKKA Laurent	Anatomie – Neurochirurgie
M.	BOURDEL Nicolas	Gynécologie-Obstétrique
M.	GUIEZE Romain	Hématologie
M.	POINCLOUX Laurent	Gastro-entérologie
M.	SOUTEYRAND Géraud	Cardiologie
M.	EVARD Bertrand	Immunologie
M.	POIRIER Philippe	Parasitologie et Mycologie
Mme	PHAM DANG Nathalie	Chirurgie Maxillo-Faciale et Stomatologie
Mme	SARRET Catherine	Pédiatrie
M.	BOUVIER Damien	Biochimie et Biologie Moléculaire
M.	BUISSON Anthony	Gastro-entérologie
Mme	CASSAGNES Lucie	Radiologie et Imagerie Médicale
M.	GAGNIERE Johan	Chirurgie Viscérale et Digestive
M.	JABAUDON-GANDET Matthieu	Anesthésiologie-Réanimation et Médecine Péri-Opératoire
M.	LEBRETON Aurélien	Hématologie
M.	MOISSET Xavier	Neurologie
M.	SAMALIN Ludovic	Psychiatrie d'Adultes
M.	THAVEAU Fabien	Chirurgie Vasculaire

## PROFESSEURS DES UNIVERSITES DE MEDECINE GENERALE

### *1<sup>ère</sup> CLASSE*

M.	VORILHON Philippe	Médecine Générale
Mme	LAPORTE Catherine	Médecine Générale

## PROFESSEURS DES UNIVERSITES

### *2<sup>ème</sup> CLASSE*

Mme	MALPUECH-BRUGERE Corinne	Nutrition Humaine
-----	--------------------------	-------------------



## PROFESSEURS ASSOCIES DES UNIVERSITES

Mme	BOTTET-MAULOUBIER Anne	Médecine Générale
M.	CAMBON Benoît	Médecine Générale
M.	TANGUY Gilles	Médecine Générale
M.	BERNARD Pierre	Médecine Générale

## MAITRES DE CONFERENCES DES UNIVERSITES - PRATICIENS HOSPITALIERS

### *HORS CLASSE*

Mme	CHAMBON Martine	Bactériologie Virologie
Mme	BOUVELOUP Corinne	Nutrition
Mme	FOGLI Anne	Biochimie Biologie Moléculaire
Mme	GOUAS Laëtitia	Cytologie et Histologie, Cytogénétique

### *1<sup>ère</sup> CLASSE*

M.	MORVAN Daniel	Biophysique et Traitement de l'Image
Mme	GOUMY Carole	Cytologie et Histologie, Cytogénétique
M.	MARCEAU Geoffroy	Biochimie Biologie Moléculaire
Mme	MINET-QUINARD Régine	Biochimie Biologie Moléculaire
M.	ROBIN Frédéric	Bactériologie
Mme	VERONESE Lauren	Cytologie et Histologie, Cytogénétique
M.	DELMAS Julien	Bactériologie
Mme	MIRAND Audrey	Bactériologie Virologie
M.	OUCHCHANE Lemlih	Biostatistiques, Informatique Médicale et Technologies de Communication
M.	LIBERT Frédéric	Pharmacologie Médicale
Mme	COSTE Karen	Pédiatrie
Mme	AUMERAN Claire	Hygiène Hospitalière
Mme	NOURRISSON Céline	Parasitologie – Mycologie
Mme	PONS Hanaë	Biologie et Médecine du Développement et de la Reproduction

### *2<sup>ème</sup> CLASSE*

M.	COLL Guillaume	Neurochirurgie
M.	GODET Thomas	Anesthésiologie-Réanimation et Médecine Péri-Opératoire
M.	LACHAL Jonathan	Pédopsychiatrie
M.	MOUSTAFA Farès	Médecine d'Urgence
M.	CHENAF Chouki	Pharmacologie Clinique

**MAITRES DE CONFERENCES DES UNIVERSITES  
DE MEDECINE GENERALE**

***1<sup>ère</sup> CLASSE***

Mme VAILLANT-ROUSSEL Hélène Médecine Générale

**MAITRES DE CONFERENCES DES UNIVERSITES**

***HORS CLASSE***

M. BLANCHON Loïc Biochimie Biologie Moléculaire  
M. MARCHAND Fabien Pharmacologie Médicale  
Mme VAURS-BARRIERE Catherine Biochimie Biologie Moléculaire

***CLASSE NORMALE***

M. BAILLY Jean-Luc Bactériologie Virologie  
Mme AUBEL Corinne Oncologie Moléculaire  
Mme GUILLET Christelle Nutrition Humaine  
M. BIDET Yannick Oncogénétique  
M. DALMASSO Guillaume Bactériologie  
M. PIZON Frank Santé Publique  
M. SOLER Cédric Biochimie Biologie Moléculaire  
M. GIRAUDET Fabrice Biophysique et Traitement de l'Image  
M. LOLIGNIER Stéphane Neurosciences – Neuropharmacologie  
Mme MARTEIL Gaëlle Biologie de la Reproduction  
M. PINEL Alexandre Nutrition Humaine

**MAITRES DE CONFERENCES ASSOCIES DES UNIVERSITES**

Mme ESCHALIER Bénédicte Médecine Générale  
Mme RICHARD Amélie Médecine Générale  
M. TESSIERES Frédéric Médecine Générale  
Mme ROUGE Laure Médecine Générale  
Mme BERTRAND-JARROUSSE Véronique Médecine Générale  
Mme VICARD-OLAGNE Mathilde Médecine Générale

*« La vie sans musique est tout simplement une erreur, une fatigue, un exil »*

*F. Nietzsche*

## REMERCIEMENTS ET DEDICACES

### ***Au Président du Jury de thèse, Monsieur FUTIER,***

Je vous remercie de m'avoir fait l'honneur de présider ce jury de thèse, de votre disponibilité ainsi que de votre soutien, notamment pour l'accompagnement de mes futurs projets scientifiques. Faire mon internat au sein de votre subdivision est une réelle chance que je mesure.

### ***A mes Directeurs de thèse, Messieurs PEREIRA et JABAUDON,***

Bruno, je ne te remercierai jamais assez pour toute l'énergie et le soutien dont tu as fait preuve à mon égard, et pour l'aide si précieuse prodiguée depuis plus d'un an. Beaucoup de notions nouvelles à acquérir pour ma part, mais tu as su m'accompagner, même si ce n'était pas tes supports de prédilection. J'espère que nous aurons l'occasion de réaliser de nouveaux travaux ensemble.

Matthieu, je te remercie sincèrement pour ta réelle confiance initiale et actuelle, que ce soit pour ce sujet, pour le futur projet de master, mais aussi pour m'avoir permis de mettre un pied dans l'enseignement, notamment par le biais des cours de LCA. Ce serait avec plaisir de pouvoir poursuivre avec toi dans cette voie, pour les années à venir, si tu l'acceptes.

### ***Aux membres du Jury de thèse, Mesdames GARRIDO et MOLLEVI,***

Madame GARRIDO, je vous remercie d'avoir accepté de participer à ce jury. S'ouvrir à un regard extérieur à la médecine permet justement d'avoir un œil nouveau et constructeur.

Madame MOLLEVI, je vous remercie de faire partie de mon jury. Une réelle collaboration mathématiques – médecine est nécessaire et complémentaire, ce que vous apportez ici.

A toutes les personnes qui ont participé de près ou de loin à ce travail, je voudrais les en remercier.

### ***A mes proches,***

Ma mère, je te remercie pour ton soutien indéfectible tout au long de mon existence, ta force incroyable, ton écoute et ta disponibilité. Je n'aurais jamais pu arriver aussi loin sans toi.

Ma grand-mère, tu as été l'initiatrice et la créatrice de mes premiers centres d'intérêt. Je me rappelle ces nombreux week-ends et cela fait partie de mes meilleurs souvenirs. Merci infiniment.

Mon frère, merci pour tous nos moments partagés et ton précieux soutien, notamment plus jeune.

A l'ensemble de ma famille et de ma belle-famille, merci pour tout ce que vous m'avez apporté.

Arthur et Pauline, vous avoir découverts a été une réelle chance dans la construction de mes connaissances, vous faites partie intégrante de ma réussite. Nos sous-colles gustatives n'étaient que le début de nos aventures !

A l'ensemble de ma promo' et de mes chefs qui m'apportent soutien et construction de mon parcours professionnel dans la bonne humeur.

Ambre, enfin, 11 ans qu'on s'est rencontré pour la première fois, 10 ans que tu arrives à me supporter, et finalement cette vie est nettement plus agréable à tes côtés. Je voudrais te dire merci pour ta patience et ta compréhension, notamment lors des phases de travail. Je ne pense pas que ce travail de thèse aurait été aussi abouti sans ton soutien, je t'aime !

# TABLE OF CONTENTS

<b>ARTICLE .....</b>	<b>14</b>
<b>1. INTRODUCTION .....</b>	<b>15</b>
1.1. BACKGROUND AND RATIONALE.....	15
1.2. HYPOTHESES AND OBJECTIVES .....	16
<b>2. MATERIAL AND METHODS .....</b>	<b>16</b>
2.1. STUDY SELECTION.....	16
2.2. OUTCOMES .....	17
2.3. RISK OF BIAS.....	17
2.4. STATISTICAL ANALYSIS .....	17
<b>3. RESULTS .....</b>	<b>18</b>
3.1. STUDY SELECTION AND CHARACTERISTICS .....	18
3.2. RISK OF BIAS AND DISAGREEMENTS .....	18
3.3. DEFINITIONS OF VENTILATOR-FREE DAYS .....	19
3.4. STATISTICAL METHODS TO ANALYZE VENTILATOR-FREE DAYS.....	27
3.5. SAMPLE SIZE ESTIMATION .....	27
3.5.1. <i>Expected parameters</i> .....	27
3.5.2. <i>Observed parameters</i> .....	30
3.5.3. <i>Simulation</i> .....	31
3.6. VENTILATOR-FREE DAYS: A BIMODAL DISTRIBUTION?.....	31
3.7. OTHER CHARACTERISTICS.....	32
<b>4. DISCUSSION.....</b>	<b>34</b>
4.1. HETEROGENEITY OF THE DEFINITIONS .....	34
4.2. SAMPLE SIZE ESTIMATION .....	34
4.2.1. <i>Parameters</i> .....	34
4.2.2. <i>Tests and distribution</i> .....	35
4.2.3. <i>Recent approaches</i> .....	36
4.3. STRENGTHS .....	37
4.4. LIMITATIONS .....	37
<b>CONCLUSION .....</b>	<b>39</b>

<b>ACKNOWLEDGEMENTS .....</b>	<b>40</b>
<b>REFERENCES.....</b>	<b>41</b>
<b>APPENDICES .....</b>	<b>47</b>
APPENDIX I: SEARCHING TERMS USED FOR REVIEWING .....	47
APPENDIX II: EXTRACTED DATA.....	48
APPENDIX III: STATISTICAL ANALYSIS PLAN.....	49
APPENDIX IV: PACKAGES R.....	53
APPENDIX V: SUPPLEMENT TABLES.....	54
<i>Table SI. Expected parameters. ....</i>	<i>56</i>
<i>Table SII. Sample size estimation (expected parameters).....</i>	<i>58</i>
<i>Table SIII. Observed ventilator-free days and mortality rate. ....</i>	<i>61</i>
<i>Table SIV. Sample size estimation (observed parameters), and author's conclusion. ....</i>	<i>63</i>
<i>Table SV. Sample size estimation (simulation from Bodet-Contentin).....</i>	<i>64</i>
<b>SERMMENT D'HIPPOCRATE .....</b>	<b>65</b>
<b>ABSTRACT.....</b>	<b>67</b>
<b>RESUME.....</b>	<b>68</b>

## TABLES AND FIGURES

### TABLES

Table I. Characteristics of the studies. ....	24
Table II. Definitions of ventilator-free days. ....	26
Table III. Statistical models of ventilator-free days and their distribution. ....	29

### FIGURES

Figure 1. PRISMA flow diagram. ....	19
Figure 2. Risk-of-bias summary based on RoB2. ....	25
Figure 3. Sample size estimation based on the Fine and Gray regression. ....	31
Figure 4. Distribution of ventilator-free days. ....	33

## LIST OF ABBREVIATIONS

<b>ANOVA:</b>	ANalysis Of VAriance
<b>ARDS:</b>	Acute Respiratory Distress Syndrome
<b>GAMLSS:</b>	Generalized Additive Model for Location, Scale, and Shape
<b>GLMM:</b>	Generalized Linear Mixed Model
<b>ICU:</b>	Intensive Care Unit
<b>IQR:</b>	InterQuartile Range
<b>MV:</b>	Mechanical Ventilation
<b>PMID:</b>	PubMed IDentifier
<b>PRISMA:</b>	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
<b>RCT:</b>	Randomized Controlled Trials
<b>REDCap:</b>	Research Electronic Data Capture
<b>RoB2:</b>	Cochrane Risk of Bias tool version 2
<b>VFD:</b>	Ventilator-Free Days



## ARTICLE

### VENTILATOR-FREE DAYS AS AN OUTCOME: APPROACHES ON SAMPLE SIZE ESTIMATION FROM A SYSTEMATIC REVIEW.

Laurent Renard-Triché<sup>1</sup>, Emmanuel Futier<sup>1,2</sup>, Manuela De Carvalho<sup>3</sup>, Nathalie Pinol-Domenech<sup>3</sup>, Laëtitia Bodet-Contentin<sup>4</sup>, Matthieu Jabaudon<sup>1,2</sup>, Bruno Pereira<sup>5</sup>.

<sup>1</sup> Department of Perioperative Medicine, CHU Clermont-Ferrand, Clermont-Ferrand, France;

<sup>2</sup> iGReD, CNRS, INSERM, Université Clermont Auvergne, Clermont-Ferrand, France;

<sup>3</sup> Université Clermont Auvergne, Health Library, Clermont-Ferrand, France;

<sup>4</sup> Medical Intensive Care Unit, INSERM CIC 1415, CRICS-TRIGGERSep Network, CHRU de Tours and Université de Tours, Tours, France.

<sup>5</sup> Biostatistics Unit, Department of Clinical Research, and Innovation (DRCI), CHU Clermont-Ferrand, Clermont-Ferrand, France;

#### **CRedit author statement:**

**LRT:** Conceptualization, Methodology, Formal analysis, Investigation, Writing – Original Draft, Visualization; **EF:** Writing – Review & Editing, Supervision; **MDC:** Investigation, Resources, Writing – Review & Editing; **NPD:** Investigation, Resources, Writing – Review & Editing; **LBC:** Resources; **MJ:** Conceptualization, Methodology, Validation, Writing – Review & Editing, Supervision; **BP:** Conceptualization, Methodology, Validation, Formal analysis, Writing – Review & Editing, Supervision.

