



Critères de jugement dans les essais randomisés en soins intensifs : intérêt et limites des critères de type durée de ventilation mécanique & ventilator free-days

Laëtitia Bodet-Contentin

► To cite this version:

Laëtitia Bodet-Contentin. Critères de jugement dans les essais randomisés en soins intensifs : intérêt et limites des critères de type durée de ventilation mécanique & ventilator free-days. Médecine humaine et pathologie. Université Sorbonne Paris Cité, 2017. Français. ⟨NNT : 2017USPCC141⟩. ⟨tel-02061799⟩

HAL Id: tel-02061799

<https://theses.hal.science/tel-02061799v1>

Submitted on 8 Mar 2019

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.



HAL Authorization

Thèse de doctorat
de l'Université Sorbonne Paris Cité
Préparée à l'Université Paris Diderot

Ecole doctorale Pierre Louis de Santé Publique à Paris (ed 393)

Equipe de recherche U1153

**Critères de jugement dans les essais randomisés en
soins intensifs : intérêt et limites des critères de type
durée de ventilation mécanique & ventilator-free days**

Par Laëtitia BODET-CONTENTIN

Thèse de doctorat d'Epidémiologie et Sciences de l'Information
Biomédicale, Spécialité : Epidémiologie Clinique

Dirigée par Bruno Giraudeau
Présentée et soutenue publiquement à Tours le 29 septembre 2017

Président du jury :	Alberti Corinne/ Professeur/Université Paris Diderot
Rapporteurs :	Laviolle Bruno/ Professeur/Université de Rennes Ragot Stéphanie/ Maitre de conférence/Université de Poitiers
Examineurs :	Ehrmann Stephan/ Professeur/Université F. Rabelais de Tours Vignon Philippe/Professeur/Université de Limoges
Directeur	Giraudeau Bruno/Professeur/Université F. Rabelais de Tours

Titre : Critères de jugement dans les essais randomisés en soins intensifs : intérêts et limites des critères de type durée de ventilation mécanique et ventilator free-days

Résumé : Historiquement, dans les essais randomisés de réanimation, le critère de jugement privilégié était la mortalité à 28 jours. Cependant, la difficulté à montrer une différence inter-groupes a conduit à utiliser des critères en rapport avec la ventilation mécanique invasive connue pour sa morbidité. La durée de ventilation a tout d'abord été retenue. Mais un taux de décès élevé peut conduire à observer une durée de ventilation faible, et faussement conclure à une efficacité. Le critère « ventilator-free days, VFDs » a été proposé pour pallier cette difficulté. Il correspond au nombre de jours vivants sans ventilation. Mais ce critère pose également des difficultés d'interprétation. Les objectifs de ce travail étaient d'identifier les éléments permettant de bien définir les critères relatifs à l'exposition à la ventilation, et de discuter leur interprétation et analyse statistique.

Dans le premier travail, à l'aide d'une revue systématique, nous avons montré qu'il existe différentes définitions et avons explicité les différents items à spécifier. Dans le deuxième travail, nous avons construit une courbe iso-VFDs qui illustre que pour une même valeur de VFDs de multiples interprétations sont possibles. Puis nous avons montré à l'aide d'une base de données fictive que les analyses statistiques classiques ne sont pas adaptées pour comparer l'exposition à la ventilation entre deux groupes et qu'il est nécessaire de recourir à des analyses statistiques prenant en compte les risques compétitifs entre extubation et décès.

Mots clés : durée de ventilation mécanique invasive, nombre de jours sans ventilation mécanique, critères de jugement, réanimation, risques compétitifs.

Title : Outcomes in intensive care randomized controlled trials: interest and limitations of “mechanical ventilation duration” and “ventilator-free days”

Abstract : The primary outcome of randomized controlled trials in intensive care was usually “mortality at day 28”. However, with reduction of mortality in intensive care units, applying an effective intervention to reduce mortality even more is difficult. Therefore, outcomes related to exposure to mechanical ventilation are used. “Mechanical ventilation duration” (MVD) was first used but raises problems when the mortality is high: the intervention might be considered effective (i.e., reduction of MVD) when actually this reduction is due to increased mortality. “Ventilator-free days” (VFDs) combines mortality and MVD. However, interpreting VFDs is challenging. This thesis collected the different outcomes definitions related to mechanical ventilation used in randomized controlled trials and analyzed the different interpretations of the VFDs outcome and the different approaches for their statistical analysis. We performed a systematic review of articles published over 10 years and found all definitions for outcomes related to mechanical ventilation and identified the items that needed to be specified to define MVD and VFDs. Then we built an iso-VFDs curve, which illustrated that a VFDs value corresponded to very different combinations in terms of mortality and MVD. We showed that parametric (Student t test) and non-parametric (Wilcoxon test) statistical analyses are not adapted to compare mechanical ventilation exposure because mortality competes with extubation. Competitive risk analysis is a more suitable approach and is illustrated.

Key words : mechanical ventilation duration; ventilator-free days; outcomes; intensive care; competitive risk.

A mon mari Charles,

A mes filles Sara et Elena,

Remerciements

Je tiens, en premier lieu, à remercier Mr Bruno Giraudeau pour avoir dirigé ce travail de thèse. Merci pour ton aide, ta rigueur et tes conseils bienveillants.

Je souhaite exprimer ma gratitude envers les membres du jury qui ont accepté d'évaluer mon travail. Je remercie Mme Stéphanie Ragot et Mr Bruno Laviolle d'avoir accepté d'être les rapporteurs de ma thèse. Je remercie Mr Philippe Vignon d'avoir accepté de participer au jury. Je remercie également Mme Corinne Alberti d'avoir accepté de présider le jury. Je tiens également à remercier Mr Stephan Ehrmann qui a accepté de participer au jury et qui a également contribué à la réalisation de ce travail.

Je tiens également à remercier l'école doctorale Pierre Louis de Santé Publique, l'équipe INSERM U1153 et l'équipe du Centre d'Investigation Clinique de Tours, de m'avoir accueillie pour la réalisation de ce travail. Merci également à mes co-doctorantes Clémence, Elsa, Bénédicte. Ainsi que Fanny Feuillet, Denis Frasca et Yohann Foucher de l'équipe SPHERE pour leurs aides dans la réalisation de ce travail.

Je remercie également l'ensemble des membres de l'équipe (actuelle et passée) de médecine intensive-réanimation de l'hôpital Bretonneau à Tours et notamment l'équipe médicale le professeur Dequin et les docteurs Emmanuelle Mercier, Denis Garot, Annick Legras, Antoine Guillon, Julie Mankikian, Youenn Jouan, Charlotte Salmon, Emmanuelle Rouve, Julien Grouille pour m'avoir soutenue et donnée le temps nécessaire à la réalisation de cette thèse. Merci au Professeur Perrotin pour ses conseils.

Et je remercie pour terminer mes parents Gilles et Chantal, ma sœur Vanessa, l'ensemble de ma famille et de mes amis Edwige et Nico, Juliette et Aurore, Hélène, Guillaume, qui m'ont soutenus pour réaliser ce travail.

Laboratoire d'accueil

INSERM UMR 1153 – CRESS Centre de Recherche Epidémiologie et Statistique Sorbonne

Paris Cité

Equipe METHODS – Méthodes de l'évaluation thérapeutique des maladies chroniques

Centre d'épidémiologie Clinique

Hôpital Hôtel Dieu

1 place du parvis Notre-Dame

75181 Paris, France

Table des matières

Remerciements	5
Laboratoire d'accueil	6
Table des matières.....	7
Liste des Figures	9
Liste des Tableaux.....	10
Publications liées à la thèse	11
INTRODUCTION.....	12
1 Contexte.....	13
1.1 La réanimation	13
1.1.1 Historique.....	13
1.1.2 La réanimation en France.....	17
1.1.3 Les caractéristiques de la réanimation.....	19
1.1.4 Le coût de la réanimation	20
1.1.5 Glossaire des termes techniques de réanimation	21
1.1.5.1 Intubation	21
1.1.5.1 Trachéotomie	21
1.1.5.2 Les différents types de ventilation invasive	22
1.1.5.3 L'extubation	22
1.1.5.4 La ventilation non invasive	22
1.1.5.5 L'oxygénothérapie	23
1.1.6 Les particularités de l'extubation.....	24
1.1.7 Le « parcours ventilatoire » d'un patient.....	25
1.2 La ventilation invasive	28
1.2.1 La morbidité de la ventilation invasive	28
1.2.2 L'aspect éthique de la ventilation invasive	29
1.3 Le syndrome de détresse respiratoire aiguë (SDRA)	32
1.3.1 Historique	32
1.3.2 Définition actuelle.....	33
1.3.3 Recommandations concernant sa prise en charge	35
1.3.4 Epidémiologie actuelle.....	36
1.4 Les critères de jugement dans les essais de réanimation.....	37
1.4.1 Généralités concernant les critères de jugement.....	37
1.4.1.1 Définition	37
1.4.1.2 Propriétés métrologiques.....	38
1.4.1.3 La pertinence clinique d'un critère de jugement	39
1.4.1.4 Les différents types de critères de jugement	39
1.4.1.5 Description des critères de jugement.....	40
1.4.2 Les critères de jugement en réanimation	43
1.4.3 Le critère de jugement : mortalité	47
1.4.3.1 Définitions	47
1.4.3.2 Historique	47
1.4.3.3 Utilisation actuelle de ce critère de jugement.....	49
1.4.3.4 Problématique de ce critère de jugement.....	49
1.4.3.5 Alternatives envisagées	51
1.4.4 Le critère de jugement : durée de ventilation mécanique.....	53
1.4.4.1 Définition	53
1.4.4.2 Utilisation actuelle de ce critère de jugement.....	53

1.4.4.3	Problématique de ce critère de jugement.....	54
1.4.5	Le critère de jugement : « ventilator-free days ».....	57
1.4.5.1	Historique.....	57
1.4.5.2	Définitions.....	57
1.4.5.3	Recommandations.....	59
1.5	Analyses statistiques des critères de jugement relatif à la ventilation mécanique.....	60
1.5.1	Pratiques actuelles.....	60
1.5.1.1	Pour la durée de ventilation mécanique.....	60
1.5.1.2	Pour le ventilator-free days.....	62
1.5.2	Particularité du ventilator-free days.....	63
2	Objectifs.....	64
TRAVAUX ET RESULTATS		65
3	Revue systématique des définitions des critères de jugement « durée de ventilation mécanique » et « ventilator-free days » et impact à l'échelle individuelle	66
3.1	Présentation du travail.....	66
3.1.1	Contexte.....	66
3.1.2	Objectifs.....	66
3.1.3	Méthodes.....	66
3.1.4	Principaux résultats.....	67
3.2	Article 1.....	70
3.3	Discussion et Perspectives.....	85
4	Interprétation du critère de jugement « ventilator-free days » et impact à l'échelle individuelle et à celle d'un essai.....	88
4.1	Présentation du travail.....	88
4.1.1	Contexte.....	88
4.1.2	Objectifs.....	88
4.1.3	Méthodes.....	88
4.1.4	Principaux Résultats.....	90
4.2	Article 2.....	91
4.3	Discussion et Perspectives.....	140
DISCUSSION ET PERSPECTIVES.....		142
5	Discussion	143
5.1	Principaux résultats.....	143
5.2	Aspect de définition des critères de jugement.....	144
5.3	Aspect de gestion de l'analyse des données.....	146
6	Perspectives	147
6.1	Les autres critères de jugement en réanimation.....	147
6.1.1	Les autres « free days ».....	147
6.1.2	L'hétérogénéité des critères de jugement.....	147
6.1.3	Les critères de jugement important pour le patient.....	149
6.2	Sensibiliser et former les réanimateurs.....	150
6.3	Réfléchir aux essais de réanimation de demain	152
CONCLUSION.....		153
Références		156

Liste des Figures

Figure 1 : Le « poumon d'acier », extrait de Kelly et al. (2).....	13
Figure 2 : Le « respirateur cuirasse », extrait de Kelly et al. (2).....	14
Figure 3 : Une petite fille ventilée de manière invasive via une trachéotomie, extrait de Kelly et al. (2).....	14
Figure 4 : Le « Pulmotor », premier respirateur, extrait de Chopin (3)	14
Figure 5 : Photographie du premier respirateur Engström, extrait de Chopin (3)	14
Figure 6 : Photographie d'un patient intubé	21
Figure 7 : Photographie d'un patient trachéotomisé.....	21
Figure 8 : Photographie d'un patient sous ventilation non invasive dont l'interface est un masque bucconasal.....	22
Figure 9 : Les 3 types d'interface pouvant délivrer de l'oxygène	23
Figure 10 : Photographie d'un patient sous oxygénothérapie à haut débit.....	23
Figure 11 : Algorithme d'extubation en réanimation, extrait de la recommandation formalisée d'experts « intubation et extubation » des sociétés savantes SRLF et SFAR (12).....	24
Figure 12 : Schéma d'un parcours ventilatoire d'un patient hospitalisé en réanimation	27
Figure 13 : Schéma des recommandations de prise en charge du syndrome de détresse respiratoire aiguë, extrait de Ferguson et al. (28).....	35
Figure 14 : Classification des critères de jugement selon la source de l'information, extrait de Acquadro et al. (36).....	39
Figure 15 : Les items à spécifier pour un critère de jugement, extrait de Zarin et al. (39) ..	41
Figure 16 : Evolution de la mortalité rapportée dans les essais randomisés concernant le syndrome de détresse respiratoire aiguë, extrait de Spragg et al. (47).....	50
Figure 17 : Exemple de calcul du nombre de jours sans ventilation invasive (VFDs) en fonction des différentes propositions faites par Schoenfeld et al. (57)	59
Figure 18 : Histogramme du critère de jugement « "ventilator-free days" chez des patients présentant un syndrome de détresse respiratoire aiguë, extrait de Bellani et al. (31) ..	63
Figure 19 : Synthèse des différents items à prendre en compte dans la définition des critères de jugement concernant la ventilation invasive (durée de ventilation invasive et nombre de jours sans ventilation).....	69
Figure 20 : Représentation des différentes analyses statistiques en présence de risque compétitif, extrait de Alberti et al. (62)	140
Figure 21 : Présentation du risque de surestimation de l'événement en présence de risque compétitif lors d'une analyse par Kaplan Meier par rapport à une analyse en incidence cumulée, extrait de Alberti et al. (62).....	141
Figure 22 : Exemple d'utilisation d'une méthode d'analyse statistique en risque compétitif, extrait de Mekontso Dessap et al. (61).....	141
Figure 23 : Les « failure free days », extrait de Bernard et al. (55).....	147

Liste des Tableaux

<i>Tableau 1 : Critères de jugement des patients ventilés ayant un diagnostic de syndrome de détresse respiratoire aiguë, extrait de Bellani et al. (31).....</i>	<i>36</i>
<i>Tableau 2 : Temps auquel la mortalité est relevée, extrait de (51).....</i>	<i>48</i>
<i>Tableau 3 : Résultats d'un essai fictif démontrant la problématique d'interprétation d'une valeur isolée de durée de ventilation mécanique.....</i>	<i>55</i>
<i>Tableau 4 : Analyses statistiques utilisées pour le critère de jugement "durée de ventilation mécanique" (101 articles), extrait de (51).....</i>	<i>61</i>
<i>Tableau 5 : Analyses statistiques utilisées pour le critère de jugement "ventilator-free days" (55 articles), extrait de (51).....</i>	<i>62</i>
<i>Tableau 6 : Critères de jugement en rapport avec la ventilation invasive, extrait de Blackwood et al. (52).....</i>	<i>86</i>
<i>Tableau 7 : Hétérogénéité de la date de début et date de fin des critères de jugement « durée de ventilation invasive» et «nombre de jours sans ventilation invasive», extrait de Blackwood et al. (52).....</i>	<i>86</i>
<i>Tableau 8 : Hétérogénéité des critères de jugement des essais évaluant la ventilation, extrait de Blackwood et al. (52).....</i>	<i>86</i>
<i>Tableau 9 : Caractéristiques de l'essai randomisé fictif utilisées pour ce travail.....</i>	<i>89</i>
<i>Tableau 10 : Recommandations pour la rédaction des essais de réanimation, extrait de Harhay et al. (50).....</i>	<i>152</i>

Publications liées à la thèse

Ce travail s'appuie sur les publications suivantes :

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.

Bodet-Contentin L, Frasca D, Tavernier E, Feuillet F, Roquilly A, Foucher Y, Giraudeau B. Ventilator-free day outcomes can be misleading. *Accepté le 3 Juin 2017 avec révisions mineures. Critical Care Medicine.*

Ces travaux ont fait l'objet de la communication suivante :

L. Contentin, S. Ehrmann, B. Giraudeau. Heterogeneity in the definition of mechanical ventilation duration and ventilator-free days. *Réanimation* (2013) 23 :S63-S66.

Le travail suivant est en rapport avec cette thèse :

D. Frasca, F. Feuillet, L. Contentin, K. Asehnoune, A. Roquilly, O. Mimos, B. Giraudeau, Y. Foucher. Le *ventilator-free days* comme critère de jugement dans les études de recherche clinique en réanimation : proposition d'une approche multi-état alternative. *Revue d'Epidémiologie et de Santé Publique* 64S (2016) S137-S161.

INTRODUCTION

1 Contexte

1.1 La réanimation

1.1.1 Historique

La réanimation est née durant l'épidémie de poliomyélite en 1952 à Copenhague. La poliomyélite est une maladie infectieuse contagieuse causée par le poliovirus (1). Elle peut être asymptomatique, bénigne ou mortelle par atteinte de la moelle épinière qui entraîne une paralysie des muscles respiratoires. Le traitement est uniquement symptomatique. Le « poumon d'acier » (Fig. 1), inventé en 1928, est la première technique utilisée pour faire survivre les patients atteints d'une paralysie des muscles respiratoires.

Figure 1 : Le « poumon d'acier », extrait de Kelly et al. (2)

Il s'agit d'un appareil de grande taille dans lequel le corps est introduit et qui permet de faire respirer la personne par une ventilation à pression négative. Des pompes permettent de faire varier la pression à l'intérieur de l'appareil de manière périodique afin de recréer les mouvements d'inspiration et d'expiration du patient. Son ancêtre le spirophore fut inventé par un médecin français Eugène Woillez en 1876. Son hypothèse de travail était la suivante « le meilleur moyen de rétablir la respiration serait la dépression extérieure appliquée sur les parois thoraciques ». Il vérifia tout d'abord son idée sur un modèle de poumon suspendu à l'intérieur d'un bocal de verre dont une paroi était mobile (le spiroscope). Puis il construisit le spirophore et le testa lui-même.

Dans le même temps se développe le « respirateur cuirasse » (Fig. 2) pour les patients ayant une atteinte des muscles respiratoires moins sévère, il s'agit d'une coquille que

l'on applique sur le devant du tronc et qui est reliée à une pompe permettant d'alterner une pression positive et négative.

Figure 2 : Le « respirateur cuirasse », extrait de Kelly et al. (2)

Lors de l'épidémie de poliomyélite de 1952, ces types de respirateur n'étaient pas disponibles en assez grand nombre pour tous les patients, ce qui a conduit à une mortalité importante (85 à 90%). Le Pr Lassen a cherché une solution et a emprunté à son collègue le Dr Bjorn une nouvelle technique : la ventilation à pression positive. Il pratique une trachéotomie à une petite fille de 12 ans et insuffle l'air dans ses poumons directement et manuellement à l'aide d'une poche en caoutchouc, avec succès (Fig. 3).

Figure 3 : Une petite fille ventilée de manière invasive via une trachéotomie, extrait de Kelly et al. (2)

Des étudiants en médecine se relayent alors auprès des patients jour et nuit pour leur insuffler l'air à l'aide du ballon et ainsi leur permettre de survivre.

« Le pulmotor », premier ventilateur, présenté à la Figure 4 fut créé en 1907 par Heinrich Dräger qui avait un savoir-faire familial de conditionnement de l'air comprimé en bouteille et la fabrication de pompe à bière. Ce ventilateur était donc alimenté par une bouteille d'air comprimé dans lequel l'augmentation de la pression à l'intérieur du circuit inspiratoire actionnait un système de bielles animées par un soufflet et fermait à plus de 20 cmH₂O l'arrivée du gaz permettant ainsi l'expiration.

Figure 4 : Le « Pulmotor », premier respirateur, extrait de Chopin (3)

Mais c'est l'Engström 150, présenté à la Figure 5 qui fut le premier respirateur permettant une ventilation à débit pré-réglé cyclée sur le temps avec une fréquence prédéterminée. C'est lui qui remplaça les étudiants en médecine au chevet des patients de réanimation pour leur permettre de respirer.

Figure 5 : Photographie du premier respirateur Engström, extrait de Chopin (3)

La spécialité « réanimation » est née officiellement en décembre 1953 lorsque le Dr Ibsen crée une unité spécifique pour les patients ventilés avec une infirmière par patient (4). Dans les suites de l'épidémie, certains patients gardaient une insuffisance respiratoire nécessitant une ventilation. Pour adapter la ventilation, les premières électrodes permettant d'analyser les gaz du sang sont créées permettant de monitorer la surveillance de la ventilation et d'améliorer la survie des patients insuffisants respiratoires.

Dans les années qui suivent, les unités de réanimation aux Etats-Unis et en Angleterre se développent, associées à la création et au développement des techniques de surveillance et de suppléance d'organes (2):

- Le développement de l'hémodialyse qui a connu son premier succès en 1945 (5),
- en 1966 le premier cathéter artériel pulmonaire,
- en 1971 le premier ventilateur avec microprocesseur permettant de contrôler la ventilation en pression positive : le Servo 900A (Siemens-Elcoma, 1972) (3),
- l'assistance circulatoire extracorporelle permet d'oxygéner le sang lors de défaillance cardiaque ou respiratoire, initialement créée en 1953 pour réaliser des interventions de chirurgie cardiaque (fermeture d'une communication inter-auriculaire) (6),
- la nutrition artificielle.

L'essor de la technologie a permis un développement important des différentes techniques utilisées en réanimation avec amélioration et miniaturisation des différentes machines de suppléance et surveillance. Devant les situations cliniques difficiles voir dramatiques, les considérations éthiques sur ce qu'il est possible de faire et ce qu'il est raisonnable de faire pour un patient de réanimation deviennent, dès 1975, une préoccupation essentielle (7).

1.1.2 La réanimation en France

En France, Claude Bernard peut être considéré comme le précurseur de la réanimation médicale dans la mesure où il consigne dans ses écrits l'importance du « milieu intérieur » le sang et le curare « poison du sang » entre 1850 et 1860 (Principes de Médecine Expérimentale). C'est le 1^{er} septembre 1954 que la première unité de réanimation fut créée par P. Mollaret (8) : « le centre de réanimation neuro-respiratoire de l'hôpital Claude Bernard à Paris » avec la mise en place des premiers respirateurs artificiels performants pour prendre en charge les patients atteints de poliomyélite. La première équipe qui s'est relayée auprès des patients de cette unité était constituée, d'un biologiste J-J Pocialdo, un interniste B. Damoiseau et un neuro-infectiologue M. Goulon. Très rapidement, ils comprirent que d'autres pathologies que la poliomyélite pouvaient être prises en charge par leur unité comme le tétanos, les asthmes, les états de chocs. Dans cette unité, la mortalité diminue de 70 à 45% en quelques mois. Cependant, M. Goulon et son collègue P. Mollaret sont également confrontés à la mort cérébrale qu'ils décriront en 1959 sous le terme de « coma dépassé » (9) et à des états végétatifs persistants qui conduiront aux réflexions éthiques qui suivirent.

Le nom de « Réanimation médicale » apparaît en 1954 dans l'ouvrage de J. Hamburger. Il définit la réanimation comme « l'ensemble des gestes thérapeutiques destinés à conserver un équilibre humoral aussi proche que possible de la normale au cours des états morbides aigus, quelle qu'en soit la nature » (8). Pourtant, ce terme de « réanimation médicale » est, dès son origine, controversé puisque, J. Hamburger s'était lui même posé la question de la bonne dénomination de cette nouvelle spécialité. En 2017, le terme de « médecine intensive réanimation » apparaît et remplace le terme de

« réanimation médicale » qui est difficile à comprendre pour le grand public et qui permet d'harmoniser le terme avec les pays anglo-saxons qui utilise le terme « intensive care » ou « critical care ».

La réanimation est une nouvelle spécialité médicale basée sur une approche physiopathologique et bio-clinique des patients ayant un risque vital à court terme et dont les objectifs sont la restauration et la récupération des fonctions vitales. Cette médecine s'exerce dans le cadre d'unité où un groupe de médecins de spécialités initiales différentes exerce en temps continu (garde de 24 heures). La réanimation médicale tient alors à se distinguer des soins post-opératoires (les réanimations chirurgicales) tenues par les anesthésistes et des services d'urgence.

Dès 1954, l'organisation de la formation en réanimation médicale se met en place progressivement avec des « journées de réanimation médicale de l'hôpital Necker et du centre de transfusion sanguine » jusqu'en 1958 puis « les journées de l'hôpital Claude Bernard ». Une société savante est créée en 1971 : la SRLF (Société de Réanimation de Langue Française) (10). Les objectifs de cette société savante sont la formation continue et l'enseignement post-universitaire d'une part et la promotion de la recherche clinique et l'évaluation de l'activité de réanimation à l'hôpital. De plus, une section universitaire voit le jour en 1969.

1.1.3 Les caractéristiques de la réanimation

La réanimation se définit donc comme une unité associant les éléments suivants :

- i. Les patients présentent un pronostic vital engagé : actuellement le taux de mortalité globale varie de 10 à 29%, mais peut aller jusqu'à 45% pour le sepsis et 50% pour les insuffisances respiratoires aiguës (11).
- ii. Les patients présentent une ou plusieurs défaillances d'organes nécessitant des suppléances :
 - ✓ Une assistance respiratoire (invasive ou non) dans le cadre de défaillance des muscles respiratoires ou d'une défaillance de la commande neurologique, qui est mise en place dans 20 à 30% des admissions de réanimation (11)
 - ✓ Une épuration extra-rénale continue ou intermittente dans le cadre d'une défaillance rénale
 - ✓ Un support médicamenteux par catécholamines dans le cadre d'une défaillance hémodynamique ou cardiaque
 - ✓ Une assistance extra corporelle dans le cadre d'une défaillance cardiaque ou pulmonaire
 - ✓ Une nutrition artificielle
- iii. L'équipe médicale et paramédicale est présente 24 heures/24 et 7 jours/7 pour adapter en temps réel la prise en charge du patient (par exemple, les résultats de la gazométrie artérielle permettent de régler les paramètres du ventilateur).

1.1.4 Le coût de la réanimation

Le coût des services de réanimation sont importants (11) lié à :

- une population hospitalisée en augmentation avec à ce jour, par exemple, plus de 5,7 millions de patients admis par an aux Etats-Unis,
- des techniques innovantes,
- la charge en personnel médical et paramédical présent en continu au lit du patient.

1.1.5 Glossaire des termes techniques de réanimation

Dans ce paragraphe, nous détaillons les différents termes techniques de réanimation en lien avec la ventilation invasive nécessaire à la bonne compréhension de ce travail de thèse.

1.1.5.1 Intubation

L'intubation (Fig. 6) est la procédure de mise en place d'une sonde dans la trachée pour réaliser une ventilation invasive. Cette sonde passe entre les cordes vocales et est équipée d'un ballonnet qui permet de faire l'étanchéité pour éviter une contamination de l'arbre bronchique. Il s'agit d'une procédure à haut risque de complication (20 à 50%) pouvant mettre en jeu le pronostic vital (12).

Le lien suivant est un film animé permettant de suivre une procédure d'intubation :

<https://www.youtube.com/watch?v=m4y-SliFRFs>

Figure 6 : Photographie d'un patient intubé

.

1.1.5.1 Trachéotomie

La trachéotomie (Fig. 7) est une procédure permettant d'assurer la perméabilité permanente des voies aériennes par la mise en place d'une canule dans la trachée sous le larynx. Cette procédure peut s'effectuer soit par une technique percutanée soit par une technique chirurgicale.

Les indications en réanimation sont sujettes à controverse (13).

Figure 7 : Photographie d'un patient trachéotomisé

1.1.5.2 Les différents types de ventilation invasive

Il existe plusieurs modes de ventilation pouvant être réglés sur le respirateur (également dénommé ventilateur). Les deux modes principaux sont : les modes « contrôlés » soit en volume soit en pression, et le mode « aide inspiratoire ».

Le mode « contrôlé » en volume est défini par le réglage d'un volume constant délivré au patient.

Le mode « contrôlé » en pression est défini par le réglage d'une pression constante, le volume délivré pouvant être variable.

Le mode « en aide inspiratoire » est défini par une aide en pression apportée par le ventilateur lors de cycle spontané. Il s'agit d'un mode qui permet au patient de conserver le contrôle du déclenchement et de la fréquence des cycles respiratoires.

1.1.5.3 L'extubation

L'extubation correspond à la procédure d'ablation de la sonde endotrachéale.

Les particularités liées à cette procédure sont explicitées dans le paragraphe 1.1.6.

1.1.5.4 La ventilation non invasive

La ventilation non invasive (Fig. 8) correspond aux méthodes d'assistance ventilatoire mécanique non invasive c'est à dire n'ayant pas recours à l'abord endotrachéal (intubation ou trachéotomie). Elle permet d'aider le travail respiratoire du patient avec administration ou non d'un air enrichi en oxygène. L'interface est un masque bucconasal ou facial.

Figure 8 : Photographie d'un patient sous ventilation non invasive dont l'interface est un masque bucconasal

1.1.5.5 L'oxygénothérapie

L'oxygénothérapie est l'administration d'oxygène afin de maintenir une pression artérielle en oxygène normale. Elle peut être délivrée par des lunettes, un masque à moyenne ou haute concentration (Fig. 9).

Figure 9 : Les 3 types d'interface pouvant délivrer de l'oxygène

L'oxygénothérapie nasale à haut débit (Fig.10) correspond à l'administration d'oxygène humidifié associé à de l'air sous pression. Elle peut être considérée pour certains comme un mode de ventilation invasive du fait d'un effet « pression expiratoire positive ».

Figure 10 : Photographie d'un patient sous oxygénothérapie à haut débit

1.1.6 Les particularités de l'extubation

L'extubation est un moment à risque pour le patient. Il peut présenter :

- un œdème des cordes vocales entraînant une dyspnée laryngée pouvant aller jusqu'à l'arrêt respiratoire par obstruction de l'arbre respiratoire par l'œdème.
- une insuffisance respiratoire aiguë dans les heures ou jours suivant l'ablation de la sonde pouvant nécessiter une réintubation.

Avant de réaliser ce geste, il existe une procédure permettant de décider d'extuber un patient (12) (Fig. 11) :

- des critères cliniques tels que l'efficacité de la toux, l'abondance faible ou l'absence de sécrétions bronchiques, l'absence de troubles de conscience,
- un succès de l'épreuve de sevrage de la ventilation invasive (soit une épreuve de ventilation spontanée sur tube ou bien d'une épreuve en ventilation en aide inspiratoire réglée à 7 cm H₂O avec une pression expiratoire positive nulle),
- un test de fuite qui correspond au dégonflement du ballonnet et à la quantification ou à l'écoute d'une « fuite » dû au passage de l'air au pourtour de la sonde, permettant de penser qu'il n'y aura pas de dyspnée laryngée en post extubation (œdème des cordes vocales).

Figure 11 : Algorithme d'extubation en réanimation, extrait de la recommandation formalisée d'experts « intubation et extubation » des sociétés savantes SRLF et SFAR (12)

Certaines équipes administrent de manière plus ou moins systématique des corticoïdes en prévention de la dyspnée laryngée. Cette pratique reste controversée et ne permet pas de réduire le risque de réintubation de manière certaine. En revanche, elle prolonge la durée de la ventilation invasive puisque lorsque le patient présente l'ensemble des

critères de sevrage de sa ventilation invasive : la ventilation est poursuivie le temps d'administrer les corticoïdes (au moins 6 heures avant l'extubation).

Après extubation, plusieurs situations sont possibles :

- Le patient ventile spontanément en air ambiant
- Le patient ventile spontanément et une oxygénothérapie aux lunettes ou au masque est nécessaire pour lui maintenir une pression artérielle en oxygène (PaO₂) suffisante.
- Le patient nécessite un support ventilatoire non invasif, soit par ventilation non invasive, soit par oxygénothérapie à haut débit.

L'échec d'extubation c'est-à-dire la nécessité de réintuber le patient après son extubation programmée est mal codifié. En effet, la durée de la période entre l'extubation et la réintubation varie de 48 heures à 7 jours. Environ 25% des patients réintubés le sont après les 48 premières heures. L'utilisation de la ventilation non invasive ou de l'oxygénothérapie à haut débit au décours de l'extubation pouvant expliquer ces réintubations au-delà de la 48^{ème} heure.

1.1.7 Le « parcours ventilatoire » d'un patient

Le « parcours ventilatoire » d'un patient de réanimation correspond à la succession des différentes techniques pouvant être utilisées chez un patient de réanimation présentant une défaillance de sa fonction respiratoire.