**Source support:** this work was supported by grants from the French Agence Nationale de la Recherche and Direction Générale de l'Offre de Soins ("Programmes de Recherche Translationnelle en Santé" ANR-13- PRTS-0010 and ANR-20-CE17-0015; "Programme Hospitalier de Recherche Clinique » 2018), and the "Fonds Européen de Développement Régional" (European Union, FEDER Auvergne) (M. Jabaudon).

This article will be submitted for publication after thesis defense.

# 1. INTRODUCTION

## 1.1. Background and rationale

Between a quarter and half of the patients admitted to the intensive care unit will present with respiratory failure, requiring invasive mechanical ventilation (MV) (1). These patients are at risk of several complications, such as infection and death, with related health-care costs (2,3).

Mortality is a strong primary endpoint that has long been used in several studies (4). However, since the improvement of therapeutics, mortality has decreased (5). Hence, the sample size needed to show a difference has also increased, and so, most published randomized clinical trials (RCT) that aim to reduce mortality have produced negative results (6,7). For this reason, other outcomes, such as ventilator-free days (VFD), have been developed.

Ventilator-free days are an outcome measure increasingly used in critical care (8). First proposed in 1994 (9), it was developed in studies focusing on acute respiratory distress syndrome (ARDS). The number of VFD is defined as the number of days from the last day of mechanical ventilation to day 28. If a patient dies during the first 28 days, their number of VFD is equal to zero. This composite outcome measure (i.e., combining survival and the duration of ventilation) is more appropriate than only the duration of ventilation because the latter disregards the mortality rate (10).

Determining the sample size can be imprecise and difficult and represents a trade-off between cost effectiveness and statistical power. Calculating this sample size involves the employment of formulae designed to obtain significant results in studies that compare several groups based on the primary endpoint. The test chosen to analyze the primary endpoint will depend on its distribution (11). The number of VFD does not follow a Gaussian distribution (12); therefore, we cannot use parametric tests. Indeed, it is a skewed distribution with inflations, especially 0s (i.e., zero-inflation), and represents a rather time-dependent event. Hence, initially, Shoenfeld et al.

proposed using the Wilcoxon test rather than Student's t-test (10). However, to use the Wilcoxon test, there should be no need to adjust to covariates or to have interim monitoring, which is quite rare in RCT. In their last review, Yehya et al. (8) recommended using competing risk regression for the VFD analysis, such as the Fine and Gray competing risk regression (13), which considers extubation as the event of interest and death as the competing risk.

## **1.2. Hypotheses and objectives**

We conducted a systematic review of RCT using VFD as their primary outcome. Our principal objective was to investigate whether the statistical analysis of RCT evaluating VFD as the primary outcome was congruent with the distribution of this measure. Our secondary objective was to review the definitions of VFD and tried to draw guidelines for future trials using an appropriate methodology for the analysis of VFD.

## **2. MATERIAL AND METHODS**

Our study protocol was registered on PROSPERO in December 2021 (ID: CRD42021282304).

### **2.1. Study selection**

The search strategies were developed with the help of two university librarians (MdC and NPD). We searched through two databases: MEDLINE via the PubMed and Embase. The search strategies used a combination of keywords related to VFD, the names of journals with a 2021 impact factor greater than 5, and study type, from inception to December 2021. The complete list of search terms is available in the Appendix I. Two investigators (LRT and MJ) independently screened the titles and abstracts of the search results. The full text of all potentially eligible studies was retrieved and reviewed for eligibility. Any disagreements between reviewers were resolved by consensus or by a third member (BP). First, we removed the duplicates between the databases. Then, we

excluded all trials that were not RCT in adults and those with a primary endpoint that was not VFD. A narrative synthesis supporting the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram was included as part of this systematic review (14).

The references for the selected studies were imported from Zotero for review, and the data were extracted independently by three investigators (LRT and, MJ for the first half and BP for the second half) and collected into standardized, pilot-tested forms using the Research Electronic Data Capture (REDCap) tools hosted at CHU Clermont-Ferrand (15,16). The data were extracted from the text when possible; otherwise, we used the supplements if they were available. The extracted data were cross-checked, and any disagreements were resolved by discussion, with the involvement of a third reviewer, if required. For each selected article, we recorded several items, as detailed in the Appendix II.

## **2.2. Outcomes**

Our primary outcome was the sample size, as estimated in RCT evaluating VFD as the primary outcome. To assess this, we extracted the sample sizes estimated and observed of the selected RCT, and simulated other sample sizes through alternative statistical approaches. Our secondary outcomes were to review the definitions of VFD, mortality rates, statistical methods, and statistical distribution of VFD in each trial.

## **2.3. Risk of bias**

To assess the risk of bias in each study and across studies, we used Version 2 of the Cochrane risk of bias tool (RoB2) for RCT (17). The studies were assessed using five fixed domains, as outlined in RoB2. Each of these studies was classified by two reviewers (LRT and MJ) as low-concern, high-concern, or with some concerns.

## **2.4. Statistical analysis**

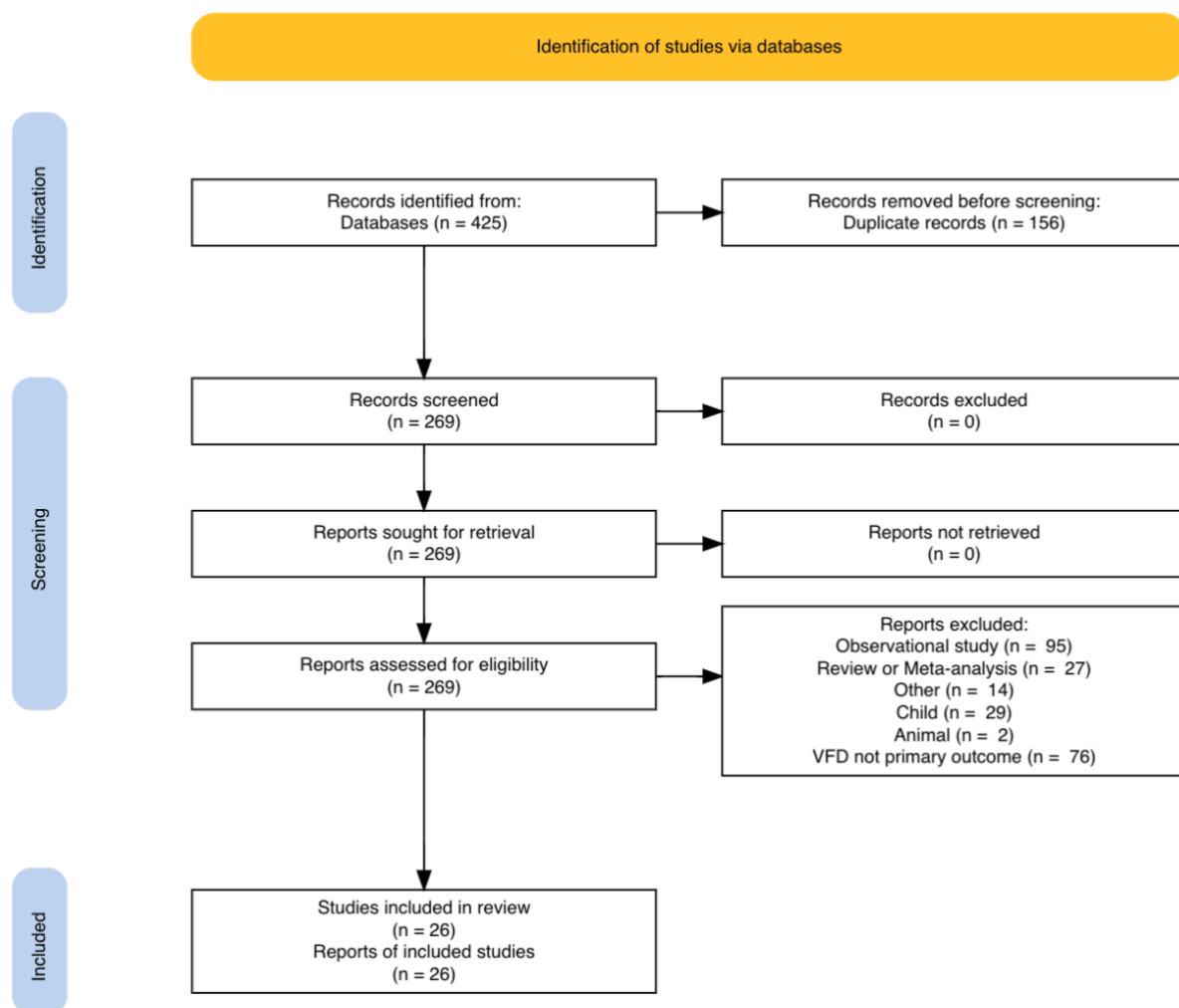
The different statistical approaches were as follows: the Student *t*-test and the Mann–Whitney–Wilcoxon rank-sum test since these are standard tests used in several studies; the Mann–Whitney–Wilcoxon rank-sum test using the Noether formula to compare if the result differ from the previous one; the Cox regression since the VFD are a time-dependent event; the zero-inflated negative binomial negative since the VFD involve a zero-inflation; and finally, the Fine and Gray regression since the VFD involve a competitive risk. The corresponding formulae are available in the Appendix III. To evaluate the distribution of VFD, we used Shapiro-Wilk normality test and also Hartigan’s dip test (18) to assess the probability of VFD following a unimodal distribution. All statistical analyses were performed using R Core Version 4.2 (19). All packages used are listed in the Appendix IV.

### **3. RESULTS**

#### **3.1. Study selection and characteristics**

We identified 425 studies from 2004 to 2021. After removing duplicates, we assessed 269 studies. First, we excluded 136 studies that were not randomized. Second, we excluded 2 animal studies and 29 child studies. Third, we excluded 76 studies in which the primary outcome was not the VFD. In the end, 26 studies were included in our systematic review (20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45) (Figure 1). These are mainly superiority studies (92%) comparing two groups (92%) with several centers (89%) and included interim analysis in most cases (58%). A significant proportion of the selected studies had to be stopped early (38%). Finally, the patients included were heterogeneous between the studies: a third of the target population were patients with ARDS, and the others had other different syndromes (Table I).

#### **3.2. Risk of bias and disagreements**









**Figure 1. PRISMA flow diagram from search in December 2021.**






*This diagram was generated with the R package DiagrammeR (14).*

Using the Excel spreadsheet provided by the RoB2 tool, we assessed the risk of bias, as synthesized in the Figure 2. 73% of the studies assessed found some concerns, mainly for the randomization process and the deviations from the intended interventions.

Among all the collected items, the median number of disagreements between the two reviewers was 1 (IQR 0–2) out of the 26 selected studies, for a total of 23 disagreements out of 769 items (3%). All disagreements were resolved by consensus. In addition, the opinion of the third reviewer was not necessary.








### 3.3. Definitions of ventilator-free days

Authors	Title	Year	Journal	Group – ratio (E/C)	Study type	Centre	Interim analysis	Early stop	Population	Overall RoB2
Morris et al.	A double-blind placebo-controlled study to evaluate the safety and efficacy of L-2-oxothiazolidine-4-carboxylic acid in the treatment of patients with acute respiratory distress syndrome. (20)	2008	CCM	2 – 1:1	Sup	Multi.	Yes	Yes	ARDS	
Paine et al.	A randomized Trial of recombinant human granulocyte-macrophage colony stimulating factor for patients with acute lung injury. (21)	2012	CCM	2 – 1:1	Sup	Multi.	No	Yes	ALI or ARDS	
Mackle et al.	Conservative oxygen therapy during mechanical ventilation in the ICU. (22)	2020	NEJM	2 – 1:1	Sup	Multi.	Yes	No	MV	
Villar et al.	Dexamethasone treatment for the acute respiratory distress syndrome: a multicentre, randomised controlled trial. (23)	2020	LRM	2 – 1:1	Sup	Multi.	Yes	Yes	ARDS	
Zhou et al.	Early application of airway pressure release ventilation may reduce the duration of mechanical ventilation in acute respiratory distress syndrome. (24)	2017	ICM	2 – 1:1	Sup	Mono.	No	No	ARDS	
Trouillet et al.	Early percutaneous tracheotomy versus prolonged intubation of mechanically ventilated patients after cardiac surgery: A Randomized Trial. (25)	2011	AIM	2 – 1:1	Sup	Mono.	No	No	Cardiac surgery + MV	

Simonis et al.	Effect of a Low vs Intermediate Tidal Volume Strategy on Ventilator-Free Days in Intensive Care Unit Patients Without ARDS: A Randomized Clinical Trial. (26)	2018	JAMA	2 – 1:1	Sup	Multi.	No	No	No ARDS + MV	
Algera et al.	Effect of a lower vs higher positive end-expiratory pressure strategy on ventilator-free days in ICU patients without ARDS: A randomized clinical trial. (27)	2020	JAMA	2 – 1:1	N-Inf	Multi.	No	No	No ARDS + MV	
Tomazini et al.	Effect of Dexamethasone on Days Alive and Ventilator-Free in Patients with Moderate or Severe Acute Respiratory Distress Syndrome and COVID-19: The CoDEX Randomized Clinical Trial. (28)	2020	JAMA	2 – 1:1	Sup	Multi.	Yes	Yes	ARDS + COVID-19	
Grieco et al.	Effect of Helmet Noninvasive Ventilation vs High-Flow Nasal Oxygen on Days Free of Respiratory Support in Patients with COVID-19 and Moderate to Severe Hypoxemic Respiratory Failure: The HENIVOT Randomized Clinical Trial. (29)	2021	JAMA	2 – 1:1	Sup	Multi.	No	No	ARF + COVID-19	
van Meenen et al.	Effect of On-Demand vs Routine Nebulization of Acetylcysteine With Salbutamol on Ventilator-Free Days in Intensive Care Unit Patients Receiving Invasive Ventilation: A Randomized Clinical Trial. (30)	2018	JAMA	2 – 1:1	N-Inf	Multi.	Yes	No	MV	