L'oxygénothérapie (oxygénothérapie aux lunettes, masque moyenne ou haute concentration, oxygénothérapie à haut débit) et la ventilation non invasive peuvent être utilisées avant une intubation ou en relais de l'extubation. Un patient peut être ventilé de manière invasive à plusieurs reprises durant son séjour en réanimation (plusieurs intubations).

Exemple :

Cet exemple permettra de comprendre les différents points évoqués dans les travaux réalisés pendant cette thèse.

Un patient se présente pour une pneumopathie. Une oxygénothérapie à haut débit est mise en place pendant 2 jours. Puis il est intubé et ventilé pendant 8 jours. Il est extubé et réintubé vingt-quatre heures plus tard pour une dyspnée laryngée. Sa deuxième période de ventilation dure 10 jours car il présente une pneumopathie acquise sous ventilation mécanique. Lors de sa deuxième extubation, un relai par ventilation non invasive est réalisé. Le patient est pendant 4 jours sous ventilation non invasive dont les 2 premiers jours en continu (24 heures/24), puis un relai par oxygénothérapie aux lunettes est réalisé. Les différentes périodes sont représentées sur la Figure 12. A tout moment de son parcours ventilatoire, le patient est à risque de décéder pendant son hospitalisation en réanimation.

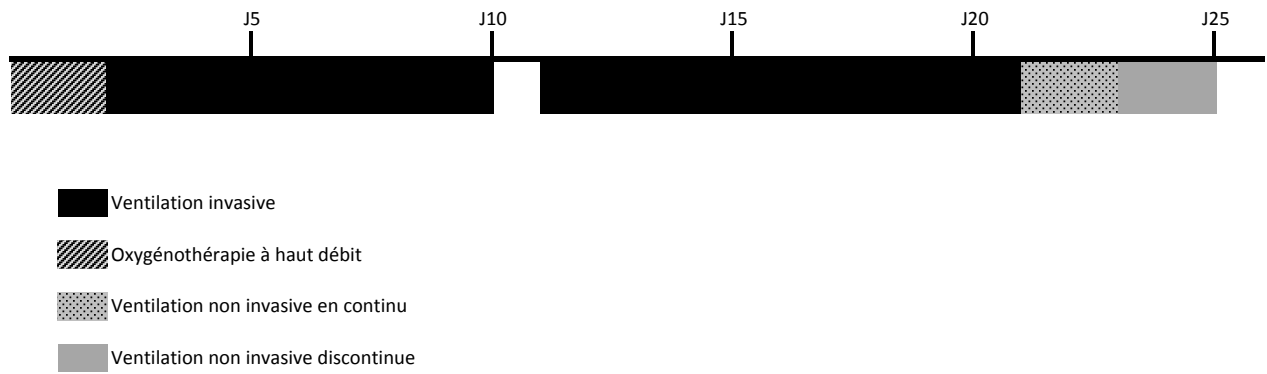


Figure 12 : Schéma d'un parcours ventilatoire d'un patient hospitalisé en réanimation

1.2 La ventilation invasive

1.2.1 La morbidité de la ventilation invasive

Les complications de la ventilation invasive (14) peuvent être liées à :

- la présence de la sonde d'intubation : sténose endotrachéale, lésions des cordes vocales par surgonflage du ballonnet, lésions cutanées sur le trajet de la fixation de la sonde
- un dysfonctionnement du matériel : obstacle sur le circuit ou débranchement du circuit pouvant conduire au décès du patient
- la ventilation invasive :
 - ✓ infectieuse : pneumopathie acquise sous ventilation mécanique
 - ✓ mécanique par baro ou volotraumatisme : pneumothorax
 - ✓ hémodynamique : hypotension artérielle
- un décubitus prolongé : thrombose veineuse profonde, fonte musculaire, escarres
- l'hospitalisation en réanimation : le syndrome post-réanimation (post intensive care syndrome)(15, 16) avec notamment altération sur la qualité de vie à long terme, difficultés de réinsertion professionnelle

Elles peuvent prolonger la durée de ventilation invasive, l'hospitalisation en réanimation ou bien à l'hôpital et elles peuvent également conduire au décès du patient.

Comme nous l'avons vu précédemment, le coût de la réanimation est important : toutes complications de la ventilation invasive va, au minimum, augmenter la durée d'hospitalisation et donc le coût du séjour du patient.

Par exemple, les patients sous ventilation invasive ont un risque de développer une pneumopathie acquise sous ventilation mécanique d'environ 10 à 70% en fonction des

population étudiée (17). Ce risque est directement dépendant de la durée d'exposition à la ventilation invasive :

- 6,5% à 10 jours,
- 19% à 20 jours,
- environ 30% à 30 jours.

Certains patients sont plus à risque de développer ce type d'infection, comme par exemple les patients présentant une bronchopathie chronique obstructive ou un syndrome de détresse respiratoire aiguë (8). Le surcoût était estimé à plus de 40 000 dollars par patient en 1999 (17, 18).

1.2.2 L'aspect éthique de la ventilation invasive

Dans certaines situations où il n'y a pas d'espoir de récupération de la fonction neurologique (anoxie cérébrale post arrêt cardiaque prolongé) ou de la fonction respiratoire (insuffisance respiratoire chronique évoluée sévère), la question de la poursuite de la ventilation invasive se pose en terme éthique : est-il raisonnable de poursuivre la ventilation invasive pour ce patient ?

La poursuite d'un traitement en l'absence d'espoir de survie ou d'une qualité de vie acceptable pour le patient va à l'encontre des quatre principes éthiques que sont l'autonomie, la bienfaisance, la non malfaisance et la justice distributive (7).

Dans ces situations, les réanimateurs doivent avoir une réflexion sur le fondement de la technique mise en place, à savoir la ventilation invasive. Après une procédure standardisée et codifiée par la loi, la décision collégiale d'arrêt de la thérapeutique peut être prise et conduire à un arrêt de la ventilation. Cette procédure va prendre un temps qui sera variable en fonction de la pathologie initiale, du patient, des complications de la réanimation, de l'avis des différents professionnels qui interviennent, de la réaction des

proches face à cette décision. Pendant ce temps, la ventilation invasive est poursuivie (19). Une fois la décision prise, deux possibilités sont décrites : soit une extubation (le terme anglophone consacré est « terminal extubation »), soit un arrêt de la ventilation sans extubation (le terme anglophone est « terminal weaning ») en modifiant progressivement les paramètres ventilatoires (diminution de la fraction inspirée en oxygène, fréquence respiratoire, volume courant, pression expiratoire positive) (20). Aucune recommandation concernant le choix de l'une des thérapeutiques par rapport à une autre n'existe actuellement.

Les limitations et arrêts des thérapeutiques en réanimation sont fréquents et interviennent pour au moins 50% des patients qui vont décéder en réanimation (19, 21, 22). Cependant, à ce jour, les articles de réanimation ne font pas la distinction entre les décès suite à une limitation ou arrêt des thérapeutiques et les décès chez les patients qui sont dans un processus de thérapeutiques curatives.

Comme nous l'avons décrit précédemment, la ventilation invasive et la réanimation sont indissociables. Le syndrome de détresse respiratoire aiguë est une entité spécifique à la réanimation et pour laquelle la ventilation invasive joue un rôle important. Un grand nombre d'essai de réanimation concernant ce syndrome ont testé des thérapeutiques en lien avec la ventilation (type de réglages du respirateur). Il nous a donc semblé nécessaire de faire un point sur ce syndrome afin que le lecteur puisse mieux appréhender les objectifs de ce travail de thèse.

1.3 Le syndrome de détresse respiratoire aiguë (SDRA)

1.3.1 Historique

Le syndrome de détresse respiratoire aiguë est nommé en anglais : Acute respiratory distress syndrome, et l'acronyme utilisé est ARDS.

Il s'agit d'une entité définie pour la première fois en 1821 par Laennec comme « un œdème pulmonaire sans dysfonction cardiaque ». C'est en 1967 qu'apparaît pour la première fois le terme « acute respiratory distress syndrome » par Ashbaugh dans le *Lancet*, indiquant qu'il s'agit d'un syndrome regroupant plusieurs étiologies, à la fois médicales et chirurgicales (4). Les premières descriptions sont faites chez des soldats américains au cours de la guerre du Viêt-Nam victimes de polytraumatisme ouvert suite à des plaies par balles qui avaient été opérés et transfusés en urgence et développaient 2 à 3 jours plus tard un tableau d'insuffisance respiratoire aiguë avec hypoxémie intense réfractaire aux hauts débits d'oxygène pur et d'opacités alvéolaires bilatérales à la radiographie thoracique (23). En 1971, Petty and Ashbaugh utilisent le terme « Adult » à la place de « Acute », probablement pour faire la distinction avec la forme pédiatrique du SDRA (24).

En 1992, la fréquence rapportée est de 150 000 cas par an aux Etats-Unis. La mortalité varie de 10 à 90% et les cas décrits sont différents entre les Etats-Unis et l'Europe. Sous la direction de la société savante américaine « American Thoracic Society » et de la société savante européenne « European Society of Intensive Care Medicine », un groupe de travail « l'American European Consensus Conference (AECC) » est alors formé et se réunit aux Etats-Unis en Floride à Miami puis en Espagne à Barcelone. L'objectif est d'améliorer la recherche clinique et épidémiologique concernant ce syndrome. Quatre sous-groupes de travail sont créés avec chacun un objectif. Le premier est en charge de

la définition du syndrome. L'AECC recommande l'utilisation du terme « Acute lung injury (ALI)» (syndrome respiratoire aigu), le SDRA devient un sous-groupe de ce syndrome pour les patients les plus hypoxémiques. Les critères définis en 1994 sont les suivants : infiltrats bilatéraux à la radiographie thoracique d'apparition brutale, une pression artérielle pulmonaire d'occlusion < 19 mmHg ou l'absence clinique d'insuffisance ventriculaire gauche et un rapport $\text{PaO}_2/\text{FiO}_2$ (rapport pression artérielle en oxygène obtenu par prélèvement sanguin artériel/fraction inspirée en oxygène réglée sur le respirateur) ≤ 300 mmHg (200 pour la définition du SRDA). Les trois autres sous-groupes traiteront : des mécanismes (sous-groupe 2), des facteurs de risque, de la prévalence et des critères de jugement pertinents (sous-groupe 3) et des mécanismes permettant de promouvoir la coordination des études cliniques (sous-groupe 4). En 1997, des experts sur le sujet sont de nouveau réunis par les sociétés savantes afin d'intégrer les résultats des recherches effectuées dans la prise en charge de ce syndrome et d'identifier les points devant faire l'objet de nouvelles études (25).

1.3.2 Définition actuelle

La définition actuelle a été publiée en 2012 (26). Il s'agit de la définition dite de « Berlin » (The Berlin Definition). Elle fait suite à une conférence de consensus auprès d'un panel d'experts qui a débuté en 2011 à l'initiative de la société savante européenne de réanimation (European Society of Intensive Care Medicine) et approuvé par les sociétés savantes américaines (American Thoracic Society and the Society of Critical Care Medicine). Le terme « Acute lung injury (ALI)» disparaît et trois classes de sévérité du SDRA sont créées. Cette nouvelle définition par rapport à celle de 1994 permet de mieux classer les patients en fonction de la sévérité de leur SDRA.

Elle repose sur quatre critères :

- 1) une insuffisance respiratoire aiguë évoluant depuis moins d'une semaine après exposition à un facteur de risque (cf. liste ci-dessous)
- 2) des opacités bilatérales sur l'imagerie thoracique
- 3) l'absence d'argument pour un œdème hydrostatique prédominant
- 4) une hypoxémie avec un rapport $\text{PaO}_2/\text{FiO}_2 < 300$ mm Hg pour une pression expiratoire positive réglée à 5 cmH₂O ou plus

Trois stades de gravité sont décrits :

- a) mineur : $200 < \text{PaO}_2/\text{FiO}_2 < 300$ mm Hg
- b) modéré : $100 < \text{PaO}_2/\text{FiO}_2 \leq 200$ mm Hg
- c) sévère : $\text{PaO}_2/\text{FiO}_2 \leq 100$ mm Hg

Le rapport PaO_2 (Pression artérielle en oxygène)/ FiO_2 (fraction inspirée en oxygène) est un index de sévérité d'une hypoxie lors de troubles du rapport ventilation/perfusion au niveau pulmonaire donnant un effet shunt (27).

Le stade mineur est autorisé chez les patients non intubés. Les stades modérés et sévères ne concernent que les patients ventilés.

Les facteurs de risque sont les suivants (27) :

- pneumopathie infectieuse
- pneumopathie d'inhalation
- noyade
- inhalation de toxique
- contusion pulmonaire
- sepsis extra-respiratoire (abdominal, urinaire..)
- traumatisme non thoracique
- transfusions multiples

- pancréatite
- circulation extra corporelle

Les deux principaux diagnostics différentiels de ce syndrome sont : l'œdème pulmonaire cardiogénique et l'exacerbation d'une pathologie pulmonaire interstitielle diffuse.

1.3.3 Recommandations concernant sa prise en charge

Lors du processus de la conférence de consensus en 2011, les experts ont également été interrogés sur la prise en charge de ce syndrome (28). Des recommandations ont été élaborées et résumées selon le schéma suivant (Fig. 13).

Figure 13 : Schéma des recommandations de prise en charge du syndrome de détresse respiratoire aiguë, extrait de Ferguson et al. (28)

« Low Tidal Volume Ventilation », en Français, la ventilation à petits volumes (6 mL/Kg de Poids idéal) est la première recommandation de la prise en charge à mettre en place lors du diagnostic de syndrome de détresse respiratoire aiguë. Cette recommandation est issue d'un essai randomisé de 2000 « petits volumes » vs. stratégie de ventilation habituelle avec des volumes à 12 mL/Kg. L'essai a été arrêté précocement car une différence significative concernant la mortalité a été observée en faveur du groupe intervention (39,8% vs 31%, $p=0.005$)(29).

D'autres essais ont testé des interventions concernant les réglages de la ventilation, comme par exemple le niveau de pression expiratoire positive. Ils n'ont pas pu démontrer une efficacité, probablement du fait d'une trop grande hétérogénéité comme l'indique la revue systématique de la Cochrane de 2013 (30).

1.3.4 Epidémiologie actuelle

En 2016, Bellani et al. ont publié les données issues d'une cohorte multicentrique prospective conduite en 2014 dans 50 pays (435 hôpitaux, 459 services). Au total, 4499 patients avec insuffisance respiratoire aiguë ont été analysés. Parmi eux, 3022 (67%) patients avaient les critères du syndrome de détresse respiratoire aiguë selon la définition de Berlin. La mortalité est évaluée en fonction de la sévérité des formes entre 34 et 46%. Dans 90% des cas, les patients sont placés sous assistance respiratoire afin de suppléer à la fonction respiratoire défaillante. Le syndrome de détresse respiratoire aiguë représente, plus de 10% des admissions sur les lits de réanimation et 23% des patients nécessitant une ventilation mécanique invasive (31). Le Tableau 1 présente les résultats des différents critères de jugement chez les patients ventilés de manière invasive en fonction de la sévérité de leur syndrome de détresse respiratoire aiguë.

Tableau 1 : Critères de jugement des patients ventilés ayant un diagnostic de syndrome de détresse respiratoire aiguë, extrait de Bellani et al. (31)

Cette cohorte prospective multicentrique a permis de montrer que le syndrome de détresse respiratoire aiguë est un problème de santé publique avec un taux de mortalité élevé et une large majorité de patients devant utiliser la ventilation invasive.

1.4 Les critères de jugement dans les essais de réanimation

1.4.1 Généralités concernant les critères de jugement

1.4.1.1 Définition

En anglais, le terme consacré est « outcome » ou « endpoint » et la définition associée est la suivante : “A component of a participant's clinical and functional status after an intervention has been applied, that is used to assess the effectiveness of an intervention” (32). Un critère de jugement permet de mesurer l’effet d’une intervention sur un patient et on considère deux types de critères de jugement (33) :

- les critères objectifs : indépendants de l’opinion du patient et de celle du clinicien, comme par exemple la mortalité ou les critères biologiques
- les critères subjectifs : basés sur l’expérience et/ou l’opinion du patient, comme par exemple la qualité de vie, ou bien sur celle du clinicien, comme par exemple, l’examen clinique d’une lésion dermatologique

Un critère de jugement doit être spécifié au moment de la rédaction du protocole de l’essai. Il peut être clinique ou paraclinique. Le critère de jugement principal doit permettre de répondre à la question de l’étude et est utilisé pour le calcul d’effectif. Il doit présenter des propriétés métrologiques qui sont définies dans le paragraphe suivant.