Spragg et al.	Effect of recombinant surfactant protein C-based surfactant on the acute respiratory distress syndrome. (31)	2004	NEJM	2 – 1:1	Sup	Multi.	No	No	ARDS	!
Welte et al.	Efficacy and safety of trimodulin, a novel polyclonal antibody preparation, in patients with severe community-acquired pneumonia: a randomized, placebo-controlled, double-blind, multicenter, phase II trial (CIGMA study). (32)	2018	ICM	2 – 1:1	Sup	Multi.	Yes	No	Community-acquired pneumonia	!
Rice et al.	Enteral omega-3 fatty acid, $\gamma$ -linolenic acid, and antioxidant supplementation in acute lung injury. (33)	2011	JAMA	2 – 1:1	Sup	Multi.	Yes	Yes	ALI	!
Bernard et al.	Evaluating the efficacy and safety of two doses of the polyclonal anti-tumor necrosis factor- $\alpha$ Fragment antibody AZD9773 in Adult patients with severe sepsis and/or septic shock: Randomized, double-blind, placebo-controlled phase IIb Study. (34)	2014	CCM	3 – 1:1:1	Sup	Multi.	No	No	Sepsis	!
Chung et al.	High-frequency percussive ventilation and low tidal volume ventilation in burns: A randomized controlled trial. (35)	2010	CCM	2 – 1:1	Sup	Multi.	Yes	Yes	Burns	!
Rice et al.	Initial trophic vs full enteral feeding in patients with acute lung injury: The EDEN randomized trial. (36)	2012	JAMA	2 – 1:1	Sup	Multi.	Yes	No	ALI	+

Bein et al.	Lower tidal volume strategy (~3 ml/kg) combined with extracorporeal CO <sub>2</sub> removal versus 'conventional' protective ventilation (6 ml/kg) in severe ARDS: the prospective randomized Xtravent-study. (37)	2013	ICM	2 – 1:1	Sup	Multi.	Yes	Yes	ARDS	
Hodgson et al.	Maximal Recruitment Open Lung Ventilation in Acute Respiratory Distress Syndrome (PHARLAP). A Phase II, Multicenter Randomized Controlled Clinical Trial. (38)	2019	AJRCCM	2 – 1:1	Sup	Multi.	Yes	Yes	ARDS	
Kacmarek et al.	Neurally adjusted ventilatory assist in acute respiratory failure: a randomized controlled trial. (39)	2020	ICM	2 – 1:1	Sup	Multi.	No	No	ARF	
Kacmarek et al.	Partial liquid ventilation in adult patients with acute respiratory distress syndrome. (40)	2006	AJRCCM	3 – 1:1:1	Sup	Multi.	Yes	No	ARDS	
Liu et al.	Randomized clinical trial of activated protein C for the treatment of acute lung injury. (41)	2008	AJRCCM	2 – 1:1	Sup	Multi.	Yes	No	ALI	
Rice et al.	Randomized trial of initial trophic versus full-energy enteral nutrition in mechanically ventilated patients with acute respiratory failure. (42)	2011	CCM	2 – 1:1	Sup	Multi.	Yes	No	ARF	
Matthay et al.	Randomized, placebo-controlled clinical trial of an aerosolized $\beta_2$ -agonist for treatment of acute lung injury. (43)	2011	AJRCCM	2 – 1:1	Sup	Multi.	Yes	Yes	ALI	

Bennett-Guerrero et al.	Severe Acute Respiratory Syndrome Coronavirus 2 Convalescent Plasma Versus Standard Plasma in Coronavirus Disease 2019 Infected Hospitalized Patients in New York: A Double-Blind Randomized Trial. (44)	2021	CCM	2 – 4:1	Sup	Mono.	No	Yes	COVID-19	
McAuley et al.	Simvastatin in the acute respiratory distress syndrome. (45)	2014	NEJM	2 – 1:1	Sup	Multi.	Yes	No	ARDS	

**Table 1. Characteristics of the studies: the authors and the title of the study, the year of publication and the journal, the number of group and the allocation ratio, the number of centres, if there are an interim analysis and if the study stop early, the population, and the overall risk of bias.**

AIM: *Annals of internal medicine*; AJRCCM: *American Journal of Respiratory and Critical Care Medicine*; CCM: *Critical Care Medicine*; ICM: *Intensive Care Medicine*; JAMA: *Journal of the American Medical Association*; LRM: *The Lancet Respiratory Medicine*; NEJM: *The New England Journal of Medicine*.  
 E: *experimental group*. C: *control group*.  
 Sup: *superiority*. N-Inf: *non-inferiority*.  
 All: *acute lung injury*; ARDS: *acute respiratory distress syndrome*; ARF: *acute respiratory failure*; ICU: *intensive care unit*; MV: *mechanical ventilation*.  
 RoB2: *Revised Cochrane risk-of bias tools for randomized trials*; low risk; some concerns; high risk.

Study ID	D1	D2	D3	D4	D5	Overall
Morris et al., 2008 (20)	+	+	+	+	!	!
Paine et al., 2012 (21)	+	+	+	+	!	!
Mackle et al., 2020 (22)	+	-	+	+	+	-
Villar et al., 2020 (23)	+	+	+	+	+	+
Zhou et al., 2017 (24)	!	+	+	+	!	!
Trouillet et al., 2011 (25)	!	+	+	+	!	!
Simonis et al., 2018 (26)	+	+	+	+	+	+
Algera et al., 2020 (27)	+	!	+	+	+	!
Tomazini et al., 2020 (28)	!	-	+	+	+	-
Grieco et al., 2021 (29)	!	+	+	+	+	!
van Meenen et al., 2018 (30)	+	-	+	+	+	-
Spragg et al., 2004 (31)	!	+	+	+	+	!
Welte et al., 2018 (32)	!	+	+	+	+	!
Rice et al., 2011 (33)	!	+	+	+	+	!
Bernard et al., 2014 (34)	!	+	+	+	+	!
Chung et al., 2010 (35)	+	!	+	+	+	!
Rice et al., 2012 (36)	+	+	+	+	+	+
Bein et al., 2013 (37)	+	!	+	+	+	!
Hodgson et al., 2019 (38)	+	-	+	+	!	-
Kacmarek et al., 2020 (39)	+	!	+	+	+	!
Kacmarek et al., 2006 (40)	+	!	+	+	+	!
Liu et al., 2008 (41)	!	+	+	+	+	!
Rice et al., 2011 (42)	+	!	+	+	+	!
Matthay et al., 2011 (43)	!	+	+	+	+	!
Bennett-Guerrero et al., 2021 (44)	!	+	+	+	+	!
McAuley et al., 2014 (45)	+	!	+	+	+	!

**Figure 2. Risk-of-bias summary: review authors' judgments about each risk of bias item for each included study based on the Revised Cochrane risk-of-bias tools for randomized trials (RoB2) (17).**

D: domain; D1: randomisation process; D2: Deviations from the intended interventions; D3: Missing outcome data; D4: Measurement of the outcome; D5: Selection of the reported result.

⊕ low risk; ⊕! some concerns; ⊖ high risk.

In Table II, we report the definitions of VFD in the considered studies. Almost all studies counted whole days without support ventilation (92%) and collected the VFD at day 28. In the rest

Variable	Proportion: n/N (%)
<b>Onset</b>	
no limit	9/26 (35)
> 4h	1/26 (4)
> 24h	3/26 (11)
> 48h	10/26 (38)
> 72h	1/26 (4)
portion with no limit	2/26 (8)
<b>Respiratory support</b>	
invasive ventilation	13/26 (50)
tracheostomy	12/26 (46)
ECMO	2/26 (8)
NIV	6/26 (23)
HFNO	1/26 (4)
no specification	13/26 (50)
<b>Reintubation</b>	
possible	7/26 (27)
reset	9/26 (35)
no specification	10/26 (38)
<b>Day</b>	
28	24/26 (92)
30	1/26 (4)
60	1/26 (4)

**Table II. Definitions of ventilator-free days.**

*ECMO: extracorporeal membrane oxygenation; NIV: non-invasive ventilation; HFNO: high-flow nasal oxygen.  
n: number of events; N: total number of studies.*

of the definitions, there was heterogeneity. The onset (i.e., the beginning of the period without support ventilation) was not the same across the studies: 35% considered the onset at extubation, and 38% at 48 hours after extubation. If a patient was intubated again after a period of extubation, 35% of the studies reset the VFD (i.e., considered the VFD equal to 0), and 27% of the studies

counted the different periods during which the patient was extubated; the remaining studies did not specify this point. Finally, half of the studies did not specify the type of respiratory support counted as not free of ventilation.

### **3.4. Statistical methods to analyze ventilator-free days**

In Table III, we report the statistical analysis methods used to assess the VFD, along with the expected and observed effect sizes. About one-third used the Mann–Whitney–Wilcoxon rank-sum test, one-third did not specify the test used, and a little less than a third used parametric tests. A minority used complex models, such as generalized linear models and generalized additive models, for location, scale, and shape. Furthermore, the main effect size used in the results of the studies was the absolute mean difference (69%).

### **3.5. Sample size estimation**

As a first step, we extracted the estimated and the observed sample size reported in the selected studies. We subsequently calculated the sample size with the parameters (e.g.,  $\alpha$  risk, power, mean difference) reported in two ways: using the expected parameters displayed in the Material and Methods section or the observed parameters displayed in the Results section.

The proper estimation of sample size necessitates the correct estimation of the VFD distribution. Thus, in Table III, we report the VFD distributions defined by the authors. Half of the studies did not explicitly state the type of assumed distribution, while the other half did not consider the VFD as normally distributed. A more precise description was available for some studies, with 8% assuming a zero-inflated  $\beta$  binomial distribution and 4% assuming a bimodal distribution. However, two recently published studies (29,32) considered the number of VFD as normally distributed.

#### **3.5.1. Expected parameters**

Study	VFD distribution	Correction	Statistical model in Methods section	Statistical model in Results section	Effect size
Morris et al., 2008 (20)	NA	NA	ANOVA	NA	Absolute mean difference
Paine et al., 2012 (21)	Asymmetric	No correction	NA	Wilcoxon	Absolute difference
Mackle et al., 2020 (22)	Not normal	+ 15%	Wilcoxon	Quantile	Absolute mean difference
Villar et al., 2020 (23)	Asymmetric	Markov chain	Wilcoxon	NA	Absolute difference
Zhou et al., 2017 (24)	Not normal	No correction	NA	Kruskal-Wallis	Absolute mean difference
Trouillet et al., 2011 (25)	NA	NA	NA	NA	Absolute mean difference
Simonis et al., 2018 (26)	NA	NA	Student	Student	Absolute difference
Algera et al., 2020 (27)	Zero inflated $\beta$	Use specific model	GAMLSS	GAMLSS	Absolute mean difference / Mean ratio
Tomazini et al., 2020 (28)	Zero inflated $\beta$ binomial	+ 15% & Use specific model	GLM	GLM	Absolute mean difference
Grieco et al., 2021 (29)	Normal	NA	NA	NA	Relative mean difference
van Meenen et al., 2018 (30)	NA	NA	Wilcoxon	Hodges-Lehmann	Absolute median difference
Spragg et al., 2004 (31)	Asymmetric	Nonparametric test	Cox	NA	Absolute mean difference
Welte et al., 2018 (32)	Normal	NA	Wilcoxon	Wilcoxon	Absolute difference
Rice et al., 2011 (33)	NA	NA	ANOVA	NA	Absolute mean difference

Bernard et al., 2014 (34)	Not normal	Nonparametric test	ANOVA	NA	Absolute mean difference
Chung et al., 2010 (35)	NA	NA	NA	NA	Absolute mean difference
Rice et al., 2012 (36)	NA	NA	ANOVA	NA	Absolute mean difference
Bein et al., 2013 (37)	NA	NA	NA	NA	Absolute mean difference
Hodgson et al., 2019 (38)	Not normal	+ 15%	NA	NA	Relative SD difference
Kacmarek et al., 2020 (39)	NA	NA	Wilcoxon + Bootstrap	NA	Absolute mean difference
Kacmarek et al., 2006 (40)	NA	NA	GLM	NA	Absolute difference
Liu et al., 2008 (41)	Not normal	Nonparametric test	Wilcoxon	NA	Absolute mean difference
Rice et al., 2011 (42)	NA	NA	Student	Wilcoxon	Relative difference
Matthay et al., 2011 (43)	NA	NA	ANOVA	NA	Absolute mean difference
Bennett-Guerrero et al., 2021 (44)	NA	NA	Wilcoxon	Wilcoxon	Absolute mean difference
McAuley et al., 2014 (45)	Bimodal	Median comparison	Student + Bootstrap	Student + Bootstrap	Absolute mean difference

**Table III. Statistical models used for sample size estimation (Material and Methods section) and for analysis (Results section) of ventilator-free days and their distribution.**

ANOVA: analysis of variance; GAMLSS: generalized additive model for location scale and shape; GLM: generalized linear model; NA: not available; SD: standard deviation.

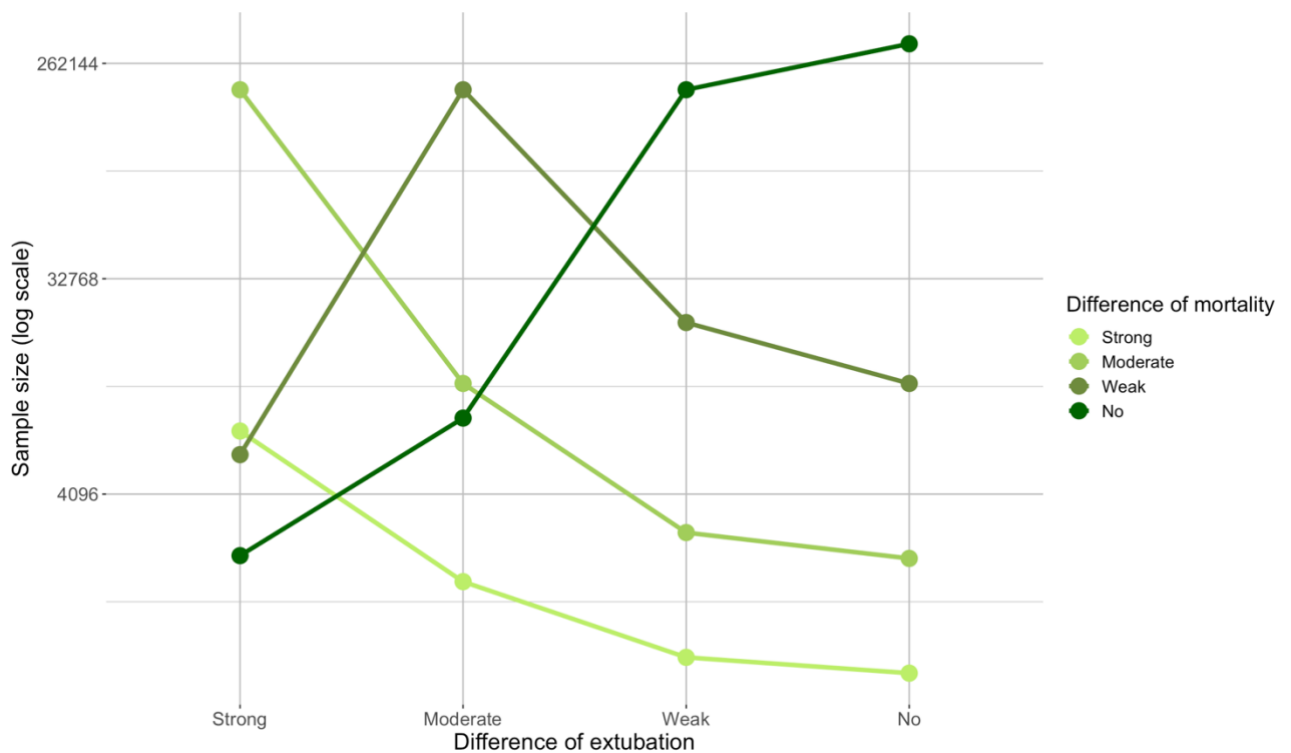


In the supplement Table SI (Appendix V), we report the expected parameters proposed by the authors for the sample size estimation in the Material and Methods sections of the selected studies. The absolute mean difference ranged from 0.5 to 7.0. In one non-inferiority study (27) the authors considered 1.6 as the non-inferiority margin, while in one superiority study (26) the authors considered 1 as the superiority margin. The standard deviation was not always reported (23%), but when it was available, it was high (median 10.0, IQR 6.8–10.5). Mortality was considered only in one study (23), who used Markov chains in considering the probabilities of death, getting off ventilation alive, and receiving ventilation. Finally, the expected drop-out rate expected was quite different between studies (0%–25%).