1.4.1.2 Propriétés métrologiques

Les différentes propriétés métrologiques attendues d'un critère de jugement sont les suivantes :

1/ Reproductibilité («reproducibility») ou fiabilité : propriété d'un test/critère dont les résultats peuvent être reproduits tant que les conditions de mesure ne changent pas (la mesure donne le même résultat à chaque fois qu'elle est répétée chez des sujets à l'état stable)

Reproductibilité inter-observateur : la mesure réalisée par des observateurs différents est la même

Reproductibilité intra-observateur : le résultat de la mesure réalisée par un même observateur ne change pas

2/ Validité (« validity ») : capacité du test/critère à mesurer ce qu'il doit mesurer, en comparaison au test de référence (« le gold standard »)

3/ Sensibilité au changement (« responsiveness ») : capacité du test/critère à détecter une différence clinique significative au cours du temps, même si la différence est faible.

Ces trois critères ont été définis en 1989 par Guyatt et al. (34). Puis en 1998, le groupe de travail de méthodologie sur les critères de jugement en rhumatologie l'OMERACT (Outcome Measures in Rheumatology) (33, 35) a ajouté :

4/ Faisabilité (« feasibility ») : la mesure doit être faisable c'est-à-dire réalisable aussi bien en terme de temps, que de coût financier.

1.4.1.3 La pertinence clinique d'un critère de jugement

La pertinence clinique d'un critère de jugement (« clinically relevant ») est fonction de la question de l'essai et de la population. Elle implique que le résultat soit directement applicable dans les soins pour les patients.

1.4.1.4 Les différents types de critères de jugement

Une façon de présenter les différents critères de jugement est la classification selon la source de l'information décrite par Acquadro et al. (36) et présentée à la Figure 14.

Figure 14 : Classification des critères de jugement selon la source de l'information, extrait de Acquadro et al. (36)

1.4.1.5 Description des critères de jugement

1.4.1.5.1 Dans le cadre de l'enregistrement des essais

L'organisation mondiale de la santé, suite à la déclaration d'Helsinki recommande que les essais cliniques soient enregistrés sur une plateforme/banque de données accessible à tous, avant l'inclusion du premier patient, dans l'objectif d'améliorer la transparence dans le domaine de la recherche. Lors de l'enregistrement sur la plateforme, des items précis doivent être renseignés. Concernant les critères de jugement, trois informations doivent être explicitées (37):

- 1/ le nom du critère de jugement (sans utiliser d'abréviation)
- 2/ la métrique ou méthode de mesure utilisée
- 3/ la date d'observation du critère

L'exemple utilisé est le suivant : 1-la dépression, 2-une échelle de dépression (Echelle de Beck) et 3-18 semaines après la fin du traitement.

En 2011, Zarin et al. (38) ont spécifié plus précisément les quatre éléments à rapporter :

- 1/ le domaine
- 2/ la méthode de mesure
- 3/ la métrique
- 4/ la méthode d'agrégation

Ils ont pris un exemple issu de la psychiatrie : l'anxiété (qui correspond au domaine) est évaluée par le Score de Hamilton (qui est une échelle spécifique correspondant à la méthode de mesure). Ils considèrent ensuite la variation brute de ce score entre la mesure finale et la mesure faite à l'inclusion (ce qui correspond à la métrique), et enfin considèrent qu'une variation de plus de 50% définit un succès (ce qui est méthode d'agrégation). Ces éléments sont indiqués sur la Figure 15.

Figure 15 : Les items à spécifier pour un critère de jugement, extrait de Zarin et al. (39)

1.4.1.5.2 Dans le cadre du « Reporting » des essais

Dans l'objectif d'améliorer la qualité de la recherche, en 1993 un groupe de 30 experts (épidémiologistes, méthodologistes, éditeurs de revue médicale) se sont réunis à Ottawa au Canada pour développer un outil pour améliorer la qualité de la présentation des résultats des essais cliniques dans les revues médicales (dénommé en anglais le « reporting »). Ils ont abouti à une recommandation qui concrètement se traduit par une liste préétablie (« check list ») de 32 items et un diagramme de flux (« flow chart »). Dans le même temps, un autre groupe de travail s'est réuni à Asilomar en Californie aux Etats-Unis et a également proposé une liste d'items à spécifier lorsque les résultats d'un essai clinique sont rapportés. En 1996 à Chicago aux Etats-Unis, l'éditeur du *Journal of the American Medical Association (JAMA)* Drummond Rennie propose aux deux groupes de se réunir (groupe EQUATOR network : Enhancing the QUALity and Transparency Of health Research) (40) et d'établir une seule liste d'items à partir des deux listes précédentes. Ce travail a abouti au Consolidated Standards of Reporting Trials (CONSORT) Statement. Depuis, plusieurs révisions du CONSORT ont eu lieu. La dernière version est le « CONSORT 2010 » (41). Concernant les critères de jugement, le CONSORT indique qu'ils doivent être complètement définis en incluant à quel moment et comment ils sont recueillis (item 6a) (42) : « *Completely defined pre-specified primary and secondary outcome measures, including how and when they were assessed* ».

1.4.1.5.3 Développer la standardisation des critères de jugement

La COMET (Core Outcome Measures in Effectiveness Trials) initiative (43) est un projet de recherche international ayant pour objectif de définir des « core outcome sets » ce qui, pour une pathologie donnée, correspond à un groupe minimal de critères de jugement à mesurer et à rapporter. L'objectif est de pouvoir améliorer la comparaison entre les essais et de pouvoir combiner les résultats dans les revues systématiques.

En réanimation, plusieurs initiatives existent pour mettre en place des « core outcome sets » dont l'un concerne la ventilation invasive : « Core Outcomes in Ventilation Trials : protocol for a core outcome set using a Delphi survey with a nested randomised trial (CoVent) » (44). A ce jour, les résultats de cette initiative ne sont pas, à notre connaissance, disponibles.

1.4.2 Les critères de jugement en réanimation

Dès 1997, l'American Thoracic Society a organisé une réunion d'experts (« workshop ») concernant les critères de jugement des essais cliniques de réanimation (45). Plusieurs éléments justifiaient cette réunion :

- le manque « d'evidence based medicine » - médecine basée sur les preuves - des soins réalisés,
- le coût financier de la réanimation,
- la validité des études non randomisées utilisées comme référence aux soins pratiqués.

Les experts indiquaient que les recherches étaient centrées sur la survie ou bien sur les désordres physiologiques, mais qu'il était important désormais d'évaluer d'autres critères et notamment ceux « centrés sur le patient » (« patient-centered outcomes ») comme par exemple la qualité de vie et le statut fonctionnel.

En 2003, dans le cadre du syndrome de détresse respiratoire aiguë, Wood et al. appelaient à améliorer la recherche sur cette thématique car la grande majorité des essais réalisés ne réussissaient pas à mettre en évidence d'effet des interventions évaluées. Ils appelaient à choisir les critères de jugement adaptés à la question et à ne pas se limiter au classique et traditionnel critère « mortalité à 28 jours » (46).

En 2009, un groupe d'experts s'est réuni à la demande du *National Heart, Lung, and Blood Institute* (NHLBI) avec l'objectif de présenter des recommandations dans la réalisation des essais concernant le syndrome de détresse respiratoire aiguë. Concernant les critères de jugement, ils ont conclu que la mortalité restait le critère essentiel mais que les mesures de qualité de vie, de statut fonctionnel et de coût étaient importantes à prendre en compte. Ils ont appelé à développer une standardisation de la manière de rapporter les critères de jugement (47).

Dans leur revue systématique parue en 2017, Gaudry et al. (48) ont, de manière arbitraire, classé en six catégories les critères de jugement qu'ils ont relevés :

1/ « Patient important outcomes » comprenant deux entités :

- la mortalité quelle que soit le moment de son évaluation
- la qualité de vie et les critères de jugement de statut fonctionnel, cognitif ou neurologique après la sortie de réanimation

2/ « Clinical outcomes in ICU and hospital » comprenant différents éléments :

- défaillances d'organes
- complications ou effets indésirables liés aux soins (par exemple : pneumopathie acquise sous ventilation mécanique, thrombose veineuse profonde)
- douleur, dyspnée en réanimation
- niveau de conscience
- tolérance des thérapeutiques, par exemple la ventilation non invasive

3/ « Biological/physiological/radiological outcomes », par exemple un dosage biologique comme le BNP (Brain Natriuretic Peptide) ou un score de sévérité radiologique

4/ « Care provider decision-related outcomes », par exemple la durée de ventilation invasive, le nombre de ré-intubations, l'exposition à la sédation (temps/dose)

5/ « Care performance outcomes » il s'agit des procédures de qualité de soins et l'exposition au bruit et à la lumière

6/ « Other outcomes » tels que la satisfaction de la famille ou des soignants, le coût, les arrêts ou limitation des thérapeutiques, la charge de travail pour l'équipe soignante, la compliance à un protocole

Actuellement, en réanimation, nous ne disposons pas de « core outcome set » standardisés.

Dans le cadre de cette thèse, nous avons étudié les critères de jugement en rapport avec la ventilation mécanique invasive. Cependant, il est essentiel pour la bonne compréhension de cette thèse de présenter le critère de jugement classique et traditionnel en réanimation « la mortalité ».

1.4.3 Le critère de jugement : mortalité

1.4.3.1 Définitions

Le critère de jugement « mortalité » est considéré depuis longtemps comme la référence (45) puisque l'objectif premier de la réanimation est de prendre en charge des pathologies mettant en jeu le pronostic vital (i.e ayant par définition un taux de mortalité élevé).

Le temps auquel ce critère est mesuré doit être précisé :

- « mortalité en réanimation »
- « mortalité hospitalière »
- « mortalité à un temps précis : 28 jours, 30 jours, 60 jours, 90 jours, 1 an »

En termes d'analyse, il y a deux approches pour gérer ces données. Alors que certains considèrent la mortalité à un temps donné comme un variable binaire (et ont donc recours à des tests statistiques comme le test du chi-deux), d'autres gèrent cette information comme une donnée de survie, estimant des courbes de survie par la méthode de Kaplan-Meier et les comparent à l'aide de test du log-rank (49).

1.4.3.2 Historique

Historiquement, dans les essais randomisés de réanimation, le critère de jugement principal est la « mortalité à 28 jours ». Cependant, dès 2003, Wood et al. se sont interrogés sur le choix de ce critère de jugement puisque des travaux ont montré que le sur-risque de mortalité en réanimation persistait après 28 jours (46). Le choix de ce « temps » doit être fonction de la pathologie et du traitement à l'étude (45). Actuellement, la mortalité évaluée à un temps choisi au-delà de 28 jours devient fréquente mais le plus souvent comme critère de jugement secondaire. Dans la revue systématique de Harhay et al. (50), parmi les 146 articles analysés, on retrouvait dans

35 (24%) d'entre eux la mortalité évaluée après le 28^{ème} jour, mais toujours comme critère de jugement secondaire et jamais comme critère de jugement principal. Dans la revue systématique que nous avons menée sur une période de 10 ans concernant les critères de jugement en réanimation, nous avons relevé les différents temps auquel la mortalité était relevée pour les 122 articles sélectionnés (51). Ces résultats sont présentés dans le Tableau 2.

Temps auquel la mortalité était relevée lorsqu'il s'agissait du critère de jugement principal (N=29)

Exclusif

Au 28 ^{ème} jour (J28)	12 (41,5)
Au 30 ^{ème} jour (J30)	1 (3,4)
Au 60 ^{ème} jour (J60)	4 (13,8)
Au 90 ^{ème} jour (J90)	2 (6,9)
Au 180 ^{ème} jour (J180)	2 (6,9)
En réanimation	3 (10,3)
A l'hôpital	5 (17,3)

Temps auquel la mortalité était relevée lorsqu'il s'agissait du critère de jugement principal (N=122)

Non exclusif

Au 28 ^{ème} jour (J28)	47 (38,5)
Au 90 ^{ème} jour (J90)	10 (8,2)
En réanimation	43 (35,2)
A l'hôpital	39 (32,0)
Non précisé	15 (12,3)
Autres	44 (36,1)

Les résultats sont présentés en nombre (pourcentage).

Tableau 2 : Temps auquel la mortalité est relevée, extrait de (51)

1.4.3.3 Utilisation actuelle de ce critère de jugement

La mortalité (tout temps t confondu) reste le critère de jugement principal des essais randomisés de réanimation :

- dans la revue systématique de 2014 de Harhay et al. (50) concernant les essais de réanimation : 40 (27%) des 146 essais analysés retenaient la mortalité comme critère principal
- dans la revue systématique de 2014 de Blackwood et al. (52) concernant les essais chez des patients ventilés de manière invasive 15 (22%) des 66 articles analysés retenaient la mortalité comme critère principal
- dans la revue systématique de 2017 de Gaudry et al. (48) concernant les essais de réanimation 21 (17%) des 112 articles analysés retenaient la mortalité comme critère principal

1.4.3.4 Problématique de ce critère de jugement

Seulement 4 des essais (10%) de la revue systématique de 2014 de Harhay et al. (50) ayant comme critère de jugement principal la mortalité ont démontré un bénéfice pour l'intervention évaluée. Cette difficulté à démontrer un effet bénéfique avait déjà été pointée par le groupe d'experts de 2009 travaillant sur la méthodologie des essais de réanimation, lors de la réunion (« workshop ») organisée par la NHLBI (47). L'un des éléments alors évoqué par le groupe d'experts comme pouvant expliquer que les essais de réanimation sont le plus souvent négatifs (dans le sens où ils n'arrivent pas à démontrer un bénéfice pour l'intervention testée) est la diminution globale de la mortalité en réanimation. En effet, le constat a été fait que dans les essais concernant le syndrome de détresse respiratoire aiguë, la diminution du taux de mortalité est nette (Fig. 16). Avec une mortalité globale plus faible, la différence escomptée (et donc l'effet

attendu d'une intervention) est plus faible, ce qui induit une majoration du nombre de patients à inclure dans les essais et questionne la faisabilité de l'essai.

Figure 16 : Evolution de la mortalité rapportée dans les essais randomisés concernant le syndrome de détresse respiratoire aiguë, extrait de Spragg et al. (47)

1.4.3.5 Alternatives envisagées

Une des réflexions des experts de la NHBLI durant cette réunion de 2009 a été de réfléchir aux autres critères de jugement possibles (47). Ils ont conclu que la mortalité devait rester le critère de jugement essentiel dans les essais concernant le syndrome de détresse respiratoire aiguë. Ils ont aussi évoqué les autres critères de jugement « importants pour le patient » tels que la qualité de vie, la fonction pulmonaire ou cognitive. Ils ont noté que les critères de jugement en lien avec les défaillances d'organes sont moins importants surtout s'ils ne s'accompagnent pas d'un impact sur la mortalité ou sur la qualité de vie.

Dans le cadre de cette thèse, nous nous sommes intéressés aux critères de jugement en lien avec la ventilation invasive, technique de suppléance du poumon qui est à l'origine de la naissance de la réanimation. Cette technique de suppléance est essentielle dans le syndrome de détresse respiratoire aiguë que nous avons exposé précédemment. Notre attention s'est portée sur la durée de ventilation invasive et sur le nombre de jours sans ventilation invasive (connu sous le terme anglophone de « ventilator-free days »).

Dans la suite du manuscrit, nous utiliserons le terme « ventilator-free days » pour le nombre de jours sans ventilation mécanique invasive.

Le terme de ventilation mécanique habituellement utilisé dans la littérature fait référence à la ventilation mécanique invasive.

1.4.4 Le critère de jugement : durée de ventilation mécanique

1.4.4.1 Définition

Il s'agit de la période entre l'intubation et l'extubation du patient, période pendant laquelle la fonction pulmonaire du patient sera en tout ou partie assurée par une machine nommée respirateur ou ventilateur.

1.4.4.2 Utilisation actuelle de ce critère de jugement

La durée de ventilation mécanique est un critère de jugement fréquent dans les essais.

Dans l'étude de Harhay et al. (50) de 2014, parmi les 146 articles analysés,

- 30 articles (21) avaient un critère de jugement principal en lien avec la ventilation.
- 55 articles (38%) avaient un critère de jugement secondaire qui était la durée de ventilation mécanique invasive.

1.4.4.3 Problématique de ce critère de jugement

1.4.4.3.1 Relation avec la mortalité

Ce critère de jugement sera impacté par une mortalité importante qui peut créer une diminution « artificielle » de la durée de ventilation invasive.