Using these parameters, we calculated the different sample sizes depending on different tests: the Student *t*-test, the Mann–Whitney–Wilcoxon rank-sum test using the Noether formula or not, the Cox regression, the zero-inflated negative binomial regression, and the Fine and Gray regression (supplement Table SII, Appendix V). Several tests could not be estimated because of the lack of expected parameters, such as the VFD in the control group and the standard deviation. Moreover, it was not possible to estimate the sample size using the Fine and Gray regression because we did not have the probability of extubation or the incidence of mortality. Consequently, we simulated different incidences of extubation and mortality (Figure 3). The size of the difference in the incidence of extubation and the incidence mortality both had a significant impact on the sample size. For the estimation using the Cox and zero-inflated negative binomial regressions, in most cases, the sample size was greater than for the other models.

### 3.5.2. Observed parameters

In the same way as above, we reported the observed parameters needed to estimate the sample size (supplement Table SIII, Appendix V). The standard deviations were almost the same as in the supplement Table SI (Appendix V) (median 11.2, IQR 9.6–16.3), without missing parameters. Furthermore, the dropout rate observed was very low (0%–2%).



**Figure 3. Sample size estimation based on the Fine and Gray regression.**

*Absolute difference of extubation incidence (experimental/control group): strong 0.60/0.45 – moderate 0.60/0.51 – weak 0.60/0.56 – no difference 0.60/0.60.*

*Absolute difference of mortality incidence (experimental/control group): strong 0.20/0.30 – moderate 0.20/0.25 – weak 0.20/0.21 – no difference 0.20/0.20.*

Using these parameters, we calculated the different sample sizes using the same tests as above (supplement Table SIV, Appendix V). Sample sizes could not be estimated using the Fine and Gray regression model because for some studies the VFD and incidence of mortality had different timeframes, and we did not have the incidence of extubation for any study. For about half of the studies, the sample sizes were much larger than those estimated, mainly because the observed mean difference was sometimes very low.

### 3.5.3. Simulation

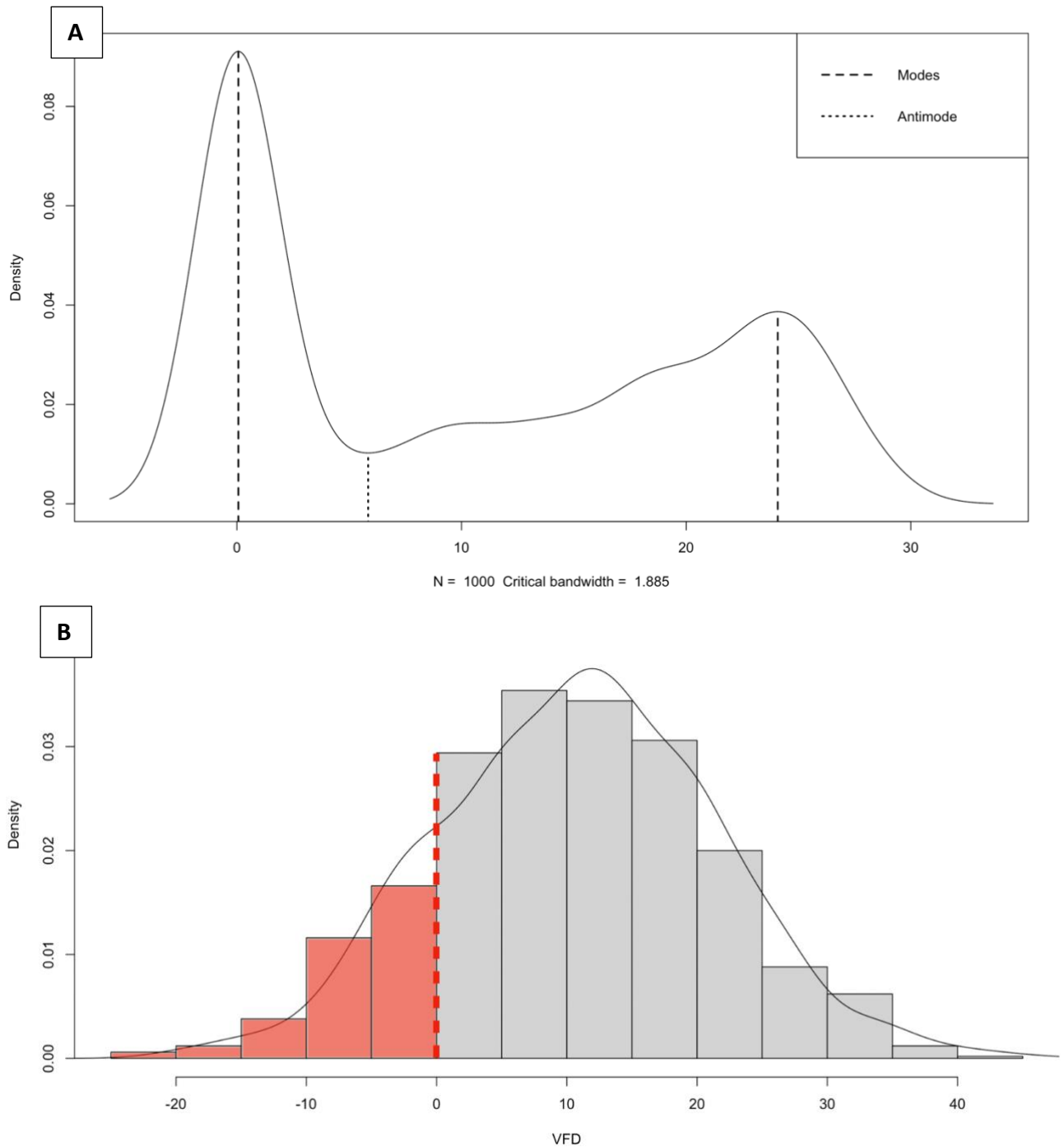
Finally, we estimated the sample size using the simulation carried out by Bodet-Contentin et al. (46) and the same tests as above (supplement Table SV, Appendix V). Only the estimation using the Fine and Gray regression model provided a realistic sample size.

## 3.6. Ventilator-free days: a bimodal distribution?

We used different tests to estimate the sample size, including the zero-inflated negative binomial regression. For this model, we need to know the overdispersion of the distribution of the data. To this end, we used two datasets: the simulation carried out by Bodet-Contentin et al. (46) and the data from the meta-analysis by Jabaudon et al. (47). In both studies, the distribution of VFD looked more like a bimodal distribution than a negative binomial distribution. To prove this, we used first the Shapiro-Wilk normality test, rejecting the normality for the two data frames ( $W = 0.8148$ ,  $p\text{-value} < 0.0001$  for the simulation and  $W = 0.8163$ ,  $p\text{-value} < 0.0001$  for the meta-analysis). Second, we used Hartigan's' dip test, rejecting the unimodality for the two data frames ( $D = 0.0826$ ,  $p\text{-value} < 0.0001$  for the simulation and  $D = 0.1007$ ,  $p\text{-value} < 0.0001$  for the meta-analysis). The first mode was 0 and the second was around 24 for both the simulation (Figure 4A) and the meta-analysis (not shown). We modelled a normal distribution with the same parameters as in the simulation (Figure 4B), since some studies assumed this type of distribution for the VFD, to visually compare the unsuitability of this model to estimate the distribution of the VFD.

### **3.7. Other characteristics**

As exploratory endpoints, we were also interested in other statistical parameters. The power ranged from 80% to 90%, the  $\alpha$  risk ranged from 2.5% to 10% (for one study (34)), and the most common  $\alpha$  risk chosen was 5% (supplement Table SI, Appendix V). As shown in the supplement Table SI (Appendix V), the reported  $\alpha$  risk was one-sided in four studies (two were a non-inferiority studies (27,30), while the other two were superiority studies (32,34)). The most used frame for mortality was 28 days corresponding to the same frame as the VFD. The frame was reported in days in 92% of the studies, the remaining two used hospital or intensive care unit mortality. Finally, the incidence was quite different across the studies but similar within studies, with a median of 24.9 days (IQR 19.2–30.9) for the control group and a median of 24.3 days (IQR 19.8–30.7) for the



**Figure 4. Distribution of ventilator-free days.**

Histograms representing the Bodet-Contentin (46) simulated distribution of ventilator free days (VFD) (A) (mean of 10.6 days with two modes) and the corresponding Gaussian distribution (B) (mean of 10.6 days, standard deviation of 10.5) used in some studies for sample size estimation. The red bars correspond to the theoretical data that should be seen if the distribution was normal.

We represented the densities and not the frequencies corresponding to the average number of VFD for a given class.

experimental group (supplement Table SIII). A significant result was obtained in only a few studies (35%), but when a complex model was used, the result was significant (supplement Table SIV, Appendix V).

## 4. DISCUSSION

To our knowledge, this is the first systematic review to report on the sample size estimation and the parameters needed for VFD. The number of VFD provides greater statistical power for detecting a treatment effect than the binary outcome measure of mortality (8,10). However, there are several limitations (discussed in the previous section) to the use of this criterion; hence, we suggest several adjustments.

### 4.1. Heterogeneity of the definitions

Following Contentin et al. (48) and Yehya et al. (8), we found important differences regarding the definitions of VFD and the methods for the analysis, such as the effect size and the test used, thus making it difficult to conduct meta-analyses. We did not detect significant changes in the definitions and methods since the publication of these two reviews, which might be because they were recently published.

Several recommendations were made by Yehya et al., including on how to define VFD explicitly: the start time, the time frame, the time of successful extubation, extubation interval, non-invasive support, tracheostomy, and value for the extubated decedent. Hence, we should use a core outcome set, such as the Core Outcomes in Ventilation Trials (49), which included standardized definitions and measures, particularly for extubation, reintubation, duration of mechanical ventilation, and mortality.

### 4.2. Sample size estimation

#### 4.2.1. Parameters

We need several parameters to estimate the sample size. First, we need the tail of the  $\alpha$  risk (one or two). In most of the superiority studies considered, the authors used a two-tailed test.

However, if we want to show a specific and unidirectional difference (e.g., treatment versus placebo), a one-sided  $\alpha$  risk should be used to test the null hypothesis (H0) for the two groups (50).

We also need the margin (i.e., the mean difference) to show the difference between two groups in a superiority study or the loss of efficacy in a non-inferiority study. Some studies used a superiority margin smaller than the non-inferiority margin, which is hardly justifiable. We advocate for a clear consensus on these margins according to the context.

Finally, we need the effect size more than the p-value, which provides information only on the statistical significance of the dichotomy and can be misinterpreted (51,52). Indeed, the effect size corresponds to the magnitude of the effect, and its precision is given by the confidence interval, which is also useful for a meta-analysis. Most of the time, the studies considered standardized differences (such as Cohen's *d*) (53), which imply a normal distribution. Furthermore, some studies used the absolute mean difference despite the use of the Mann–Whitney–Wilcoxon rank-sum test when they should use the absolute median difference.

#### 4.2.2. Tests and distribution

There are two concepts to consider when describing VFD: the presence of competitive risks and the distribution required in complex models.

First, regarding the presence of competitive risks, Yehya et al. (8) recommended reporting the mortality, since the number of VFD combines mortality and ventilator duration. Because mortality is a competitive event of extubation, competing risk regression using the Fine and Gray regression is more appropriate than usual tests such as Student *t*-test and Mann–Whitney–Wilcoxon rank-sum test (13). Moreover, this test allows for the adjustment of the covariate and stratification variables and interim monitoring, which are common in RCT. For this model, it is recommended to report the overall effect size (subdistribution hazard ratio) and the components of this composite outcome—that is, the incidence of mortality and the incidence of extubation—which are useful for sample size estimation. Nevertheless, the cumulative incidence function provided by the Fine and

Gray regression does not have a natural interpretation (54,55) and it does not take the zero inflation into account; but, it is the best way to distinguish between extubation and mortality.

Second, the distribution of the VFD renders the statistical analysis more arduous. On one hand, using the Bodet-Contentin simulation (46), we obtained a bimodal distribution that is consistent with those in other studies: a type U distribution according to Galtung's classification (56). In contrast, we may categorize the VFD into several classes, such as the three groups of weaning (short weaning, difficult weaning, and prolonged weaning) (12). These distributions require mixture models corresponding to a mixture of two or more distributions, models available in the *mixtools* R package (57). On the other hand, some articles reported a beta distribution (27,28). For this distribution, we can use the generalized linear mixed model (GLMM) or the generalized additive model for location, scale, and shape (GAMLSS) (27)). Moreover, using complex models allow for adjusting covariates even in RCT, which may reduce the sample size and increase the power (58). Therefore, there is heterogeneity in the distributions and tests used, which warrants further investigation.

In conclusion, the use of parametric tests and the Mann–Whitney–Wilcoxon rank-sum test should be discouraged for VFD. We advocate for the use of complex models considering competitive risks, the distribution of VFD, and any covariates.

#### 4.2.3. Recent approaches

To estimate the sample size, we used commonly applied formulae (11). In contrast, some authors have recently explored simulations. Indeed, for complex models, such as the GLMM and GAMLSS, these formulae cannot be employed; hence, simulation-based power analyses were developed (59). Instead of the frequentist approach with null hypothesis significance testing, Bayesian approaches are increasingly used to simulate the sample size (60) without arbitrarily fixing the power (61).

On the other hand, the Markov chain model is an interesting approach. This stochastic model describes a sequence of possible events in which the probability of each event depends only on the state attained in the previous event (62). For VFD, there are three possible states: intubated, extubated, or dead. Based on this model, there is the Monte Carlo simulation method can be used to estimate the sample size (63). This type of simulation allows the modelling of a range of intervention effects and the calculation of the corresponding sample sizes required to detect the significance at a specific power (e.g., 80%).

In conclusion, we discourage the use of simplistic formulae but instead advocate for simulations to estimate the sample size, especially for VFD, given their complex probability distribution, in the continuity of Yehya et al. (8) who also applied simulations.

### **4.3. Strengths**

We followed all the guidelines from the PRISMA 2020 statement (64). Two authors independently reviewed the studies, and in case of a disagreement, a third author intervened. We used two databases (PubMed and Embase), to be as comprehensive as possible. In addition, we did not place any limit on the publication year: all the studies referenced since the creation of the database were included. Finally, we used the RoB2 tool to evaluate any potential biases.

### **4.4. Limitations**

On the one hand, we chose only studies involving adults to focus on one type of population and reduce the heterogeneity of the populations. On the other hand, we included only RCT. Indeed, these studies specify the sample size estimation and are less prone to bias. Moreover, since we knew there was heterogeneity in how VFD are defined, we selected the studies published only in journals with a 2021 impact factor greater than 5 in the hope of a more rigorous methodology.



The selected studies had different populations, mainly since the inclusion criteria of the reported studies were quite different, thus reducing the possibility of generalizing our results. However, we were interested in the sample size estimation, so this limitation did not impact our conclusions. Next, several studies had missing data, which did not allow us to estimate the sample size in these cases. Hence, the dataset was not comprehensive enough for us to draw conclusions with more certainty. Harhay et al. (6) also found that 63% of the power parameters were missing for the selected RCT. Then, the overdispersion parameter was not available for the use of the zero-inflated negative binomial regression. Consequently, we chose to estimate it with one large database and a simulation. Finally, some studies had to stop early the inclusion, resulting in a very low absolute mean difference in VFD between the two study groups. As a result, the sample size estimation was very high and, therefore, unrealistic.

Other outcomes should be evaluated and compared to VFD to better assess a potential treatment effect. For instance, alive and ventilator-free is a recent hierarchical composite outcome that does not treat mortality as equivalent to prolonged intubation (7).

## CONCLUSION

We conducted a systematic review of ventilator-free days as a primary outcome in randomized controlled studies. First, we observed heterogeneity in the definitions of VFD, making it difficult for us to perform meta-analyses of such studies. Novel approaches, such as a core outcome set, may be more appropriate as they ensure standardization of the outcome. Second, we observed a particular distribution regarding the number of VFD, which warrants further investigation. Finally, like other authors before us, we recommend using a complex model combining the distribution of VFD and the competitive risk event of mortality. To estimate the corresponding sample size, we encourage simulating it using a Monte Carlo Markov chain model.

Le Doyen de l'UFR de Médecine,

**CLAVELOU Pierre**



Le Président du Jury de Thèse,

**FUTIER Emmanuel**

A handwritten signature in black ink, appearing to read 'Emmanuel Futier', with a horizontal line extending to the right.

## ACKNOWLEDGEMENTS

Conflicts of interests/Competing interests: the authors have no conflicts of interest to declare.

Availability of data and material / Code availability: the data and R code will be made available upon reasonable request.

Ethics approval: not applicable.

Consent to participate: not applicable.

Consent for publication: not applicable.

Acknowledgements: We thank Ambre Salis for her help with programming and proofreading the article. Research reported in this publication was supported by DaCCoTA (the National Institute of General Medical Sciences of the National Institutes of Health under Award Number U54GM128729).

## REFERENCES

1. Marshall DC, Hatch RA, Gerry S, Young JD, Watkinson P. Conditional Survival With Increasing Duration of ICU Admission: An Observational Study of Three Intensive Care Databases. *Crit Care Med.* 2020 Jan;48(1):91–7.
2. Herridge MS, Chu LM, Matte A, Tomlinson G, Chan L, Thomas C, et al. The RECOVER Program: Disability Risk Groups and 1-Year Outcome after 7 or More Days of Mechanical Ventilation. *Am J Respir Crit Care Med.* 2016 Oct;194(7):831–44.
3. Hill AD, Fowler RA, Burns KEA, Rose L, Pinto RL, Scales DC. Long-Term Outcomes and Health Care Utilization after Prolonged Mechanical Ventilation. *Ann Am Thorac Soc.* 2017 Mar;14(3):355–62.
4. Veldhoen RA, Howes D, Maslove DM. Is Mortality a Useful Primary End Point for Critical Care Trials? *Chest.* 2020 Jul 1;158(1):206–11.
5. Spragg RG, Bernard GR, Checkley W, Curtis JR, Gajic O, Guyatt G, et al. Beyond Mortality: future clinical research in acute lung injury. *Am J Respir Crit Care Med.* 2010 May 15;181(10):1121–7.
6. Harhay MO, Wagner J, Ratcliffe SJ, Bronheim RS, Gopal A, Green S, et al. Outcomes and Statistical Power in Adult Critical Care Randomized Trials. *Am J Respir Crit Care Med.* 2014 Jun 15;189(12):1469–78.
7. Novack V, Beitler JR, Yitshak-Sade M, Thompson BT, Schoenfeld DA, Rubenfeld G, et al. Alive and Ventilator-Free: A Hierarchical, Composite Outcome for Clinical Trials in the Acute Respiratory Distress Syndrome. *Crit Care Med.* 2020 Feb;48(2):158–66.
8. Yehya N, Harhay MO, Curley MAQ, Schoenfeld DA, Reeder RW. Reappraisal of Ventilator-Free Days in Critical Care Research. *Am J Respir Crit Care Med.* 2019 Oct 1;200(7):828–36.
9. Bernard GR, Artigas A, Brigham KL, Carlet J, Falke K, Hudson L, et al. The American-European Consensus Conference on ARDS. Definitions, mechanisms, relevant outcomes, and clinical trial coordination. *Am J Respir Crit Care Med.* 1994 Mar;149(3 Pt 1):818–24.
10. Schoenfeld DA, Bernard GR, ARDS Network. Statistical evaluation of ventilator-free days as an efficacy measure in clinical trials of treatments for acute respiratory distress syndrome. *Crit Care Med.* 2002 Aug;30(8):1772–7.
11. Chow SC, Shao J, Wang H, Lokhnygina Y. *Sample Size Calculations in Clinical Research.* 3<sup>rd</sup> ed. New York: Chapman and Hall/CRC; 2017, 510 p.

12. Béduneau G, Pham T, Schortgen F, Piquilloud L, Zogheib E, Jonas M, et al. Epidemiology of Weaning Outcome according to a New Definition. The WIND Study. *Am J Respir Crit Care Med*. 2017 Mar 15;195(6):772–83.
13. Fine JP, Gray RJ. A Proportional Hazards Model for the Subdistribution of a Competing Risk. *J Am Stat Assoc*. 1999;94(446):496–509.
14. Haddaway NR, Page MJ, Pritchard CC, McGuinness LA. PRISMA2020: An R package and Shiny app for producing PRISMA 2020-compliant flow diagrams, with interactivity for optimised digital transparency and Open Synthesis. *Campbell Syst Rev*. 2022 Jun;18(2):e1230.
15. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap) -- a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform*. 2009 Apr 1;42(2):377–81.
16. Harris PA, Taylor R, Minor BL, Elliott V, Fernandez M, O’Neal L, et al. The REDCap consortium: Building an international community of software platform partners. *J Biomed Inform*. 2019 Jul 1;95:103208.
17. Sterne JAC, Savović J, Page MJ, Elbers RG, Blencowe NS, Boutron I, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ*. 2019 Aug 28;366:l4898.
18. Hartigan JA, Hartigan PM. The Dip Test of Unimodality. *Ann Stat*. 1985 Mar;13(1):70–84.
19. R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing; 2022 [accessed 2022 Apr 22]. Available from: <https://www.r-project.org/>.
20. Morris PE, Papadakos P, Russell JA, Wunderink R, Schuster DP, Truwit JD, et al. A double-blind placebo-controlled study to evaluate the safety and efficacy of L-2-oxothiazolidine-4-carboxylic acid in the treatment of patients with acute respiratory distress syndrome. *Crit Care Med*. 2008 Mar;36(3):782–8.
21. Paine R, Standiford TJ, Dechert RE, Moss M, Martin GS, Rosenberg AL, et al. A randomized trial of recombinant human GM-CSF for patients with acute lung injury. *Crit Care Med*. 2012 Jan;40(1):90–7.
22. Mackle D, Bellomo R, Bailey M, Beasley R, Deane A, Eastwood G, et al. Conservative Oxygen Therapy during Mechanical Ventilation in the ICU. *N Engl J Med*. 2020 Mar 12;382(11):989–98.
23. Villar J, Ferrando C, Martínez D, Ambrós A, Muñoz T, Soler JA, et al. Dexamethasone treatment for the acute respiratory distress syndrome: a multicentre, randomised controlled trial. *Lancet Respir Med*. 2020 Mar 1;8(3):267–76.

24. Zhou Y, Jin X, Lv Y, Wang P, Yang Y, Liang G, et al. Early application of airway pressure release ventilation may reduce the duration of mechanical ventilation in acute respiratory distress syndrome. *Intensive Care Med.* 2017 Nov;43(11):1648–59.
25. Trouillet JL, Luyt CE, Guiguet M, Ouattara A, Vaissier E, Makri R, et al. Early Percutaneous Tracheotomy Versus Prolonged Intubation of Mechanically Ventilated Patients After Cardiac Surgery: a randomized trial. *Ann Intern Med.* 2011 Mar 15;154(6):373–83.
26. Simonis FD, Neto AS, Binnekade JM, Braber A, Bruin KCM, Determann RM, et al. Effect of a Low vs Intermediate Tidal Volume Strategy on Ventilator-Free Days in Intensive Care Unit Patients Without ARDS: A Randomized Clinical Trial. *JAMA.* 2018 Nov 13;320(18):1872–80.
27. Algera AG, Pisani L, Neto AS, den Boer SS, Bosch FFH, Bruin K, et al. Effect of a Lower vs Higher Positive End-Expiratory Pressure Strategy on Ventilator-Free Days in ICU Patients Without ARDS: A Randomized Clinical Trial. *JAMA.* 2020 Dec 22;324(24):2509–20.
28. Tomazini BM, Maia IS, Cavalcanti AB, Berwanger O, Rosa RG, Veiga VC, et al. Effect of Dexamethasone on Days Alive and Ventilator-Free in Patients With Moderate or Severe Acute Respiratory Distress Syndrome and COVID-19: The CoDEX Randomized Clinical Trial. *JAMA.* 2020 Oct 6;324(13):1307–16.
29. Grieco DL, Menga LS, Cesarano M, Rosà T, Spadaro S, Bitondo MM, et al. Effect of Helmet Noninvasive Ventilation vs High-Flow Nasal Oxygen on Days Free of Respiratory Support in Patients With COVID-19 and Moderate to Severe Hypoxemic Respiratory Failure: The HENIVOT Randomized Clinical Trial. *JAMA.* 2021 May 4;325(17):1731–43.
30. van Meenen DMP, van der Hoeven SM, Binnekade JM, de Borgie CAJM, Merkus MP, Bosch FH, et al. Effect of On-Demand vs Routine Nebulization of Acetylcysteine With Salbutamol on Ventilator-Free Days in Intensive Care Unit Patients Receiving Invasive Ventilation: A Randomized Clinical Trial. *JAMA.* 2018 Mar 13;319(10):993–1001.
31. Spragg RG, Lewis JF, Walmrath HD, Johannigman J, Bellingan G, Laterre PF, et al. Effect of Recombinant Surfactant Protein C-Based Surfactant on the Acute Respiratory Distress Syndrome. *N Engl J Med.* 2004 Aug 26;351(9):884–92.
32. Welte T, Dellinger RP, Ebelt H, Ferrer M, Opal SM, Singer M, et al. Efficacy and safety of trimodulin, a novel polyclonal antibody preparation, in patients with severe community-acquired pneumonia: a randomized, placebo-controlled, double-blind, multicenter, phase II trial (CIGMA study). *Intensive Care Med.* 2018;44(4):438–48.
33. Rice TW, Wheeler AP, Thompson BT, deBoisblanc BP, Steingrub J, Rock P, et al. Enteral Omega-3 Fatty Acid, Gamma-Linolenic Acid, and Antioxidant Supplementation in Acute Lung Injury. *JAMA.* 2011 Oct 12;306(14):1574–81.

34. Bernard GR, Francois B, Mira JP, Vincent JL, Dellinger RP, Russell JA, et al. Evaluating the Efficacy and Safety of Two Doses of the Polyclonal Anti-Tumor Necrosis Factor- $\alpha$  Fragment Antibody AZD9773 in Adult Patients With Severe Sepsis and/or Septic Shock: Randomized, Double-Blind, Placebo-Controlled Phase IIb Study\*. *Crit Care Med*. 2014 Mar;42(3):504–11.
35. Chung KK, Wolf SE, Renz EM, Allan PF, Aden JK, Merrill GA, et al. High-frequency percussive ventilation and low tidal volume ventilation in burns: A randomized controlled trial. *Crit Care Med*. 2010 Oct;38(10):1970–7.
36. Rice TW, Wheeler AP, Thompson BT, Steingrub J, Hite RD, Moss M, et al. Initial Trophic vs Full Enteral Feeding in Patients With Acute Lung Injury: The EDEN Randomized Trial. *JAMA*. 2012 Feb 22;307(8):795–803.
37. Bein T, Weber-Carstens S, Goldmann A, Müller T, Staudinger T, Brederlau J, et al. Lower tidal volume strategy ( $\approx 3$  ml/kg) combined with extracorporeal CO<sub>2</sub> removal versus 'conventional' protective ventilation (6 ml/kg) in severe ARDS: the prospective randomized Xtravent-study. *Intensive Care Med*. 2013 May;39(5):847–56.
38. Hodgson CL, Cooper DJ, Arabi Y, King V, Bersten A, Bihari S, et al. Maximal Recruitment Open Lung Ventilation in Acute Respiratory Distress Syndrome (PHARLAP). A Phase II, Multicenter Randomized Controlled Clinical Trial. *Am J Respir Crit Care Med*. 2019 Dec 1;200(11):1363–72.
39. Kacmarek RM, Villar J, Parrilla D, Alba F, Solano R, Liu S, et al. Neurally adjusted ventilatory assist in acute respiratory failure: a randomized controlled trial. *Intensive Care Med*. 2020 Dec;46(12):2327-37.
40. Kacmarek RM, Wiedemann HP, Lavin PT, Wedel MK, Tütüncü AS, Slutsky AS. Partial Liquid Ventilation in Adult Patients with Acute Respiratory Distress Syndrome. *Am J Respir Crit Care Med*. 2006 Apr 15;173(8):882–9.
41. Liu KD, Levitt J, Zhuo H, Kallet RH, Brady S, Steingrub J, et al. Randomized Clinical Trial of Activated Protein C for the Treatment of Acute Lung Injury. *Am J Respir Crit Care Med*. 2008 Sep 15;178(6):618–23.
42. Rice TW, Mogan S, Hays MA, Bernard GR, Jensen GL, Wheeler AP. Randomized Trial of Initial Trophic versus Full-Energy Enteral Nutrition in Mechanically Ventilated Patients with Acute Respiratory Failure. *Crit Care Med*. 2011 May;39(5):967–74.
43. Matthay MA, Brower RG, Carson S, Douglas IS, Eisner M, Hite D, et al. Randomized, Placebo-controlled Clinical Trial of an Aerosolized  $\beta_2$ -Agonist for Treatment of Acute Lung Injury. *Am J Respir Crit Care Med*. 2011 Sep 1;184(5):561–8.
44. Bennett-Guerrero E, Romeiser JL, Talbot LR, Ahmed T, Mamone LJ, Singh SM, et al. Severe Acute Respiratory Syndrome Coronavirus 2 Convalescent Plasma Versus Standard Plasma in

Coronavirus Disease 2019 Infected Hospitalized Patients in New York: A Double-Blind Randomized Trial. *Crit Care Med*. 2021 Jul 1;49(7):1015–25.