Par exemple, si on considère (Tableau 3) :

- un essai randomisé avec 200 patients par groupe
- l'intervention (par exemple un nouveau traitement) est délivrée au groupe A. Le groupe B reçoit le placebo.
- les critères de jugement (mortalité et durée de ventilation mécanique) sont évalués à 28 jours
- un taux de mortalité de 40% dans le groupe A et de 30% dans le groupe B
- une durée de ventilation mécanique identique chez les survivants de 17 jours.
- une durée de ventilation mécanique identique chez les décédés de 4 jours.
- Le résultat sera le suivant en termes de durée moyenne de ventilation mécanique : 11,8 jours dans le groupe A et 13,1 jours dans le groupe B.

En pratique, si le résultat de durée de ventilation mécanique est lu de manière isolée, la conclusion sera que l'intervention est efficace (si la significativité est obtenue) avec une durée de ventilation moindre dans le groupe A par rapport au groupe B, alors même que les patients décèdent plus dans le groupe A (groupe intervention) que dans le groupe B. Il y a donc un risque de conclure à tort que l'intervention est efficace.

	Groupe Intervention (A) (N=200)	Groupe Placebo (B) (N=200)
Mortalité	80 (40%)	60 (30%)
Moyenne de durée de ventilation mécanique (jours)	11,8	13,1

Tableau 3 : Résultats d'un essai fictif démontrant la problématique d'interprétation d'une valeur isolée de durée de ventilation mécanique

1.4.4.3.2 Prise en compte des ré-intubations

Comme nous l'avons vu précédemment avec l'exemple du parcours ventilatoire d'un patient (Fig. 12), un patient qui a été extubé peut être de nouveau intubé (ré-intubation) au cours d'un même séjour en réanimation. Il faudra alors déterminer si l'ensemble des périodes de ventilation invasive s'additionne au cours du séjour ou bien si uniquement la première période de ventilation doit être prise en compte. Dans l'hypothèse où les périodes de ventilation invasive seront comptabilisées, il faudra également déterminer si le temps entre les périodes de ventilation non invasive doit être pris en compte dans le cadre de l'étude comme une période de ventilation invasive, soit du fait de l'échec de l'extubation, soit du fait d'une suppléance de la fonction respiratoire par un autre support que la ventilation invasive, par exemple une ventilation non invasive continue avec certains paramètres de réglages (par exemple, une aide inspiratoire importante). Il s'agirait alors plutôt de parler d'une dépendance à la ventilation.

1.4.4.3.3 Un critère objectif ?

Le critère de jugement « durée de ventilation mécanique invasive » semble être un critère objectif puisque la durée de cette période est fixée par la mise en place jusqu'au retrait de la sonde d'intubation. Cependant, comme nous avons pu le voir précédemment de nombreux éléments (choix du moment pour réaliser l'extubation,

manière de considérer les suppléances ventilatoires mises en place après l'extubation comme la ventilation non invasive) interviennent et peuvent modifier cette période. Il s'agit donc plutôt d'un critère de jugement objectif mais potentiellement influencé par le clinicien (« outcomes objectively measured but potentially influenced by clinician judgement ») (53, 54).

1.4.5 Le critère de jugement : « ventilator-free days »

1.4.5.1 Historique

R.J Mangialardi et G.R. Bernard ont proposé un critère de jugement prenant en compte à la fois la mortalité et la durée de ventilation mécanique invasive, nommé le nombre de jours sans ventilation mécanique « ventilator-free days ». Ce concept de « free days » c'est-à-dire « nombre de jours sans » apparaît initialement dans un essai contrôlé randomisé concernant le syndrome de détresse respiratoire aiguë (55). Les auteurs ont dans le même temps rédigé un chapitre dans un livre destiné à une mise au point sur le syndrome de détresse respiratoire aiguë dans la collection mise au point en réanimation « update in intensive care » (56). Ils justifient l'utilisation de ce critère en indiquant qu'il permet de contenir deux informations et de réduire le nombre de patients nécessaire pour démontrer l'effet d'un traitement sur la dépendance à la ventilation invasive.

1.4.5.2 Définitions

La définition donnée dans le livre du nombre de jours sans ventilation invasive est la suivante : « Une période prédéfinie, classiquement 28 jours doit être spécifiée avant le début de l'essai. Chaque patient aura un point d'attribué chaque jour de cette période lorsqu'il sera vivant et libre de la ventilation invasive. Un patient qui est extubé le deuxième jour et reste vivant jusqu'à la fin de la période aura un score de 26, alors qu'un patient qui décède ventilé le deuxième jour aura un score à zéro (56).

En 2002, Schoenfeld et al. (57) ont expliqué comment réaliser un calcul d'effectif avec ce critère. Pour cela, ils ont précisé la définition en l'écrivant ainsi :

Le nombre de jours sans ventilation invasive (VFDs pour ventilator-free days) est défini comme suit :

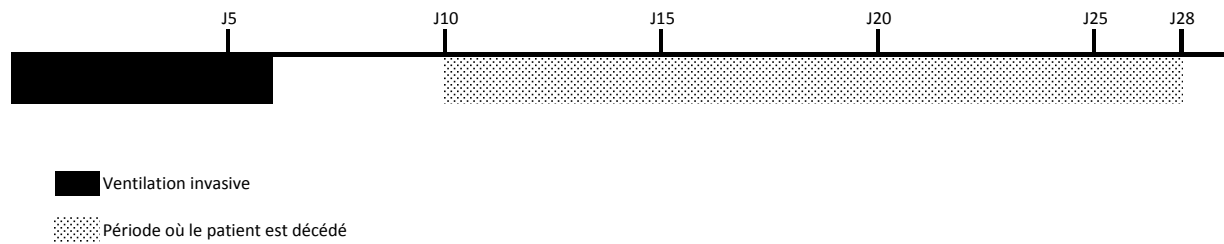
- VFDs = 0 : si le patient décède avant 28 jours
- VFDs = (28-x) : si la patient est sevré avec succès de son respirateur avant 28 jours, avec x le nombre de jours pendant lequel il a été sous ventilation invasive
- VFDs = 0 : si le patient est sous ventilation invasive à 28 jours ou au-delà.

Il proposaient une définition alternative pour ce critère de jugement concernant la manière de compter le score de VFDs pour les patients décédés ayant été sevrés de la ventilation invasive avant de décéder : VFDs = nombre de jours entre l'extubation et le décès (c'est-à-dire la période où le patient est vivant sans ventilation invasive).

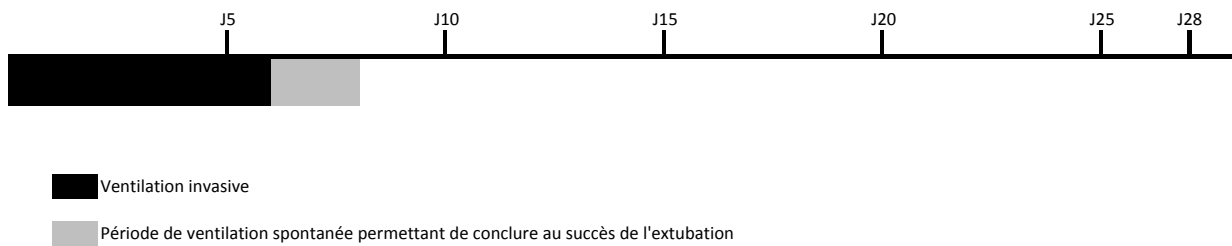
Il définissaient une « extubation réussie » : le patient doit pouvoir respirer sans l'aide du respirateur pendant une période de 48 heures consécutives.

Ils proposaient deux façons de comptabiliser la période « libre de ventilation », soit on comptabilise les VFDs à partir de la fin de la période des 48 heures, soit on comptabilise les VFDs à partir du dernier jour de ventilation invasive.

Pour que ces différentes possibilités de calcul des VFDs soient bien comprises du lecteur, nous avons ajouté un exemple présenté Figure 17.



Le patient est décédé à J10 et a été ventilé jusqu'à J6 : $VFDs = 0$ ou $VFDs = 10 - 6 = 4$



Le patient est vivant à J28 et a été ventilé jusqu'à J6 : $VFDs = 28 - 6 = 22$ ou $VFDs = 28 - 6 - 2 = 20$

Figure 17 : Exemple de calcul du nombre de jours sans ventilation invasive (VFDs) en fonction des différentes propositions faites par Schoenfeld et al. (57)

1.4.5.3 Recommandations

En 2011, un consensus d'experts australiens et néo-zélandais a établi des recommandations concernant les critères de jugement à utiliser dans les essais de réanimation de phase II (58). Ils indiquaient dans la recommandation 6 que si un critère de jugement « nombre de jours sans (free-days) » était utilisé, il fallait :

- indiquer la durée du nombre de jours où le patient est dépendant de la thérapeutique (thérapeutique au sens large, support comme par exemple la ventilation invasive ou médicamenteux comme par exemple l'utilisation des catécholamines)
- rapporter les résultats pour l'ensemble des patients, pour les survivants et pour les décédés.

1.5 Analyses statistiques des critères de jugement relatif à la ventilation mécanique

1.5.1 Pratiques actuelles

Les données suivantes sont issues de notre revue systématique de la littérature (51). Cette revue s'est intéressée à une période de 10 ans (janvier 2000 à décembre 2011) des essais contrôlés randomisés de réanimation adulte dans 10 revues à facteur d'impact élevé (6 revues généralistes et 4 revues de réanimation), avec les mots clés suivants : « mechanical ventilation » et « ventilator-free days ». Nous avons relevé pour chacun des critères de jugement « durée de ventilation mécanique » et « ventilator-free days » l'analyse statistique utilisée.

1.5.1.1 Pour la durée de ventilation mécanique

Dans notre revue systématique, la durée de ventilation mécanique est indiquée dans 101 articles parmi les 122 sélectionnés. Les résultats concernant les analyses statistiques utilisées dans ces articles sont présentés dans le Tableau 4.

Critère de jugement considéré comme une variable continue	
Approches paramétriques	29 (28,7)
Test de Student	24 (23,4)
ANOVA	4 (4,0)
Régression linéaire	2 (2,0)
Approches non paramétriques	24 (23,4)
Test de Mann Whitney/Wilcoxon	20 (19,8)
Test de Breslow Gehan Wilcoxon	1 (1,0)
Non précisées	2 (2,0)
Critère de jugement considéré comme une variable de survie	
Test du Log Rank	5 (5,0)
Modèle de Cox	4 (4,0)
Risque compétitif	1 (1,0)
Ambigu	2 (1,9)
Non précisé	37 (36,6)

Les données sont présentées en Nombre (pourcentage).

Tableau 4 : Analyses statistiques utilisées pour le critère de jugement "durée de ventilation mécanique" (101 articles), extrait de (51)

1.5.1.2 Pour le ventilator-free days

Dans notre revue systématique, la durée de ventilation mécanique est indiquée dans 55 articles parmi les 122 sélectionnés. Les résultats concernant les analyses statistiques utilisées dans ces articles sont présentés dans le Tableau 5.

Critère de jugement considéré comme une variable continue	
Approches paramétriques	19 (34,5)
Test de Student	10 (18,1)
ANOVA	6 (10,9)
Régression linéaire	1 (1,8)
Modèle généralisé linéaire suivant une distribution de Poisson	1 (1,8)
Régression logistique	1 (1,8)
Approches non paramétriques	24 ()
Test de Mann Whitney/Wilcoxon/Kruskall	23 (41,8)
Wallis	
Bootstrap	1 (1,8)
Critère de jugement considéré comme une variable de survie	
Modèle de Cox	1 (1,8)
Ambigu	1 (1,8)
Non précisé	14 (25,4)

Les données sont présentées en Nombre (pourcentage).

Tableau 5 : Analyses statistiques utilisées pour le critère de jugement "ventilator-free days" (55 articles), extrait de (51)

1.5.2 Particularité du ventilator-free days

Si on considère la définition princeps (56), pour une valeur de zéro du ventilator-free days, les patients peuvent être soit décédés soit vivants sous ventilation invasive à la fin de la période considérée. Ce qui explique probablement que Bellani et al. (31) ont distingué ces deux types de patients sur l'histogramme des résultats de la VFD (Fig. 18).

Figure 18 : Histogramme du critère de jugement « "ventilator-free days" chez des patients présentant un syndrome de détresse respiratoire aiguë, extrait de Bellani et al. (31)

La distribution de probabilité des variables de type « free-days » est donc typiquement une distribution mixte, avec un pic en zéro, et une distribution continue. Or, comme on a pu le voir en section 1.5.1.2, cette particularité n'a à l'évidence pas été prise en compte dans les stratégies d'analyse statistique classiquement retenues.

2 Objectifs

Dans un essai randomisé de réanimation, s'intéresser à la durée d'exposition à la ventilation invasive des patients est cliniquement pertinent. Ainsi qu'explicité, la difficulté à laquelle on est confronté est le décès d'une proportion non négligeable de patients, alors qu'ils sont encore intubés. Actuellement, l'exposition à la ventilation est classiquement quantifiée soit par la durée de ventilation, soit par le « nombre de jours sans ventilation invasive (ventilator-free days) ». L'objectif principal de cette thèse est de sensibiliser la communauté médicale de réanimation aux limites de ces deux critères de jugement, et au fait que l'analyse de l'exposition à la ventilation invasive doit d'être conduite dans les règles de l'art.

Ce travail s'est déroulé en deux étapes.

La première étape de ce travail avait pour objectifs :

1/ de recueillir les différentes définitions utilisées dans les essais randomisés contrôlés en réanimation chez l'adulte,

2/ d'appliquer ces définitions à un patient fictif afin d'évaluer l'impact de définitions différentes à l'échelon individuel.

La deuxième étape de ce travail avait pour objectifs :

1/ de montrer qu'une valeur donnée de nombre de jours sans ventilation invasive (ventilator-free days) prise isolément pouvait s'interpréter de différentes façons,

2/ de montrer les limites des approches statistiques classiquement utilisées pour analyser les ventilator free-days,

3/ d'expliciter les résultats d'une analyse en risques compétitifs d'un essai réel de réanimation.

TRAVAUX ET RESULTATS

3 Revue systématique des définitions des critères de jugement « durée de ventilation mécanique » et « ventilator-free days » et impact à l'échelle individuelle

3.1 Présentation du travail

3.1.1 Contexte

Actuellement, il n'existe pas de définition reconnue par les sociétés savantes de réanimation des critères de jugement : « durée de ventilation mécanique » et « ventilator-free days ».

3.1.2 Objectifs

L'objectif était de réaliser une revue de la littérature des définitions de ces deux critères de jugement telles qu'appliquées dans les essais randomisés de réanimation. Après avoir identifié les différentes définitions utilisées, nous les avons appliquées à des patients fictifs.

3.1.3 Méthodes

Nous avons réalisé une revue systématique de la littérature référencée dans MEDLINE via Pubmed sur une période de 10 ans (janvier 2000 à décembre 2011) des essais contrôlés randomisés de réanimation adulte dans 10 revues à facteur d'impact élevé (6 revues généralistes et 4 revues de réanimation), avec les mots clés « mechanical ventilation » et « ventilator-free days ». Nous avons recueilli des données concernant les caractéristiques de l'essai (nombre de sujets, critère de jugement principal) et des données concernant la définition de chacun des deux critères de jugement (la date de début et de fin de ventilation mécanique considérée, la manière de prendre en compte différentes périodes de ventilation, la période pendant laquelle le patient est observé).

Afin de préciser l'impact des différentes définitions, nous avons choisi de prendre l'exemple de deux patients fictifs observés sur une période de 28 jours (l'un décédant pendant cette période et l'autre étant vivant à la fin de la période observée) et de calculer leurs durées de ventilation mécanique et de nombre de jours sans ventilation invasive (ventilator-free days), en appliquant les définitions identifiées.

3.1.4 Principaux résultats

Parmi les 128 articles retenus, 105 (82%) utilisaient le critère de jugement « durée de ventilation mécanique » et 55 (44%) le critère « ventilator-free days ». Le critère était défini dans 20 (19%) articles pour la durée de ventilation invasive et dans 34 (61%) pour le ventilator-free days. Lorsqu'il s'agissait du critère de jugement principal, 5 articles parmi 11 donnaient une définition de la durée de ventilation invasive et 14 parmi 15 pour le ventilator-free days. Parmi les articles donnant une définition, que ce soit en critère de jugement principal ou en critère de jugement secondaire, 9 définitions différentes étaient retrouvées pour la durée de ventilation invasive et 13 pour le ventilator-free days.