45. McAuley DF, Laffey JG, O’Kane CM, Perkins GD, Mullan B, Trinder TJ, et al. Simvastatin in the Acute Respiratory Distress Syndrome. *N Engl J Med*. 2014 Oct 30;371(18):1695–703.

46. Bodet-Contentin L, Frasca D, Tavernier E, Feuillet F, Foucher Y, Giraudeau B. Ventilator-Free Day Outcomes Can Be Misleading. *Crit Care Med*. 2018 Mar;46(3):425–9.

47. Jabaudon M, Blondonnet R, Pereira B, Cartin-Ceba R, Lichtenstern C, Mauri T, et al. Plasma sRAGE is independently associated with increased mortality in ARDS: a meta-analysis of individual patient data. *Intensive Care Med*. 2018 Sep 1;44(9):1388–99.

48. Contentin L, Ehrmann S, Giraudeau B. Heterogeneity in the Definition of Mechanical Ventilation Duration and Ventilator-Free Days. *Am J Respir Crit Care Med*. 2014 Apr 15;189(8):998–1002.

49. Blackwood B, Ringrow S, Clarke M, Marshall JC, Connolly B, Rose L, et al. A Core Outcome Set for Critical Care Ventilation Trials. *Crit Care Med*. 2019 Oct;47(10):1324–31.

50. Ludbrook J. Should we use one-sided or two-sided P values in tests of significance? *Clin Exp Pharmacol Physiol*. 2013 Jun;40(6):357–61.

51. Schober P, Bossers SM, Schwarte LA. Statistical Significance Versus Clinical Importance of Observed Effect Sizes: What Do P Values and Confidence Intervals Really Represent? *Anesth Analg*. 2018 Mar;126(3):1068–72.

52. Nakagawa S, Cuthill IC. Effect size, confidence interval and statistical significance: a practical guide for biologists. *Biol Rev Camb Philos Soc*. 2007 Nov;82(4):591–605.

53. Cohen J. *Statistical Power Analysis for the Behavioral Sciences*. 2<sup>nd</sup> ed. New York: Routledge; 1988, 567 p.

54. Tai BC, Chen ZJ, Machin D. Estimating sample size in the presence of competing risks - Cause-specific hazard or cumulative incidence approach? *Stat Methods Med Res*. 2018 Jan 1;27(1):114–25.

55. Latouche A, Allignol A, Beyersmann J, Labopin M, Fine JP. A competing risks analysis should report results on all cause-specific hazards and cumulative incidence functions. *J Clin Epidemiol*. 2013 Jun; 66(6):648-53.

56. Goldthorpe JH, Galtung J. *Theory and Methods of Social Research*. *Econ J*. 1969 Sep 1;79(315):665–6.



57. Benaglia T, Chauveau D, Hunter DR, Young D. *mixtools*: An R Package for Analyzing Mixture Models. *J Stat Soft.* 2009 Oct 21;32(6):1-29.
58. Steingrímsson JA, Hanley DF, Rosenblum M. Improving precision by adjusting for prognostic baseline variables in randomized trials with binary outcomes, without regression model assumptions. *Contemp Clin Trials.* 2017 Mar 1;54:18–24.
59. Kumle L, Vö MLH, Draschkow D. Estimating power in (generalized) linear mixed models: An open introduction and tutorial in R. *Behav Res Methods.* 2021 Dec;53(6):2528–43.
60. Kelter R. Bayesian Hodges-Lehmann tests for statistical equivalence in the two-sample setting: Power analysis, type I error rates and equivalence boundary selection in biomedical research. *BMC Med Res Methodol.* 2021 Aug 17;21(1):171.
61. Sidebotham D, Popovich I, Lumley T. A Bayesian analysis of mortality outcomes in multicentre clinical trials in critical care. *Br J Anaesth.* 2021 Sep 1;127(3):487–94.
62. Gagniuc PA. *Markov Chains: From Theory to Implementation and Experimentation.* Hoboken: Wiley; 2017 Jul, 256 p.
63. Morton SE, Chiew YS, Pretty C, Moltchanova E, Scarrott C, Redmond D, et al. Effective sample size estimation for a mechanical ventilation trial through Monte-Carlo simulation: Length of mechanical ventilation and Ventilator Free Days. *Math Biosci.* 2017 Feb 1;284:21–31.
64. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ.* 2021 Mar 29;372:n71.

## APPENDICES

### Appendix I: Searching terms used for reviewing

#### MEDLINE

("Ventilator free days"[TW] OR "Ventilator free day"[TW] OR "Ventilation free days"[TW] OR "MV-free days"[TW] OR (Ventilat\*[Title] AND free[Title] AND (days[Title] OR day[Title])))

AND

("the new england journal of medicine"[Journal] OR "Lancet (London, England)"[Journal] OR "Nature"[Journal] OR "Nature medicine"[Journal] OR "Science (New York, N.Y.)"[Journal] OR "JAMA"[Journal] OR "Annals of internal medicine"[Journal] OR "plos medicine"[Journal] OR "BMJ (Clinical research ed.)"[Journal] OR "BMJ"[journal] OR "american journal of respiratory and critical care medicine"[Journal] OR "Chest"[Journal] OR "Critical care (London, England)"[Journal] OR "Critical care medicine"[Journal] OR "Intensive care medicine"[Journal] OR "The Lancet. Respiratory medicine"[Journal] OR "Nature immunology"[Journal] OR "Nature reviews. Immunology"[Journal])

AND

(randomized controlled trial[pt] OR controlled clinical trial[pt] OR randomized[TW] OR placebo[TW] OR clinical trials as topic[mesh:noexp] OR randomly[TW] OR trial[TW])

#### Embase

(('ventilator free days' OR 'ventilator-free days' OR 'ventilator free day' OR 'ventilation free days' OR 'mv-free days' OR (ventilat\*:ti AND free:ti AND (days:ti OR day:ti))) AND ('new england journal of medicine':jt OR 'nature'/jt OR 'lancet (london, england)'/jt OR 'the lancet'/jt OR 'nature medicine':jt OR 'science'/jt OR 'science (new york, n.y.)'/jt OR 'jama'/jt OR 'jama - journal of the american medical association'/jt OR 'jama : the journal of the american medical association'/jt OR 'annals of internal medicine':jt OR 'plos medicine'/jt OR 'bmj'/jt OR 'bmj (clinical research ed.)'/jt OR 'bmj (online)'/jt OR 'the bmj'/jt OR 'american journal of respiratory and critical care medicine'/jt OR 'chest'/jt OR 'critical care'/jt OR 'critical care (london, england)'/jt OR 'critical care medicine'/jt OR 'intensive care medicine'/jt OR 'the lancet respiratory medicine'/jt OR 'nature immunology':jt OR 'nature reviews immunology'/jt)) AND ('article'/it OR 'article in press'/it)

## **Appendix II: Extracted data**

### **Features**

PMID (PubMed Identifier) if available, year of publication, design, definition of ventilator-free days (VFD), type of ventilation, onset of VFD, length of VFD, PaO<sub>2</sub>/FiO<sub>2</sub> ratio.

### **Methods**

Power,  $\alpha$  risk, inflation  $\alpha$  risk, type of study (superiority/non-inferiority/equivalence), type of test (unilateral/bilateral), type of analysis (intention-to-treat/per-protocol), margin of non-inferiority if applicable, VFD expected in control and experimental groups, distribution of VFD and correction if applicable, mortality rate expected in control and experimental groups, expected difference between groups and justification, expected drop-out patients, management of missing values, statistical method for sample size estimation and for analysis of the primary outcome of VFD, sample size estimation.

### **Results**

Sample size, missing values, premature cessation, drop-out patients, VFD in control and experimental groups, mortality rate in control and experimental groups, difference between groups and confidence interval, effect size, p-value.

### **Research Electronic Data Capture (REDCap)**

REDCap is a secure, web-based software platform designed to support data capture for research studies, providing 1) an intuitive interface for validated data capture, 2) audit trails for tracking data manipulation and export procedures, 3) automated export procedures for seamless data downloads to common statistical packages, and 4) procedures for data integration and interoperability with external sources.

## Appendix III: Statistical analysis plan

All sample size estimation were achieved with R version 4.2.0 using the following equations with  $\alpha$ : type I error,  $(1 - \beta)$ : statistical power,  $n$ : size of one group and  $N$ : size of all groups.

For these estimations, we used the DaCCoTA Project (E1) and the Sample Size Calculations in Clinical Research (E2).

When the mean is missing, we estimated it from the median, the interquartile range, and the sample size (E3).

First, we calculated the sample size with the expected parameters displayed in the Material and Methods section (corresponding to the estimated sample size), with the different statistical tests presented below. Second, we calculated the sample size with the observed parameters displayed in the Results section (corresponding to the observed sample size), with the same statistical tests.

### Student *t*-test based on Gaussian distribution

For the sample size estimation based on expected parameters, we used the standard deviation of the control group when it was available, displayed in the Material and Methods section.

For the sample size estimation based on observed parameters, we used the pooled standard deviation  $\sigma$  :

$\sqrt{\left(\frac{\sigma_1^2 + \sigma_2^2}{2}\right)}$ , where  $\sigma_1$  is the standard deviation of the control group and  $\sigma_2$  the standard deviation of the interventional group, given that the two  $\sigma$  are available in the Results section.

For the sample size estimation, we used the following equation:

$n = \frac{2\sigma^2}{\Delta^2} \left( Z_{1-\frac{\alpha}{2}} + Z_{1-\beta} \right)^2$ , with  $\sigma$  the standard deviation and  $\Delta$  the absolute difference between groups.

We used the *pwr* package to estimate the sample size for the Student *t*-test for the superiority studies.

For the non-inferiority study, we used the following equation according to one group:

$n = \frac{2\sigma^2}{(\Delta - \delta)^2} \left( Z_{1-\frac{\alpha}{2}} + Z_{1-\beta} \right)^2$ , with  $\delta$  the margin of the non-inferiority.

We used the *TrialSize* package to estimate the sample size for the Student *t*-test for the non-inferiority studies.

### Mann–Whitney–Wilcoxon

This non-parametric test uses the sample size estimation of the student *t*-test. We therefore simply multiplied the sample size estimated according to the student *t*-test by 1.15 to have an increase of 15% in the sample size (E4).

### Noether formula (E5)

First, we estimated the effect size  $\frac{\delta}{\sigma} = \frac{\mu_1 - \mu_2}{\sigma}$ , where  $\delta$  is the absolute difference between groups,  $\sigma$  the standard deviation,  $\mu_1$  the mean of the control group and  $\mu_2$  the standard deviation of the interventional group.

Then, we calculated the relative effect: there is a relation between the corresponding relative effect  $p$  and the effect size for a normal distribution (E6).

$\delta/\sigma$	-2	-1.5	-1	-0.5	0	0.5	1	1.5	2
$p$	0.08	0.14	0.24	0.36	0.50	0.64	0.76	0.86	0.92

Finally, for the sample size estimation, we used the following equation:

$N = \frac{(Z_\alpha + Z_\beta)^2}{12c(1-c)(p - \frac{1}{2})^2}$ , where  $c$  is the proportion of subjects in the control group and  $p$  the relative effect.

We used the *rankFD* package to estimate the sample size for the Mann–Whitney–Wilcoxon rank-sum test using the Noether formula.

### **One-way ANOVA (ANalysis Of VAriance)**

We used ANOVA when there were 3 or more groups in the selected studies.

We estimated first the Cohen's  $f$  effect size:

$f = \frac{\sigma_m}{\sigma}$ , where  $\sigma_m = \sqrt{\frac{\sum_{i=1}^k ((m_i - m)^2)}{k}}$ , with  $k$  the number of groups,  $m_i$  the mean of each group and  $m$  the grand mean corresponding to the mean of the means, and  $\sigma$  the pooled standard deviation.

Then, for the sample size estimation, we used the following equation:

$n = \frac{n_{.05}}{400 \times (f^2)} + 1$ , where  $n_{.05}$  is the necessary sample size for the given  $\alpha$  risk,  $u$  (the degree of freedom i.e.,  $k - 1$ ) and desired power  $\beta$  at  $f = 0.5$  read from the Table (E7).

We used the *pwr* package to estimate the sample size for the one-way ANOVA.

### **Kruskal-Wallis**

We used Kruskal-Wallis when there were 3 or more groups in the selected studies. This non-parametric test uses the sample size estimation of the ANOVA. Like above, we multiply the sample size estimated according to the ANOVA test by 1.15 to have an increase of 15% in the sample size (E4).

### **Cox regression**

We did not use the log-rank test because testing superiority or non-inferiority is difficult to be carried out (E2).

For the sample size estimation, we used the following equation:

$$N = \frac{\left( Z_{1-\frac{\alpha}{2}} + Z_{1-\beta} \right)^2}{(\log(HR) - \delta)^2 \times p_e \times p_c \times d}$$
, where  $HR$  is the Hazard Ratio of the VFD,  $\delta$  the margin of the non-inferiority if applicable,  $p_e$  the proportion of patients within the experimental group,  $p_c$  the proportion of patients within the control group and  $d$  the probability of observing an event.

We used the *TrialSize* package to estimate the sample size for the Cox regression.

When the HR isn't available, we estimated it as follows:

$\frac{\mu(VFD_e)}{\mu(VFD_c)}$ , where  $\mu(VFD_e)$  is the mean of the VFD in the experimental group and  $\mu(VFD_c)$  the mean of the VFD in the control group, as described by Zhu and Lakkis (E8).

### **Zero-inflated negative binomial regression (E8)**

First, we calculated the mean rates of events ratio:

$r = \frac{\mu(VFD_E)}{\mu(VFD_C)}$ , where  $\mu(VFD_E)$  is the mean of the VFD in the experimental group and  $\mu(VFD_C)$  the mean of the VFD in the control group.

Then, we estimated the overdispersion  $\kappa$  by using the simulation of Bodet-Contentin et al. (E9) and the meta-analysis of Jabaudon et al. (E10) with a generalized linear mixed model. The parameters of this model were the negative binomial family with a link logit and a zero inflation. The center was considered as random effect. We found the same overdispersion in the two studies:  $\kappa = 0.057$ .

For the sample size estimation, we used the "approach 2" as recommended in their article:

$$n = \frac{\left( Z_{\alpha} \sqrt{V_{0|2}} + Z_{\beta} \sqrt{V_A} \right)^2}{(\log r)^2},$$

$V_A$  is the variance estimated under the alternative hypothesis:

$V_A = \frac{1}{\mu_t} \left( \frac{1}{\lambda_C} + \frac{1}{\lambda_E \theta} \right) + \frac{(1+\theta)\kappa}{\theta}$ , where  $\mu_t$  is the average subject exposure time (same unit as the one used for the mean event rate per time ( $\lambda$ )),  $\lambda_C$  the mean event rate per time in the control group,  $\lambda_E$  the mean event rate per time in the interventional group and  $\theta = n_E/n_C$  the allocation ratio with  $n_E$  the sample size in the interventional group and  $n_C$  the sample size in the control group.

$V_{0|2}$  is the variance estimated under the null hypothesis according to the method 2 using the true event rates ( $\lambda_C = \mu(VFD_C)$  and  $\lambda_E = \mu(VFD_E)$ ):

$$V_{0|2} = V_A$$

We used the *MKmisc* package to estimate the sample size for the zero-inflated negative binomial regression.

### **Fine and Gray regression (E11)**

First, we estimated the  $\Xi$  corresponding of:

$\Xi = \frac{\lambda_{1E}(t) + \lambda_{2E}(t)}{\lambda_{1C}(t) + \lambda_{2C}(t)}$ , where  $\lambda_{1E}$  is the proportion of patients within the experimental group with the event of interest (i.e. extubation),  $\lambda_{2E}$  is the proportion of patients within the experimental group with the competing event (i.e. death),  $\lambda_{1C}$  is the proportion of patients within the control

group with the event of interest and  $\lambda_{2C}$  is the proportion of patients within the control group with the competing event.

Then, for the sample size estimation, we used the following equation:

$$N = \frac{\left( Z_{1-\frac{\alpha}{2}} + Z_{1-\beta} \right)^2}{(\log(\Xi) - \delta)^2 \times p_E \times p_C \times d}$$
, where  $\delta$  is the margin of the non-inferiority if applicable,  $p_E$  the proportion of patients within the experimental group,  $p_C$  the proportion of patients within the control group and  $d$  the probability of observing an event.

Since the package to estimate the sample size for a Fine and Gray regression is not available in R, we built it.

## **References**

- E1. The DaCCoTA (Dakota Cancer Collaborative on Translational Activity) Project is supported by the National Institute of General Medical Sciences of the National Institutes of Health under Award Number U54GM128729.
- E2. Chow SC, Shao J, Wang H, Lokhnygina Y. Sample Size Calculations in Clinical Research. 3<sup>rd</sup> ed. New York: Chapman and Hall/CRC; 2017, 510 p.
- E3. Wan X, Wang W, Liu J, Tong T. Estimating the sample mean and standard deviation from the sample size, median, range and/or interquartile range. BMC Med Res Methodol. 2014 Dec 19;14:135.
- E4. Lehmann EL. Nonparametrics: statistical methods based on ranks. San Francisco: Holden-Day; 1975, 457 p.
- E5. Noether GE. Sample Size Determination for Some Common Nonparametric Tests. J Am Stat Assoc. 1987 Jun;82(398):645–7.
- E6. Brunner E, Bathke AC, Konietzschke F. Rank and pseudo-rank procedures for independent observations in factorial designs: using R and SAS. Cham: Springer; 2018 Jan, 521 p.
- E7. Cohen J. Statistical Power Analysis for the Behavioral Sciences. 2<sup>nd</sup> ed. New York: Routledge; 1988, 567 p.
- E8. Zhu H, Lakkis H. Sample size calculation for comparing two negative binomial rates. Stat Med. 2014 Feb 10;33(3):376–87.
- E9. Bodet-Contentin L, Frasca D, Tavernier E, Feuillet F, Foucher Y, Giraudeau B. Ventilator-Free Day Outcomes Can Be Misleading. Crit Care Med. 2018 Mar;46(3):425–9.
- E10. Jabaudon M, Blondonnet R, Pereira B, Cartin-Ceba R, Lichtenstern C, Mauri T, et al. Plasma sRAGE is independently associated with increased mortality in ARDS: a meta-analysis of individual patient data. Intensive Care Med. 2018 Sep 1;44(9):1388–99.
- E11. Schulgen G, Olschewski M, Krane V, Wanner C, Ruf G, Schumacher M. Sample sizes for clinical trials with time-to-event endpoints and competing risks. Contemp Clin Trials. 2005 Jun;26(3):386–96.

## Appendix IV: Packages R

*Hmisc* for label and describe.

Harrell FEJ. Hmisc: Harrell Miscellaneous. 2022 R package version 4.7-0. Available from: <https://CRAN.R-project.org/package=Hmisc>.

*pwr* for the sample size estimation of the superiority study using Student.

Champely S. pwr: Basic Functions for Power Analysis. 2020 R package version 1.3-0. Available from: <https://CRAN.R-project.org/package=pwr>.

*TrialSize* for the sample size estimation of the non-inferiority studies using Student *t*-test or the Cox regression.

Zhang E, Wu VQ, Chow SC, Zhang HG. TrialSize: R Functions for Chapter 3, 4, 6, 7, 9, 10, 11, 12, 14, 15 of Sample Size Calculation in Clinical Research. 2020 R package version 1.4. Available from: <https://CRAN.R-project.org/package=TrialSize>.

*rankFD* for the sample size estimation using Noether formula.

Konietschke F, Friedrich S, Brunner E, Pauly M. rankFD: Rank-Based Tests for General Factorial Designs. 2022 R package version 0.1.1. Available from: <https://CRAN.R-project.org/package=rankFD>.

*MKmisc* for the sample size estimation using the zero-inflated negative binomial regression.

Kohl M. MKmisc: Miscellaneous functions from M. Kohl. 2021 R package version 1.8. Available from: <https://www.stamats.de>.

*readxl* to read Excel files.

Wickham H, Bryan J. readxl: Read Excel Files. 2022 R package version 1.4.0. Available from: <https://CRAN.R-project.org/package=readxl>.

*glmmTMB* to use a GLMM (Generalized Linear Mixed Model).

Brooks ME, Kristensen K, van Benthem KJ, Magnusson A, Berg CW, Nielsen A, et al. glmmTMB Balances Speed and Flexibility Among Packages for Zero-inflated Generalized Linear Mixed Modeling. *The R Journal*. 2017 Dec;9(2):378-400.

*PRISMA2020* for the DiagrammeR.

Haddaway NR, Page MJ, Pritchard CC, McGuinness LA. PRISMA2020: An R package and Shiny app for producing PRISMA 2020-compliant flow diagrams, with interactivity for optimised digital transparency and Open Synthesis. *Campbell Syst Rev*. 2022 Jun;18(2):e1230.

*diptest* and *multimode* for the bimodal distribution.

Maechler M. diptest: Hartigan's Dip Test Statistic for Unimodality - Corrected. 2021 R package version 0.76-0. Available from: <https://CRAN.R-project.org/package=diptest>.



## Appendix V: supplement tables

Study	VFD (ctrl. Group)	VFD (exp. Group)	Mean difference	Standard deviation	$\alpha$ risk (%)	Power (%)	Tail(s)	Drop-out (%)	Mortality considered
Morris et al., 2008 (20)	NA	NA	NA	NA	5	NA	NA	0	No
Paine et al., 2012 (21)	NA	NA	3.9	NA	5	80	Two	0	No
Mackle et al., 2020 (22)	16.4	19	2.6	11.3	5	90	Two	1	No
Villar et al., 2020 (23)	9	11	2	10.7	5	80	Two	0	Yes
Zhou et al., 2017 (24)	14.5	19.5	5	10.4	5	80	Two	25	No
Trouillet et al., 2011 (25)	23	30	7	17	5	80	Two	9	No
Simonis et al., 2018 (26)	23	24	1	5	5	80	Two	20	No
Algera et al., 2020 (27)	16	16	1.6	10	5	80	One	10	No
Tomazini et al., 2020 (28)	8	11	3	9	5	80	Two	4	No
Grieco et al., 2021 (29)	11.6	14.5	2.9	5	5	80	Two	10	No
van Meenen et al., 2018 (30)	NA	NA	0.5	NA	5	80	One	5	No
Spragg et al., 2004 (31)	NA	NA	4	10	5	80	Two	0	No
Welte et al., 2018 (32)	NA	NA	2	3	2.5	80	One	20	No

Rice et al., 2011 (33)	14	16.25	2.25	10.5	5	90.7	Two	0	No
Bernard et al., 2014 (34)	NA	NA	4.2	NA	10	90	One	0	No
Chung et al., 2010 (35)	13	15	2	7	5	80	Two	0	No
Rice et al., 2012 (36)	14	16.25	2.25	10.5	4.29	90.7	Two	0	No
Bein et al., 2013 (37)	6	11	5	10	5	90	Two	10	No
Hodgson et al., 2019 (38)	NA	NA	NA	NA	5	80	Two	5	No
Kacmarek et al., 2020 (39)	21	23	2	6	5	80	Two	0	No
Kacmarek et al., 2006 (40)	NA	NA	4	NA	5	80	Two	0	No
Liu et al., 2008 (41)	12	18.5	6.5	11	5	80	Two	0	No
Rice et al., 2011 (42)	17.6	29.6	3	7.3	5	80	Two	5	No
Matthay et al., 2011 (43)	NA	NA	2.25	10.5	5	90.7	Two	0	No
Bennett-Guerrero et al., 2021 (44)	NA	NA	2.5	6	5	90	Two	0	No
McAuley et al., 2014 (45)	12.7	15.3	2.6	10.6	5	80	Two	3	No

**Table S1. Expected parameters reported for the sample size estimation (Material and Methods section).**  
Ctrl: control; exp: experimental; NA: not available; VFD: ventilator-free days.

<b>Study</b>	<b>SS estimated</b>	<b>SS Student</b>	<b>SS Wilcoxon</b>	<b>SS Noether</b>	<b>SS Cox</b>	<b>SS ZINB</b>	<b>SS F&amp;G</b>
Morris et al., 2008 (20)	352	NA	NA	NA	NA	NA	NA
Paine et al., 2012 (21)	200	NA	NA	NA	NA	NA	NA
Mackle et al., 2020 (22)	1000	796	915	1273	1582	1524	NA
Villar et al., 2020 (23)	320	901	1036	1441	614	1054	NA
Zhou et al., 2017 (24)	138	138	158	218	282	296	NA
Trouillet et al., 2011 (25)	216	187	215	297	351	130	NA
Simonis et al., 2018 (26)	952	787	905	1258	13653	10448	NA
Algera et al., 2020 (27)	980	483	556	1549	2228	1823	NA
Tomazini et al., 2020 (28)	350	285	327	453	244	446	NA
Grieco et al., 2021 (29)	110	95	110	150	497	663	NA
van Meenen et al., 2018 (30)	950	NA	NA	NA	51194	155	NA
Spragg et al., 2004 (31)	440	198	228	315	NA	NA	NA
Welte et al., 2018 (32)	160	73	84	113	NA	NA	NA
Rice et al., 2011 (33)	1000	941	1082	1505	1586	1765	NA

Bernard et al., 2014 (34)	300	NA	NA	NA	NA	NA	NA	NA	NA
Chung et al., 2010 (35)	170	387	445	297	1208	1499	NA	NA	NA
Rice et al., 2012 (36)	1000	978	1125	1565	1665	1835	NA	NA	NA
Bein et al., 2013 (37)	120	128	147	201	67	146	NA	NA	NA
Hodgson et al., 2019 (38)	340	NA	NA	NA	NA	NA	NA	NA	NA
Kacmarek et al., 2020 (39)	306	285	327	453	2988	2432	NA	NA	NA
Kacmarek et al., 2006 (40)	309	NA	NA	NA	NA	NA	NA	NA	NA
Liu et al., 2008 (41)	90	92	106	144	132	157	NA	NA	NA
Rice et al., 2011 (42)	200	188	216	298	998	928	NA	NA	NA
Matthay et al., 2011 (43)	1000	941	1082	1505	NA	NA	NA	NA	NA
Bennett-Guerrero et al., 2021 (44)	500	363	418	606	NA	NA	NA	NA	NA
McAuley et al., 2014 (45)	540	524	602	837	713	888	NA	NA	NA

**Table SII. Sample size estimation according to the statistical test analysis, based on expected parameters.**

F&G: Fine and Gray regression; NA: not available; SS: sample size; ZINB: zero-inflated negative binomial regression.