Les différents items retrouvés constituant la définition de ces critères de jugement étaient les suivants :

- la date de début de la période observée : date d'intubation ou date de randomisation,
- la définition de l'extubation : durée minimale sans réintubation permettant de considérer que le patient est sevré de la ventilation,
- la manière de considérer les suppléances ventilatoires pouvant être mises en place dans les suites de l'extubation : ventilation non invasive, oxygénothérapie à haut débit, en fonction de leur durée d'utilisation sur la journée et selon les paramètres de réglages,
- la manière de considérer les patients trachéotomisés,
- la manière de prendre en compte les différentes périodes de ventilation,
- la date de fin de la période observée : majoritairement exprimée en nombre de jours (par exemple période de 28 jours : J28, est classique),
- pour le critère de jugement ventilator-free days, la manière de calculer le critère pour les patients décédés pendant la période observée : attribuer zéro ou le nombre de jours entre l'extubation et le décès pour les patients qui décèdent après extubation.

Ces différentes définitions étaient complètes ou non pour chacun des items.

Ces différents items sont synthétisés Figure 19.

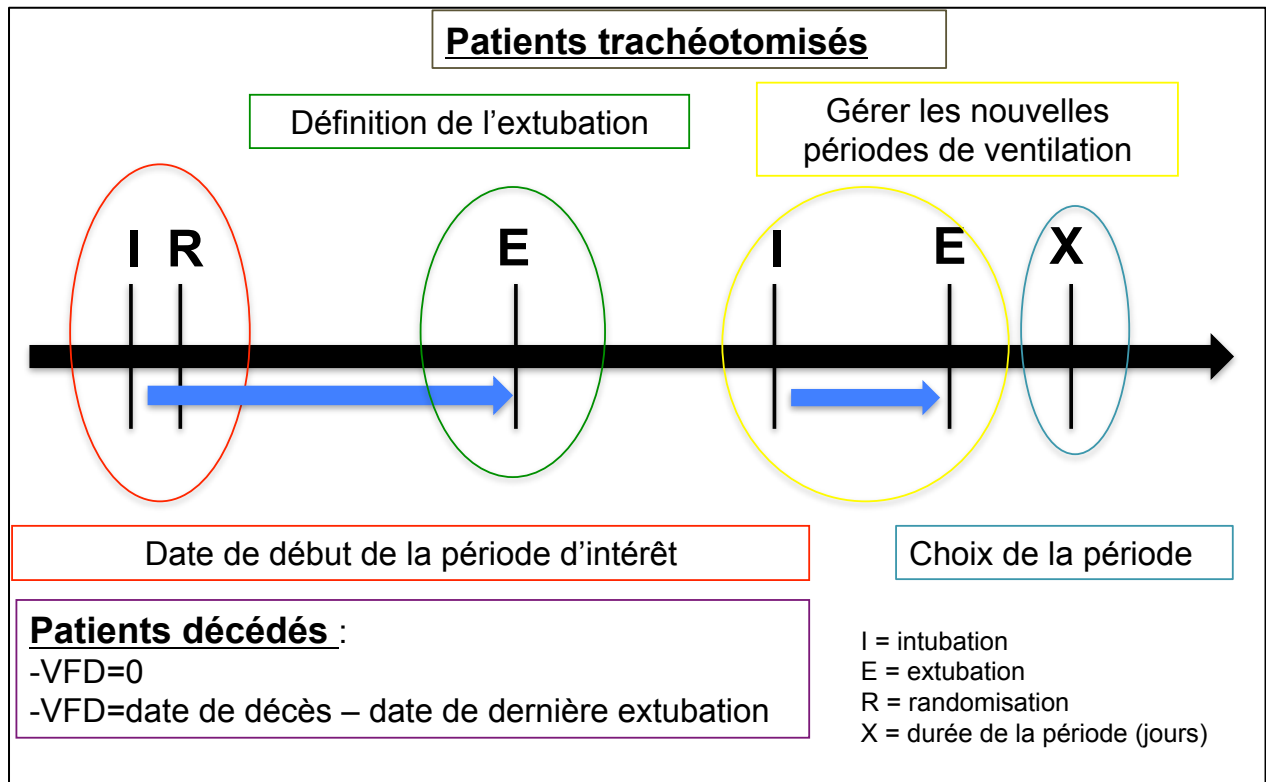


Figure 19 : Synthèse des différents items à prendre en compte dans la définition des critères de jugement concernant la ventilation invasive (durée de ventilation invasive et nombre de jours sans ventilation)

Lorsqu'on applique les différentes définitions à un même cas, la durée de ventilation invasive et le nombre de jours sans ventilation invasive (ventilator-free days) sont très variables allant respectivement de 4 à 12 jours et de 0 à 4 jours.

En résumé, nous avons donc identifié que :

- 1/ Les articles définissant ces deux critères de jugement sont peu fréquents.
- 2/ Lorsqu'une définition est indiquée, elle est dans la majorité des cas partielle : l'ensemble des items nécessaires à un calcul précis du critère de jugement ne sont pas tous indiqués.
- 3/ Les définitions retrouvées sont hétérogènes.

3.2 Article 1

Ce travail a fait l'objet d'une publication dans *l'American Journal of Respiratory and Critical Care Medicine* :

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.

CORRESPONDENCE

Heterogeneity in the Definition of Mechanical Ventilation Duration and Ventilator-Free Days



To the Editor:

An estimated one-third of patients admitted to an intensive care unit (ICU) will present with respiratory failure, thus requiring invasive mechanical ventilation (MV) to help or replace spontaneous breathing (1). This therapy is associated with high morbidity (2). Therefore, duration of MV is often used as an outcome. Death, which is frequent (an estimated one-third of patients who require MV die [2]), is a competing event to extubation. During the 1990s, exposure to MV was proposed to be assessed by an index that combines both MV duration and death status: ventilator-free days (VFDs) (3). VFDs are defined as follows: “one point [for] each day during the measurement period that [patients] are both alive and free of mechanical ventilation. A patient who is extubated on Day 2 of the study and remains alive and free of the ventilator for the remainder of the 28-day study period would receive a VFD score of 26, whereas the patient who is ventilated until death on Day 2 would receive a score of zero” (4). VFDs are now widely used as an outcome in randomized controlled trials (RCTs) (5) and are also advised as a potential primary end point (6). However, the definitions of both MV duration and VFDs are not standardized. For instance, to calculate VFDs, a dead patient may receive “a score of 0” or “the number of days between extubation and death” (7). We aimed to review the definitions of MV duration and VFDs in RCTs of adults in ICUs and to illustrate how the definitions affect individual patient values for these outcome variables.

Methods

This study was a review of reports of RCTs of adult patients in ICUs that assessed MV duration and/or VFDs as outcomes. Further specifications of the methods used are presented in the online supplement (Section E1). To illustrate how different definitions may translate into different individual values, we applied them to fictitious patients (Section E2).

Results

We retained 128 reports (Section E3).

MV duration was reported in 105 (82.0%) articles. In 20 of these articles (19.0%), the authors detailed how the MV duration was calculated, with 9 distinct definitions (Figure 1a). Among the 11 reports for which MV duration was the primary outcome, only 5 gave a definition of the outcome.

VFDs were reported in 55 (43.0%) articles. In 34 of these (61.8%), the authors detailed how VFDs were calculated, with 13 distinct definitions (Figure 1b). In four more reports, the authors specified that the VFDs for dead patients was set to zero but did not specify whether interintubation periods were taken into account

or not. In 46 of the 55 reports (83.6%), VFDs were assessed on Day 28. Other periods, when specified, ranged from 7 to 180 days in the ICU. Among the 15 reports for which VFDs were the primary outcome, 14 gave a definition of the outcome and in only 6 was a complete definition provided.

As an example, applying these different definitions to a fictitious patient with one ventilation episode and who died before Day 28 could lead to values as different as 4 to 12 days for the duration of MV and 0 to 4 days for VFDs (Section E4).

Discussion

This review of RCT reports of adults in ICUs assessing MV duration and/or VFDs reveals that such outcomes are rarely precisely defined and greatly heterogeneous in definition, which could translate into great variability in data for some patients.

For each trial, at the level of a trial, heterogeneity in definitions does not challenge the results. Indeed, internal validity is not questioned, provided that the same definition is applied to the different groups. Nevertheless, if the assessed intervention does not allow for blinding, internal validity may be challenged because of the subjective nature of MV duration and VFD outcomes (8). Moreover, if definitions are not specific enough, results may show center differences (e.g., what constitutes successful extubation). Such an issue is associated with increased variability, and the interpretation of the study results may be difficult. Finally, if outcomes are *a posteriori* defined (rather than in the protocol), we cannot exclude that the definition used was data oriented, a form of selective reporting bias (9).

However, heterogeneity in definitions raises several concerns regarding external validity. Indeed, defining and reporting outcomes is of major importance (10) because trial results must be “expressed in a standardized format to allow for comparison between studies” (11). Lack of a consensus definition for MV duration or VFDs has at least two consequences. First, when planning a new trial, variables required for the sample size calculation are usually derived from previous studies, which supposes that the same outcome was assessed. Second, meta-analyses aim to pool results from several independent studies, which implies that the same outcome was assessed in these studies. Indeed, findings of systematic reviews and meta-analyses can be affected by disagreement in data extraction (12) or multiplicity of data (13). In the same way, heterogeneity in outcome definitions may affect the conclusions of meta-analyses.

Our study has several limitations. First, our review was not exhaustive; we limited it to 10 journals with high impact factor. Nevertheless, there is no reason that the reporting of definitions may have been more complete in other journals. We may have underestimated the heterogeneity of definitions. Second, some deficiencies may occur because of poor reporting, which does not necessarily mean that definitions were absent from the protocol (14). This situation would have led to an underestimation of the proportion of trials for which a precise definition was actually used but not reported. Nevertheless, because no standardized definitions exist and given the numerous different definitions reported, heterogeneity would possibly have been even greater if the protocol did not report definitions of MV duration and VFDs were included.

Author Contributions: L.C. and B.G. designed the study. L.C. performed the study and drafted the manuscript. B.G. and S.E. provided critical review of the manuscript. All authors read and approved the manuscript.

This letter has an online supplement, which is accessible from this issue's table of contents online at www.atsjournals.org

CORRESPONDENCE

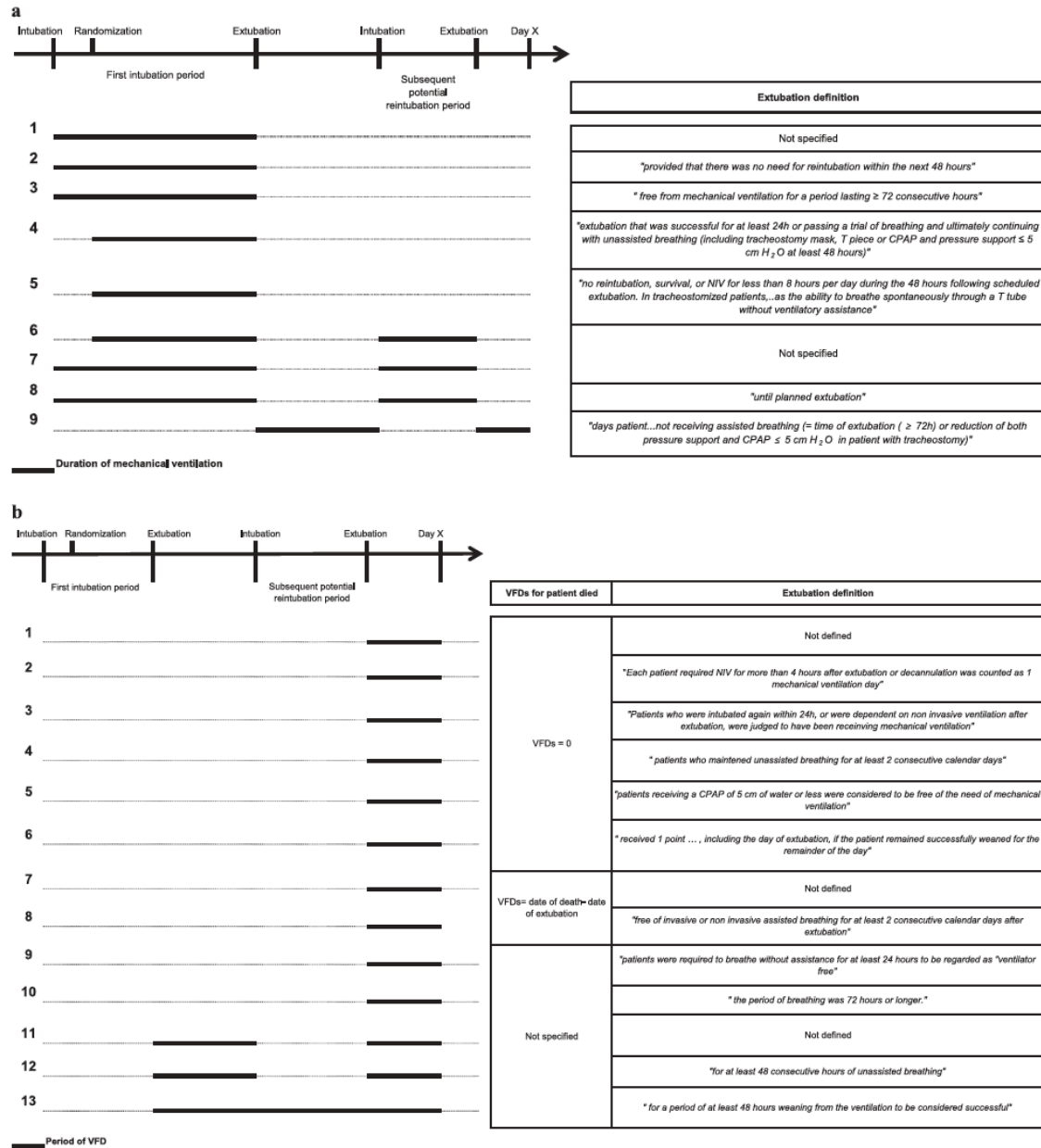


Figure 1. (a) Definitions of mechanical ventilation duration. Definitions 1, 6, and 7 are incomplete. Of note, definition 9 seems to be a ventilator-free days (VFDs) definition, although the authors explicitly stated that "the prospectively defined primary efficacy endpoint for this trial was the duration of mechanical ventilation measured by the number of days patients were alive and not receiving assisted breathing" (15). (b) Definitions of VFDs. Start of the observational period was intubation or randomization. Definitions 1, 7, and 9–13 are incomplete. CPAP = continuous positive airway pressure; Day X = observational period (e.g., Day 28); ICU = intensive care unit; NIV = noninvasive ventilation.

CORRESPONDENCE

Table 1: Items That Need to Be Defined to Calculate Invasive Mechanical Ventilation Duration and Ventilator-Free Days

Item	Issues Potentially Leading to Variability	Suggestions for Reducing Variability	Semantic Issues
Which date corresponds to the beginning of the observational period?	Randomization Intubation	From a methodological point of view, the “date of randomization” may be preferred because it corresponds to the initiation of the assessed intervention, and MV duration before randomization is expected to be balanced between randomization groups MV duration before randomization should be reported as a baseline characteristic	If intubation precedes randomization, the outcome assessed is the postrandomization MV duration rather than the entire MV duration
How to define a successful extubation?	Which minimal duration without reintubation should be considered?	48 h appears to be a reasonable option. This clinically sound cutoff is frequently used in studies to define successful extubation	<i>Extubation</i> means withdrawal of the endotracheal tube but not necessarily the end of ventilation if noninvasive respiratory support is used
How to consider noninvasive respiratory support?	Noninvasive ventilation: With two levels of assistance or continuous positive airway pressure High-flow and/or standard oxygen therapy	One solution would be to focus on MV through an endotracheal tube exclusively, thus discarding periods when another respiratory support is used However, if other respiratory supports are of interest for the study, their associated duration must be reported. As well, it seems necessary to specify the minimal period without noninvasive respiratory support (48 h appears to be a reasonable option) In the specific case of patients with chronic noninvasive ventilation, the return to chronic ventilation settings may be considered the end of acute support	If noninvasive respiratory support after extubation is considered: • For MV duration, the outcome assessed is the entire invasive and noninvasive mechanical ventilation duration • For VFDs, the outcome assessed is the invasive and noninvasive ventilator-free days
How to consider patients with tracheostomy?	MV through tracheostomy Breathe spontaneously Decannulation is performed	Permanent spontaneous breathing through a tracheostomy may be considered the end of ventilatory support. In specific studies decannulation may be considered It seems necessary to specify the minimal period without MV through tracheostomy (48 h appears to be a reasonable option)	
How to calculate VFDs for patients who die?	Assign systematically 0 d Number of days between date of extubation and date of death	A 0 value allows giving an important weight to death, which remains the most pejorative outcome	

(Continued)

In conclusion, this review of reports of RCTs of adults in ICUs, revealing that MV duration and/or VFDs are rarely precisely defined and greatly heterogeneous in definition,

calls for an international expert-based work to define MV duration and VFDs. From a practical point of view, items that need to be specified to calculate the MV duration and

CORRESPONDENCE

Table 1: (CONTINUED)

Item	Issues Potentially Leading to Variability	Suggestions for Reducing Variability	Semantic Issues
How new periods of MV are taken into account?	For MV duration: • All periods of MV are taken into account • Only the first period of MV is considered For VFDs: • All periods free of MV are taken into account • Only the last period without MV is considered	For MV duration, it seems reasonable to consider any MV episodes and to sum their respective durations For VFDs, all periods without MV should be considered In fact, every day free of the ventilator reduces the MV-associated risks such as ventilator-associated pneumonia	If only the first period of MV duration is considered, ideal reporting should spell out "duration of first period of MV" If only the period since the last successful extubation is considered, the outcome assessed should be spelled out "VFDs since the last successful weaning"
Which date corresponds to the end of the observational period?	A fixed day: e.g., 28 d, 60 d, 90 d A period: e.g., discharge from the intensive care unit	A fixed day seems a reasonable choice for the end of the observational period, thus allowing homogeneity between patients Because patients typically undergo about 1 wk of ventilator support in most intensive care units, assessing ventilator status beyond 28 d seems questionable	It is important to specify the end-point day: • Postrandomization MV duration over a 28-d period • Postrandomization VFDs on Day 28

Definition of abbreviations: MV = mechanical ventilation; VFDs = ventilator-free days.