Study	VFD ctrl. group	Ctrl. group	VFD exp. group	Exp. group	Mean difference	Standard deviation	Mort. rate ctrl. group	Mort. rate exp. group	Mortality frame	Dropout rate (%)
Morris et al., 2008 (20)	13.5	Placebo	8.3	OTZ	5.2	10.2	15.8	29.7	28	0.5
Paine et al., 2012 (21)	10.7	Placebo	10.8	GM-CSF	0.1	10.4	23 / 27	17 / 22	28 / 180	1.5
Mackle et al., 2020 (22)	16	Conservative O <sub>2</sub>	15.5	Usual O <sub>2</sub>	0.5	11.7	32.5 / 34.5	34.7 / 35.7	90 / 180	0
Villar et al., 2020 (23)	7.5	Standard care	12.3	Standard care + Dexamethasone	4.8	9.5	36	21	60	0
Zhou et al., 2017 (24)	5.7	LTV	16.3	APRV	10.6	11.0	34.3 / 37.3	19.7 / 23.9	ICU / hosp.	0
Trouillet et al., 2011 (25)	28.3	Prolonged ventilation	30.4	Early tracheotomy	2.1	23.1	21 / 28 / 30	16 / 26 / 30	28 / 60 / 90	0
Simonis et al., 2018 (26)	15.5	Intermediate tidal volume	15.2	Low tidal volume	0.3	11.5	32.1 / 37.8	34.9 / 39.1	28 / 90	0.6
Algera et al., 2020 (27)	14.4	Higher PEEP	14.8	Lower PEEP	0.4	19.8	42 / 44.3	38.4 / 41.6	28 / 90	1.1
Tomazini et al., 2020 (28)	4	Standard care	6.6	Standard care + Dexamethasone	2.6	9.1	61.5	56.3	28	0

Grieco et al., 2021 (29)	13	HFNO	15	Helmet NIV	2	11.0	18 / 22	15 / 24	28 / 60	0.9
van Meenen et al., 2018 (30)	15.3	Routine nebulization	15.6	On demand nebulization	0.3	19.3	31.9 / 36.7	31 / 38.5	28 / 90	2.3
Spragg et al., 2004 (31)	9	Standard care	8.2	Standard care + surfactant	0.8	15.8	32	36	28	0
Welte et al., 2018 (32)	9.6	Placebo	11	Trimodulin	1.4	9.5	27.8	22.2	28	0
Rice et al., 2011 (33)	17.2	Control	14	n-3 supplement	3.2	10.7	17.6	25.1	28	0
Bernard et al., 2014 (34)	18.3	Placebo	19.7	AZD9773 (low or high dose)	1.4	15.2	20	15	28	1.3
Chung et al., 2010 (35)	11	LTV	12	HFPV	1	9.0	19	19	28	0
Rice et al., 2012 (36)	15	Full feeding	14.9	Trophic feeding	0.1	10.9	22.2	23.2	60	0
Bein et al., 2013 (37)	9.3	Conventional tidal	10	Lower tidal + avECCO <sub>2</sub> -R	0.7	8.5	15.4	17.5	hosp.	0
Hodgson et al., 2019 (38)	12	Protective ventilation	12.3	Max lung recruitment	0.3	16.2	26.8 / 30.4 / 30.4	24.6 / 24.6 / 24.6	28 / 90 / 180	1.7
Kacmarek et al., 2020 (39)	14	Conventional ventilation	16.7	NAVA	2.7	17.3	32.7	26.8	90	1.3

Kacmarek et al., 2006 (40)	13	Conventional ventilation	7.4	PLV (low or high dose)	5.6	8.9	15	26.3	28	0
Liu et al., 2008 (41)	14.3	Placebo	18.3	Activated protein C	4	13.8	13.5	13.5	60	0
Rice et al., 2011 (42)	17.8	Full feeding	17.9	Trophic feeding	0.1	10.5	19.6	22.4	28	0
Matthay et al., 2011 (43)	16.6	Placebo	14.4	Albuterol	2.2	0.9	17.7 / 18.5	23 / 24.3	60 / 90	0
Bennett-Guerrero et al., 2021 (44)	18.7	Standard plasma	19.3	Convalescent plasma	0.6	21.4	27 / 33	24 / 27	28 / 90	0
McAuley et al., 2014 (45)	11.5	Placebo	12.6	Simvastatin	1.1	10.2	26.8	22	28	0.6

**Table SIII. Observed ventilator-free days and mortality rate according to the group (control or experimental).**

APRV: airway pressure-release ventilation; HFNO: high-flow nasal oxygen; HFPV: high-frequency percussive ventilation; hosp.: hospital; ICU: intensive care unit; LTV: low-tidal volume ventilation; Mort.: mortality; NA: not available; NAVA: neurally-adjusted ventilation assist; NIV: non-invasive ventilation; OTZ: oxothiazolidine-4-carboxylic acid; PEEP: positive end-expiratory pressure; PLV: partial liquid ventilation; VFD: ventilator-free day.



Study	SS estimated	Mean difference	SS Student	SS Wilcoxon	SS Noether	SS Cox	SS ZINB	Authors' conclusion
Morris et al., 2008 (20)	215	5.2	123	141	194	105	165	Significant difference
Paine et al., 2012 (21)	132	0.1	339606	390547	544407	285785	453378	No difference between groups
Mackle et al., 2020 (22)	965	0.5	22823	26247	36584	33984	36327	No difference between groups
Villar et al., 2020 (23)	277	4.8	124	143	196	101	184	Significant difference
Zhou et al., 2017 (24)	138	10.6	36	41	54	22	45	Significant difference
Trouillet et al., 2011 (25)	216	2.1	3787	4355	6068	4827	1624	No difference between groups
Simonis et al., 2018 (26)	961	0.3	46139	53060	73961	64740	73337	No difference between groups
Algera et al., 2020 (27)	980	0.4	1884	2167	96648	24019	30811	Significant difference
Tomazini et al., 2020 (28)	299	2.6	383	440	611	99	329	Significant difference
Grieco et al., 2021 (29)	110	2	952	1094	1522	1208	1499	No difference between groups
van Meenen et al., 2018 (30)	944	0.3	18424	21187	164077	47820	581678	Significant difference
Spragg et al., 2004 (31)	448	0.8	12248	14085	19631	2854	5619	No difference between groups
Welte et al., 2018 (32)	160	1.4	1432	1647	2293	1334	2215	No difference between groups

Rice et al., 2011 (33)	272	3.2	480	552	767	831	903	Significant difference
Bernard et al., 2014 (34)	300	1.4	1877	2159	4933	4977	5314	No difference between groups
Chung et al., 2010 (35)	62	1	2545	2927	4077	3267	4868	No difference between groups
Rice et al., 2012 (36)	1000	0.1	529465	608884	848759	826467	916043	No difference between groups
Bein et al., 2013 (37)	79	0.7	4647	5344	7447	4696	8270	No difference between groups
Hodgson et al., 2019 (38)	115	0.3	91565	105300	146781	40560	57276	No difference between groups
Kacmarek et al., 2020 (39)	310	2.7	1286	1479	2058	795	907	Significant difference
Kacmarek et al., 2006 (40)	311	5.6	151	174	191	117	210	Significant difference
Liu et al., 2008 (41)	75	4	375	432	599	407	442	No difference between groups
Rice et al., 2011 (42)	200	0.1	342856	394284	549616	787957	775495	No difference between groups
Matthay et al., 2011 (43)	282	2.2	10	11	14	1742	1894	No difference between groups
Bennett-Guerrero et al., 2021 (44)	74	0.6	4622	5316	133999	53664	229079	No difference between groups
McAuley et al., 2014 (45)	540	1.1	2677	3078	4288	2964	4226	No difference between groups

**Table SIV. Sample size estimation according to the statistical test analysis based on observed parameters, and author's conclusion.**

NA: not available; SS: sample size; ZINB: zero-inflated negative binomial regression.

We did not include the sample size estimation based on the Fine and Gray regression since the extubation incidences were missing.

	Ctrl. Group	Exp. group
$\alpha$ risk (%)	5	
Power (%)	80	
VFD	10.8	10.4
Mean difference	0.46	
Standard deviation	10.8	10.2
Extubation rate	0.572	0.596
Mortality rate	0.28	0.15
SS Student	16519	
SS Wilcoxon	18997	
SS Noether	26478	
SS Cox	13339	
SS ZINB	21437	
SS FG	1777	

**Table SV. Sample size estimation of the simulation from Bodet-Contentin et al. (48) according to the statistical test analysis.**

*Ctrl.: control; exp.: experimental; F&G: Fine and Gray; SS: sample size; VFD: ventilator-free days; ZINB: zero-inflated negative binomial regression.*

## SERMENT D'HIPPOCRATE

*(Conseil National de l'Ordre des Médecins)*

Au moment d'être admis à exercer la Médecine, je promets et je jure d'être fidèle aux lois de l'Honneur et de la Probité.

Mon premier souci sera de rétablir, de préserver ou de promouvoir la santé dans tous ses éléments, physiques et mentaux, individuels et sociaux.

Je respecterai toutes les personnes, leur autonomie et leur volonté, sans aucune discrimination selon leur état ou leurs convictions. J'interviendrai pour les protéger si elles sont affaiblies, vulnérables ou menacées dans leur intégrité ou leur dignité. Même sous la contrainte, je ne ferai pas usage de mes connaissances contre les lois de l'humanité.

J'informerai les patient·e·s des décisions envisagées, de leurs raisons et de leurs conséquences. Je ne tromperai jamais leur confiance et n'exploiterai pas le pouvoir hérité des circonstances pour forcer les consciences.

Je donnerai mes soins à l'indigent et à quiconque me les demandera. Je ne me laisserai pas influencer par la soif du gain ou la recherche de la gloire.

Admis dans l'intimité des personnes, je tairai les secrets qui me seront confiés. Reçu à l'intérieur des maisons, je respecterai les secrets des foyers et ma conduite ne servira pas à corrompre les mœurs.

Je ferai tout pour soulager les souffrances. Je ne prolongerai pas abusivement les agonies. Je ne provoquerai jamais la mort délibérément.

Je préserverai l'indépendance nécessaire à l'accomplissement de ma mission. Je n'entreprendrai rien qui dépasse mes compétences. Je les entretiendrai et les perfectionnerai pour assurer au mieux les services qui me seront demandés.

J'apporterai mon aide à mes pair·e·s ainsi qu'à leurs familles dans l'adversité.

Que les femmes, les hommes et mes pair·e·s m'accordent leur estime si je suis fidèle à mes promesses ; que je sois déshonoré et méprisé si j'y manque.

RENARD-TRICHE Laurent

Signature :

## SERMENT D'HIPPOCRATE

*(Conseil National de l'Ordre des Médecins)*

En présence des Maîtres·ses de cette Faculté et de mes cher·e·s condisciples, je promets et je jure d'être fidèle aux lois de l'Honneur et de la Probité dans l'exercice de la Médecine.

Je donnerai mes soins gratuits à l'indigent et je n'exigerai jamais un salaire au-dessus de mon travail. Admis dans l'intérieur des maisons, mes yeux ne verront pas ce qui s'y passe, ma langue taira les secrets qui me seront confiés et mon état ne servira pas à corrompre les mœurs ni à favoriser le crime.

Respectueux et reconnaissant envers mes Maîtres·ses, je rendrai à leurs enfants l'instruction que j'ai reçue de leurs pères et mères.

Que les femmes et les hommes m'accordent leur estime si je suis fidèle à leur promesse. Que je sois couvert d'opprobre et méprisé de mes pair·e·s si j'y manque.

RENARD-TRICHE Laurent

Signature :

## **VENTILATOR-FREE DAYS AS AN OUTCOME: APPROACHES ON SAMPLE SIZE ESTIMATION FROM A SYSTEMATIC REVIEW.**

### **Abstract:**

**RATIONALE** Ventilator-free days are a composite endpoint increasingly used as primary outcome in critical care trials. Its distribution is skewed and involves competitive risk, which complexifies the sample size estimation.

**OBJECTIVES** To investigate whether the sample size estimation on ventilator-free days was congruent with its distribution, to review its definition and to propose an appropriate methodology for its analysis.

**METHODS** We conducted a systematic review with the databases PubMed and Embase, in adult randomized trials with ventilator-free days as primary outcome. After reviewing its definitions, we extracted the sample size and parameters to estimate it in each study and simulated it through alternative statistical approaches. We assessed the risk of bias with Cochrane risk of bias tool.

**MEASUREMENTS AND MAIN RESULTS** We included 26 studies from 2004 to 2021, mainly superiority studies comparing two groups with several centers. When assessing the risk of bias, most studies were classified as 'some concerns'. The definitions of ventilator-free days were heterogeneous, such as the onset and the respiratory support. Sample size estimation methods were also heterogeneous in most studies. Moreover, they did not take the specific parameters about ventilator-free days into account such as the bimodal distribution and the death as competitive risk. Simulations studies highlighted the need for complex models that few studies used.

**CONCLUSION** We recommend a standardization of both the definition and the methodology of ventilator-free days. Further investigations are also required to choose the best statistical approaches given its distribution and the competitive risk event of mortality. Moreover, the sample size estimation by simulation should be encouraged.

**Systematic review registration:** PROSPERO (ID: CRD42021282304)

### **Keywords:**

- respiration, artificial
- time factors
- ventilators, mechanical
- respiratory insufficiency
- treatment outcome

## LES VENTILATOR-FREE DAYS COMME CRITERE DE JUGEMENT PRINCIPAL : APPROCHES POUR L'ESTIMATION DU NOMBRE DE SUJETS NECESSAIRES A PARTIR D'UNE REVUE SYSTEMATIQUE.

### Résumé :

**CONTEXTE** Les ventilator-free days (VFD) sont de plus en plus utilisés dans les essais cliniques de réanimation comme critère de jugement principal. Leur distribution asymétrique et l'implication de risques compétitifs complexifient le calcul du nombre de sujets nécessaires (NSN).

**OBJECTIFS** Vérifier si l'analyse statistique des essais mesurant les VFD était conforme à leur distribution, examiner leurs définitions et proposer des recommandations sur la méthodologie la plus appropriée à employer pour leur analyse.

**METHODES** Nous avons conduit une revue systématique à l'aide de deux bases de données (PubMed et Embase), d'essais contrôlés randomisés ayant pour critère de jugement principal les VFD, chez les adultes. Après avoir examiné leurs définitions, nous avons extrait le nombre de sujets inclus dans chaque étude ainsi que les paramètres permettant de l'estimer, et nous avons simulé le NSN à travers différentes approches statistiques. Enfin, nous avons évalué le risque de biais à l'aide d'un outil de la Cochrane (RoB2).

**PRINCIPAUX RESULTATS** Nous avons inclus 26 études de 2004 à 2021, principalement des essais de supériorité, multicentriques, avec deux groupes de comparaison. Concernant l'évaluation du risque de biais, la plupart des études ont été classées comme préoccupantes. Les définitions des VFD étaient hétérogènes, notamment pour le début et le support ventilatoire. Les méthodes pour le calcul du NSN étaient également hétérogènes dans la plupart des études. De plus, ils ne prenaient pas en compte les paramètres spécifiques des VFD comme la distribution bimodale et la mortalité en risque compétitif. Enfin, les simulations ont mis en évidence la nécessité de modèles complexes n'ayant à ce jour pas fait l'objet d'études approfondies.

**CONCLUSION** Nous recommandons une standardisation de la définition et de la méthodologie des VFD. Sélectionner la meilleure approche statistique, étant donné la distribution particulière des VFD et la mortalité comme évènement compétitif, nécessite des investigations supplémentaires. Par ailleurs, l'estimation du NSN à l'aide de simulations devrait être encouragée.

### Mots-clés :

- facteurs temps
- insuffisance respiratoire
- respirateurs artificiels
- résultat thérapeutique
- ventilation artificielle