VFDs are presented in Table 1. Such standardized definitions may also be developed for other "-free" days outcomes (e.g., vasopressor-free days) that are frequently encountered in the ICU literature. ■

Author disclosures are available with the text of this letter at www.atsjournals.org.

Laëtitia Contentin, M.D.
Institut National de la Santé et de la Recherche Médicale (INSERM) U1153
Paris, France
INSERM CIC 1415
Tours, France
Université François-Rabelais de Tours, Pôle de Recherche et d'Enseignement Supérieur (PRES), Centre-Val de Loire Université
Tours, France
and
Centre Hospitalier Régional Universitaire (CHRU) de Tours, Service de Réanimation Polyvalente
Tours, France

Stephan Ehmann, M.D., Ph.D.
Université François-Rabelais de Tours, PRES, Centre-Val de Loire Université
Tours, France
and
CHRU de Tours, Service de Réanimation Polyvalente
Tours, France

Bruno Giraudeau, M.D., Ph.D.
INSERM U1153
Paris, France
INSERM CIC 1415
Tours, France
Université François-Rabelais de Tours, PRES, Centre-Val de Loire Université
Tours, France
and
CHRU de Tours
Tours, France

References

1. Esteban A, Frutos-Vivar F, Muriel A, Ferguson ND, Peñuelas O, Abaira V, Raymontos K, Rios F, Nin N, Apezteguía C, *et al.* Evolution of mortality over time in patients receiving mechanical ventilation. *Am J Respir Crit Care Med* 2013;188:220–230.
2. Bums KEA, Adhikari NKJ, Keenan SP, Meade MO. Noninvasive positive pressure ventilation as a weaning strategy for intubated adults with respiratory failure. *Cochrane Database Syst Rev*. 2010 [accessed February 2014];8:CD004127. Available from: <http://summaries.cochrane.org/>
3. Bernard GR, Artigas A, Brigham KL, Carlet J, Falke K, Hudson L, Lamy M, Legall JR, Morris A, Spragg R. The American–European Consensus Conference on ARDS: definitions, mechanisms, relevant outcomes, and clinical trial coordination. *Am J Respir Crit Care Med* 1994;149:818–824.
4. Marini JJ, Evans TW. Acute lung injury. Heidelberg: Springer-Verlag; 1997.
5. Rubenfeld GD, Angus DC, Pinsky MR, Curtis JR, Connors AF Jr, Bernard GR; Members of the Outcomes Research Workshop. Outcomes research in critical care: results of the American Thoracic Society Critical Care Assembly Workshop on Outcomes Research. *Am J Respir Crit Care Med* 1999;160:358–367.
6. Young P, Hodgson C, Dulhunty J, Saxena M, Bailey M, Bellomo R, Davies A, Finfer S, Kruger P, Lipman J, *et al.*; ANZICS Clinical Trials Group. End points for phase II trials in intensive care: recommendations from the Australian and New Zealand Clinical Trials Group consensus panel meeting. *Crit Care Resusc* 2012;14:211–215.
7. 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;30:1772–1777.
8. Spragg RG, Bernard GR, Checkley W, Curtis JR, Gajic O, Guyatt G, Hall J, Israel E, Jain M, Needham DM, *et al.* Beyond mortality: future clinical research in acute lung injury. *Am J Respir Crit Care Med* 2010;181:1121–1127.
9. Moher D, Hopewell S, Schulz KF, Montori V, Gøtzsche PC, Devereaux PJ, Elbourne D, Egger M, Altman DG; Consolidated Standards of Reporting Trials Group. CONSORT 2010 explanation and elaboration:

CORRESPONDENCE

- updated guidelines for reporting parallel group randomised trials. *J Clin Epidemiol* 2010;63:e1–e37.
10. Zarin DA, Tse T, Williams RJ, Califf RM, Ide NC. The ClinicalTrials.gov results database—update and key issues. *N Engl J Med* 2011;364:852–860.
 11. Egger M, Smith GD, Phillips AN. Meta-analysis: principles and procedures. *BMJ* 1997;315:1533–1537.
 12. Tendal B, Higgins JP, Jüni P, Hróbjartsson A, Trelle S, Nüesch E, Wandel S, Jørgensen AW, Gesser K, Iltis-Kristensen S, et al. Disagreements in meta-analyses using outcomes measured on continuous or rating scales: observer agreement study. *BMJ* 2009;339:b3128.
 13. Tendal B, Nüesch E, Higgins JPT, Jüni P, Gotzsche PC. Multiplicity of data in trial reports and the reliability of meta-analyses: empirical study. *BMJ* 2011;343:d4829.
 14. Soares HP, Daniels S, Kumar A, Clarke M, Scott C, Swann S, Djulbegovic B; Radiation Therapy Oncology Group. Bad reporting does not mean bad methods for randomised trials: observational study of randomised controlled trials performed by the Radiation Therapy Oncology Group. *BMJ* 2004;328:22–24.
 15. Taylor RW, Zimmerman JL, Dellinger RP, Straube RC, Criner GJ, Davis K Jr, Kelly KM, Smith TC, Small RJ; Inhaled Nitric Oxide in ARDS Study Group. Low-dose inhaled nitric oxide in patients with acute lung injury: a randomized controlled trial. *JAMA* 2004;291:1603–1609.

Copyright © 2014 by the American Thoracic Society

A Novel Predictor of Cancer Malignancy: Up-regulation of Myristoylated Alanine-Rich C Kinase Substrate Phosphorylation in Lung Cancer



To the Editor:

Lung cancer currently remains the leading cause of cancer-related deaths because of its aggressive nature. The 5-year survival rates for localized and regional disease are 54 and 26%, respectively, but only 4% for patients with late-stage (stage IV) disease (1). Thus, development of biomarkers to identify patients at high risk for aggressive progression is of urgent need. Recently, we have reported myristoylated alanine-rich C kinase substrate (MARCKS), predominantly its phosphorylated state, as a risk factor associated with lung cancer invasiveness and metastasis (2). MARCKS is a substrate of protein kinase C, and also a membrane-associated protein. Upon phosphorylation at Ser159 and Ser163 within its phosphorylation site domain, phosphorylated MARCKS (phospho-MARCKS) is detached from the plasma membrane and is able to regulate various cellular processes, including cell migration and exocytic vesicle release (2–4). In the lungs, MARCKS has been extensively studied because of its role in regulating mucus secretion and inflammation. Inhibition of MARCKS activity not only reduces airway mucus hypersecretion both *in vitro* and *in vivo* (3, 5), but also represses inflammatory leukocyte migration and degranulation (6, 7). There have been limited studies on MARCKS in cancer

metastasis, but the results have been conflicting (8–13). This is because MARCKS expression is ubiquitous in various normal and tumor tissues. Despite this, there is a consensus that phospho-MARCKS, a post-translational modification, is associated with cell motility, and has a role in the regulation of cancer cell invasiveness and metastasis (2, 4, 14, 15). Of note, our laboratory discovered that inhibition of MARCKS phosphorylation was able to reduce lung cancer metastasis in murine models (2). However, the clinical significance of phospho-MARCKS in different cancers remains to be determined. In particular, there is limited information regarding its relevance in cancer progression, especially lung cancer.

Based on 18 pairs of normal and malignant lung cancer tissue sections, we previously reported that elevated phospho-MARCKS was seen in malignant lung cancer tissue sections, but not in their adjacent normal counterparts (2), suggesting a potential association between MARCKS phosphorylation and more aggressive lung cancer histological grades. To investigate more fully this previous finding, we analyzed samples from a cohort of 110 human patients with lung cancer using immunohistochemical staining with an anti-pSer159/163 MARCKS monoclonal antibody (see the online supplement). The clinical characteristics of these patients are summarized in Table 1. Consistent with our previous reports (2), high levels of MARCKS phosphorylation were found in tumor tissues compared with normal lung tissues (Figures 1A–1F). Weak phospho-MARCKS staining was observed in the cytoplasm of lung cancer tissue samples from patients in stage I (Figure 1C). In contrast, strong MARCKS phosphorylation occurred in advanced-stage lung cancer tissue samples (Figures 1D–1F). The levels of MARCKS phosphorylation correlated significantly with advanced stages of disease (Figure 1G, Pearson's chi-square test).

To quantitatively investigate these impressions, bivariate logistic regression models to predict the likelihood of high phospho-MARCKS levels from advanced tumor stages were estimated, and the probabilities of high phospho-MARCKS with stage I to III are shown in Figure 1H. The analyses demonstrated that, for a one-unit increase in stages II and III, the log odds of high expression of phospho-MARCKS levels increased by 1.00 and 2.46 compared with stage I. There were significant differences in the logistic probabilities of high phospho-MARCKS levels between stages I and II ($P = 0.039$), as well as stages I and III ($P < 0.001$), respectively. These results suggest that phospho-MARCKS may be a promising clinical predictor of tumor stages in patients with lung cancer.

Moreover, we also investigated the significance of phospho-MARCKS in lymph node status and found that higher levels of MARCKS phosphorylation correlated with lymph node metastasis (Figure 1I, N0 versus N1–2). Notably, MARCKS phosphorylation was lower in a subtype of adenocarcinoma, bronchoalveolar carcinoma, which shows a less invasive phenotype than adenocarcinoma (Figure 1J, AC versus bronchoalveolar carcinoma). Because tumor necrosis is a common event in aggressive cancers, we further checked phospho-MARCKS levels in the 10 tumor tissues with necrosis in this set of tissue arrays. Interestingly, we found higher staining intensity and increased numbers of cells stained with anti-phospho-MARCKS antibody in these tumors. These data raise the possibility that high phospho-MARCKS levels may contribute to cancer progression in non-small

Supported by National Institutes of Health grants HL077902, HL096373, and HL36982.

This article has an online supplement, which is accessible from this issue's table of contents at www.atsjournals.org

**Heterogeneity in the definition of mechanical ventilation duration
and ventilator-free days**

Contentin L, Ehrmann S, Giraudeau B

“Online Data Supplement”

Online data supplement E1: Search strategy, selection criteria, data collection, outcomes definitions and data analysis of our study.

Search strategy

We searched MEDLINE via PubMed with the keywords “mechanical ventilation” OR “ventilator free days” for reports of randomized controlled trials (RCTs) involving human adults ≥ 19 years and published in the English language from January 1, 2000 to December 31, 2011 in 10 journals with high impact factor: 6 general journals (*New England Journal of Medicine*, *Lancet*, *Journal of the American Medical Association*, *Annals of Internal Medicine*, *PloS Medicine*, and *British Medical Journal*) and 4 intensive care journals (*American Journal of Respiratory and Critical Care Medicine*, *Critical Care Medicine*, *Intensive Care Medicine*, and *Critical Care*). One author (LC) screened the titles and abstracts of retrieved articles. We excluded reports of cross-over trials and cluster randomized trials. We also excluded reports of trials for which the mean observed MV duration was < 48 hr because they evaluated “rapid” withdrawal of MV in the postoperative setting, a situation with low rates of death, but we were interested in a “long” MV duration associated with morbidity occurrences of death in the ICU. We also excluded reports that evaluated non-invasive ventilation. Reports of post-hoc analyses or intermediate analyses were not considered. In case two reports were associated to a common 2x2 factorial design trial and published concomitantly, both reports were considered.

Data collection

One author (LC) extracted and collected data from the full text of selected articles using a standardized and pilot-tested data abstraction form. At the request of reviewers, a second author (SE) performed a second extraction from the full text of selected articles using the

same data abstraction form as independently as possible from the first extraction. We resolved disagreements between the two authors (LC and SE) by discussion with the third author (BG). For each selected article, we recorded the trial type (superiority or non-inferiority), primary outcome, and number of randomized patients. The trial type was determined by sample size calculation (i.e., whether a relevant difference or a non-inferiority margin was specified). If no sample size calculation was reported, the trial type was determined from the objective stated in the report. An outcome was considered the primary outcome if it was used for sample size calculation. If no sample size calculation was reported, we considered the outcome declared by the authors as the primary outcome.

Outcome definition

We recorded whether a definition of MV duration and/or VFDs was provided (i.e., whether the report specified how such outcomes were calculated). For VFDs, we recorded also the “day” endpoint that defined the measurement period: day 28 or 60 VFDs or other.

Data analysis

Descriptive statistics were estimated.

Online data supplement E2: Impact of heterogeneity of definitions of invasive mechanical ventilation (MV) duration and ventilator-free days (VFDs) on individual patient values.

Method

To illustrate how different definitions may translate into different individual values, we considered 2 hypothetical patients, the first undergoing only one MV period and the second 2 distinct MV periods. In both cases, we considered 2 extreme situations at the end of the 28-day measurement period: being alive and without breathing assistance and death.

Online data supplement E3: Results.

Data collection Results

Among the 15 collected items, the median number of disagreements between the two reviewers was 9 (inter-quartile range: 1-11) out of the 128 selected reports, for a total of 144 disagreements. Most of them were resolved by consensus. The opinion of a third reviewer was necessary for 43 disagreements.

Figure E3: Selection of reports.

MV, invasive mechanical ventilation

*short duration of MV = mean observed MV duration was < 48 hr (evaluating “rapid” withdrawal of MV in the post-operative setting, a situation with low rates of death, but we were interested in “long” MV duration associated with morbidity occurrences of deaths in the ICU).

Table E3: General characteristics of articles of randomized controlled trials of adults in intensive care units with duration of invasive mechanical ventilation (MV) and ventilator-free days (VFDs) used as outcomes.

Characteristics	No. of articles n=128
Type of journal (high-impact-factor)	
General journal	36 (28.1)
Intensive care journal	92 (71.9)
Trial type ^a	
Superiority	100 (78.1)
Non-inferiority	3 (2.3)
Unclear	26 (20.3)
Primary outcome	
Mortality	30 (23.4)
MV duration	11 (8.6)
VFDs	15 (11.7)
Other	72 (56.3)
Outcome assessed (non-exclusive)	
Mortality	124 (96.9)
MV duration	105 (82.0)
VFDs	55 (43.0)
No. of patients randomized ^b	138 (63-324)

Data are number (percentage), unless otherwise specified.

^a For one trial, the primary objective was dual with both a non-inferiority and a superiority objective.

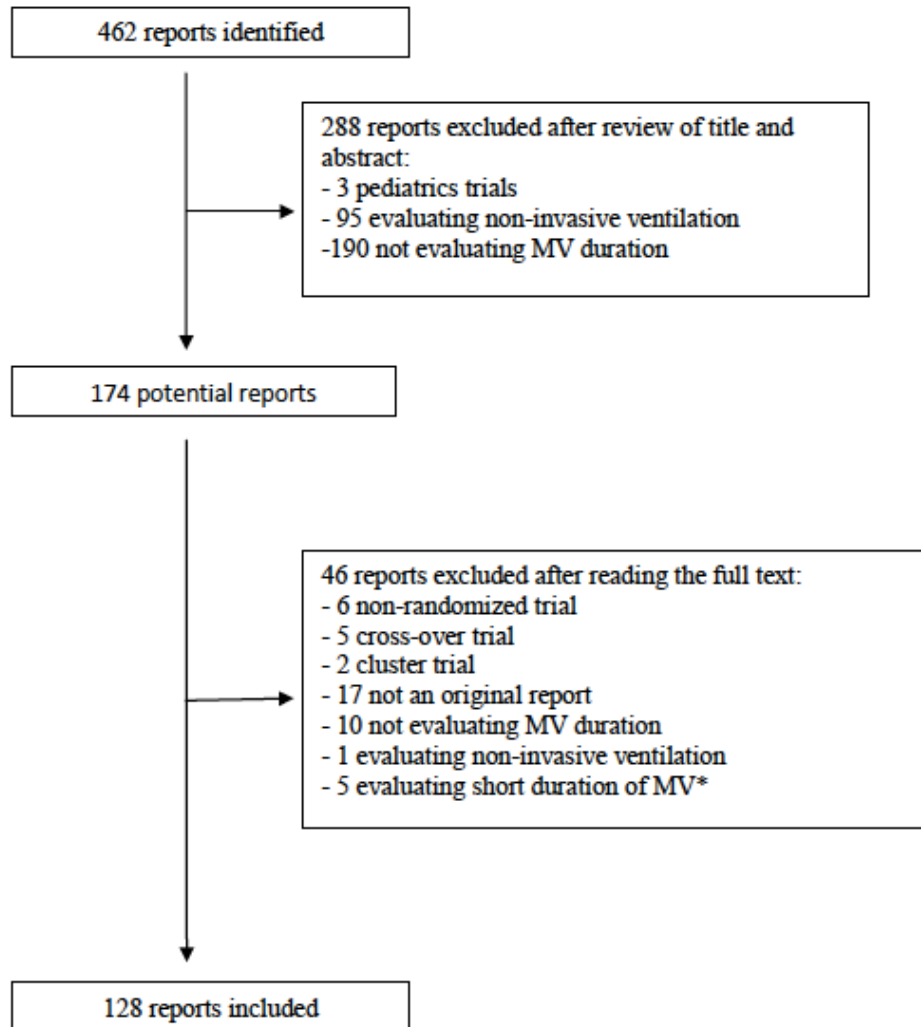
^b Median (1st and 3rd quartile).

Online data supplement E4: Impact of heterogeneity of definitions of invasive mechanical ventilation (MV) duration and ventilator-free days (VFDs) on individual patient values.

Results

Depending on the definition used, the MV duration ranged from 10 to 18 days for the first patient (Figure E4a) if the patient was alive without any breathing assistance at day 28 (the end of the measurement period) and between 4 and 12 days if the patient died before the end of the measurement period (or between 10 and 12 days if definition 9 of MV duration is not considered, cf supra). For the second patient (Figure E4b), the MV duration ranged from 10 to 19 days if the patient was alive without any breathing assistance at day 28 and between 5 and 16 days if the patient died before the end of the measurement period (or between 10 and 16 days if definition 9 of MV duration is not considered, cf supra).

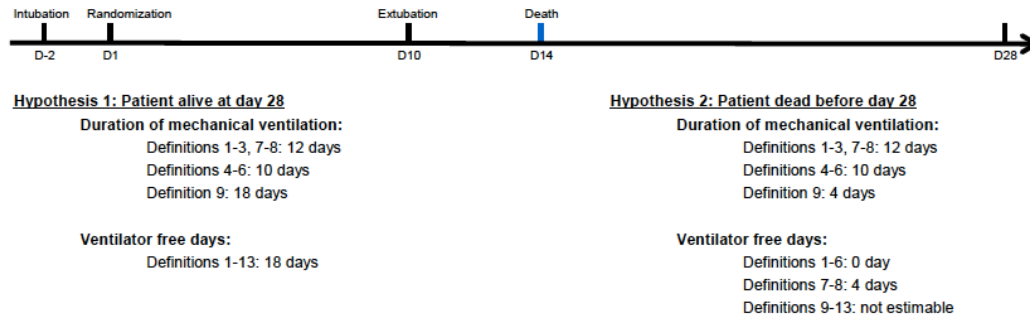
VFDs was estimated at 18 days for the first patient (Figure E4a) if the patient was alive without any breathing assistance at day 28 and ranged from 0 to 4 days if the patient died before the end of the measurement. For the second patient (Figure E4b), VFDs ranged from 6 to 18 days if the patient was alive without any breathing assistance at day 28 and was 0 days if the patient died before the end of the measurement period (and not estimable for some definitions).

Figure E3: Selection of reports.

MV, invasive mechanical ventilation

*short duration of MV = mean observed MV duration was < 48 hr (evaluating “rapid” withdrawal of MV in the post-operative setting, a situation with low rates of death, but we were interested in “long” MV duration associated with morbidity occurrences of deaths in the ICU).

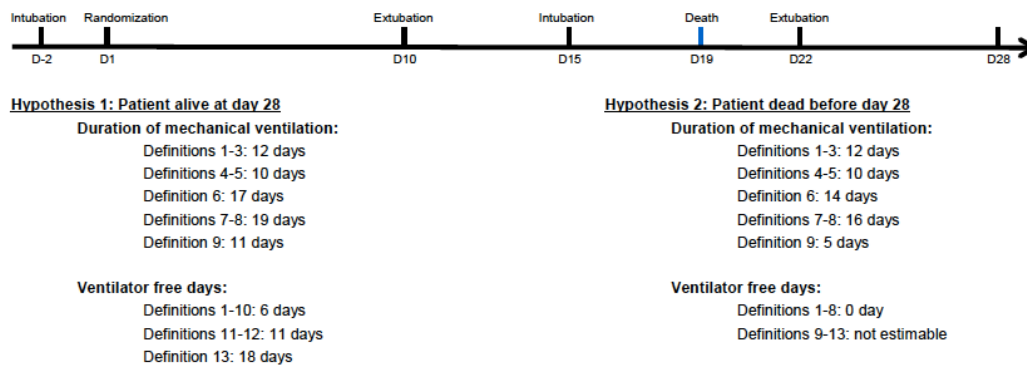
Figure E4a: Impact of heterogeneity of definitions of invasive mechanical ventilation (MV) duration and ventilator-free days (VFDs) on individual patient values: patient with only 1 MV period.



Definitions for MV duration and VFDs are reported in Figures 1a and 1b, except that we did not consider the extubation definition.

1

Figure E4b: Impact of heterogeneity of definitions of invasive mechanical ventilation (MV) duration and ventilator-free days (VFDs) on individual patient values: patient with 2 MV periods.



Definitions for MV duration and VFDs are reported in Figures 1a and 1b, except that we did not consider the extubation definition.

2

3.3 Discussion et Perspectives

La multiplicité des définitions possibles pour les critères de jugement en rapport avec la ventilation, a pour conséquence des difficultés dans l'analyse des essais randomisés de réanimation:

1/ pour interpréter les résultats de l'essai :

- lors d'études multicentriques, si tous les centres n'utilisent pas la même définition (par exemple, l'item concernant le succès de l'extubation) cela entraine un effet centre avec des difficultés d'interprétation des résultats de l'étude.
- il existe la possibilité d'utiliser une définition à posteriori en fonction des données, créant une forme de biais de sélection de l'information rapportée « reporting bias ».

2/ pour comparer les essais entre eux :

- lors de la réalisation d'un nouvel essai, le calcul d'effectif suppose l'utilisation du même critère de jugement, pour cela sa définition doit être disponible.
- lors de la réalisation de méta-analyse, l'hétérogénéité des définitions risque d'affecter les résultats de la méta-analyse.

Le problème de la définition des critères de jugement utilisés en réanimation a également été soulevé dans l'article de Bronagh Blackwood et al. (52) paru dans le même numéro de *l'American Journal of Respiratory and Critical Care Medicine*, et dans l'édito Dale M. Needham (59) associé à ce dernier article et au nôtre.

Blackwood et al ont montré qu'il existe une très grande variabilité des critères de jugement concernant la ventilation dans les essais de réanimation. Ils ont réalisé une revue systématique entre janvier 2007 et décembre 2012 de 8 revues (4 généralistes et 4 de réanimation). Ils ont analysé 66 essais randomisés. Parmi eux, ils ont retrouvé 30 critères de jugement principaux différents. Les 7 critères de jugement les plus fréquents sont les suivants : la durée de ventilation invasive, la mortalité, la durée de séjour en réanimation, la durée de séjour à l'hôpital, l'acquisition d'une pneumopathie sous ventilation artificielle, et l'arrêt d'un traitement (perfusion). Neuf critères de jugement en rapport avec la ventilation invasive ont été rapportés. Ils les ont classés en deux catégories : ceux en rapport avec une mesure de la dépendance à la ventilation et ceux en rapport avec des événements (Tableau 6). Ils ont également noté que la date de début et de fin des critères de jugement « durée de ventilation invasive » et « nombre de jours sans ventilation invasive » étaient très hétérogènes (Tableau 7). Cette hétérogénéité des critères de jugement utilisés de manière générale et concernant la ventilation persiste lorsqu'on s'intéresse uniquement aux essais évaluant la ventilation (Tableau 8).

Tableau 6 : Critères de jugement en rapport avec la ventilation invasive, extrait de Blackwood et al. (52)

Tableau 7 : Hétérogénéité de la date de début et date de fin des critères de jugement « durée de ventilation invasive » et « nombre de jours sans ventilation invasive », extrait de Blackwood et al. (52)

Tableau 8 : Hétérogénéité des critères de jugement des essais évaluant la ventilation, extrait de Blackwood et al. (52)

Ils ont également fait le même constat d'hétérogénéité ou d'absence d'information concernant la période à laquelle ou pendant laquelle le critère de jugement était évalué.

Blackwood et al. appellent de nouveau à standardiser les critères de jugement, comme l'avaient fait des groupes d'experts à la demande de la NHLBI et des sociétés savantes la Society of Critical Care Medicine et Multisociety Task Force for Critical Care Research (15, 47).

Needham (59) dans son éditto qui présente les 2 articles (Blackwood et al. et Contentin et al.) appelle :

- de nouveau à évaluer les critères de jugement utilisés en réanimation et à utiliser des critères appropriés et standardisés,
- à déterminer de manière optimale la durée de suivi,
- à ne pas négliger les censures consécutives au décès dans l'analyse des données.

4 Interprétation du critère de jugement « ventilator-free days » et impact à l'échelle individuelle et à celle d'un essai

4.1 Présentation du travail

4.1.1 Contexte

Tout d'abord, comme nous l'avons exprimé dans le paragraphe 1.5.2 la distribution des valeurs du critère de jugement « ventilator-free days » est spécifique avec un pic à zéro pouvant correspondre à des patients vivants ou décédés. Ensuite, lors de notre revue systématique, nous avons constaté une hétérogénéité des méthodes statistiques utilisées pour analyser le critère de jugement « ventilator-free days ». Cependant, nous avons noté une prédominance de l'utilisation du test de Student et du test de Wilcoxon, comme indiqué dans le Tableau 5.

4.1.2 Objectifs

L'objectif de ce travail était d'alerter les cliniciens sur les erreurs d'interprétation possible du critère de jugement « ventilator-free days ». Il s'agissait :

- d'étudier l'interprétation d'une valeur de ventilator-free days,
- d'étudier l'impact de différentes analyses statistiques possibles pour ce critère de jugement à partir d'un jeu de données fictif et d'une base de données réelles.

4.1.3 Méthodes

Pour réaliser la première partie de ce travail, nous avons utilisé une méthode analytique pour expliciter le concept de courbe isoVentilator-free days (isoVFDs) où chaque point de la courbe correspond à une combinaison (taux de décès, distribution de la durée de ventilation chez les extubés vivants) qui conduit à une même durée médiane de ventilator-free days. Nous avons pris comme exemple une médiane de ventilator-free

days de 10 jours sur une période de 28 jours (J28). Pour chaque taux de mortalité envisageable (10, 20, 30 ou 40%), nous en avons déduit la durée médiane de ventilation mécanique associée chez les vivants et avons construit la courbe.

Pour réaliser la deuxième partie de ce travail, nous avons utilisé un essai randomisé fictif réaliste dans lequel la mortalité et la durée de ventilation invasive étaient celles décrites dans la dernière étude descriptive épidémiologique du syndrome de détresse respiratoire aiguë publiée par Bellani et al. (31). Les différentes informations concernant cet exemple sont décrites dans le Tableau 9.

A J28	Groupe Contrôle (n=500)	Groupe Expérimental (n=500)
Taux de mortalité	34,2 %	21,6 %
Patient toujours sous ventilation invasive	8,8 %	18,8 %

Tableau 9 : Caractéristiques de l'essai randomisé fictif utilisées pour ce travail

Pour réaliser la troisième partie de ce travail, nous avons utilisé une base de données réelles issue de l'étude de Roquilly et al. (60) afin d'explicitier comment lire des courbes d'incidence cumulée, méthode prenant en compte les risques compétitifs.

4.1.4 Principaux Résultats

Dans la première partie du travail, nous nous sommes intéressés à la signification d'une valeur de ventilator-free days. Nous avons montré que pour une même valeur de ventilator-free days la situation clinique pour un groupe ou pour un patient peut prendre des significations différentes.

Dans la deuxième partie du travail, nous avons montré, à partir de la base de données fictive, que :

1/ le test de Student était impacté par le choix de la période de suivi (J60/J90/J180). La p-value diminue avec le temps (0,08 à J60 puis 0,004 à J90 et 0,0002 à J180), ce qui laisse penser qu'en augmentant l'horizon temporel, on finit par obtenir un test statistiquement significatif.

2/ le test de Wilcoxon était robuste (équivalence des médianes quelle que soit la période de suivi choisie). Cependant, dans notre exemple, il ne montrait pas de différence statistiquement significative, alors que les statistiques résumées des deux groupes laissaient penser le contraire.

3/ L'approche prenant en compte le risque compétitif du décès sur l'analyse de la durée de ventilation avec réalisation de courbes d'incidence cumulée d'extubation et comparaison avec un test de Gray permettait de montrer une différence.

Dans la troisième partie, nous avons utilisé une base de données réelles (60) à partir de laquelle nous avons réalisé les courbes d'incidence cumulée et le test de Gray, à titre d'exemple. Le lecteur peut ainsi déduire des courbes d'incidence cumulée la probabilité d'être extubé vivant à chaque temps souhaité (par exemple J28 et J60).

4.2 Article 2

Ce travail a été accepté pour publication dans la revue *Critical Care Medicine*.

Bodet-Contentin L, Frasca D, Tavernier E, Feuillet F, Roquilly A, Foucher Y, Giraudeau B.

Ventilator-free day outcomes can be misleading.

Ventilator-free day outcomes can be misleading

Bodet-Contentin Laëtitia, MD ^{1, 2}, Frasca Denis, MD-PhD ^{1, 3, 4},

Tavernier Elsa, PhD ^{1, 5}, Feuillet Fanny, PhD ^{1, 6}, Antoine Roquilly, MD ⁶, Foucher

Yohann, PhD ^{1, 6}, Giraudeau Bruno, PhD ^{1, 5}

1: Université de Tours, Université de Nantes, INSERM, SPHERE U1246, Tours, France

2: CHRU de Tours, Tours, France

3: Université de Poitiers, Poitiers, France

4: CHU de Poitiers, Poitiers, France

5: INSERM CIC 1415, CHRU de Tours, Tours, France

6: CHU de Nantes, Nantes, France

Corresponding author:

Bodet-Contentin Laëtitia

Service de Médecine Intensive Réanimation

CHRU de Tours

2 boulevard Tonnelé

37044 TOURS Cedex 09

Tel: +33(0)2.34.37.37.18

Fax: +33(0)2.47.39.65.36

e-mail: laetitia.contentin@univ-tours.fr

Author contribution: L.B-C. and B.G. designed the study. L.B-C., D.F. and E.T. performed the study. L.B-C. drafted the manuscript. A.R. provided data. B.G., F.F., A.R., and Y.F. provided critical review of the manuscript. All authors read and approved the manuscript.

Conflict of interest: The authors do not have any conflict of interest.

Source of support: No funding was required for this work.

Key words: ventilator-free days – competing events – intensive care – invasive mechanical ventilation duration – acute respiratory distress syndrome

Word count: 2241

Abstract

Introduction

Acute respiratory distress syndrome often requires invasive mechanical ventilation (MV), with both mortality and MV duration as outcomes of interest. The concept of ventilator-free days (VFD) has been proposed as an outcome combining these two outcomes. Here we analyzed the construction of the VFD outcome and provide a hypothetical scenario to alert physicians that such an outcome can lead to misleading interpretations.

Methods

We proposed the iso-VFD curve concept and, using an analytical development, illustrate how a median VFD value can actually result from very different combinations of death rates and MV durations. We also used a hypothetical example to compare the Student t test, Wilcoxon rank-sum (WRS) test and Gray test (which accounts for death as a competing event with extubation) in comparing exposition to MV.

Results

A median VFD value of 10 days may mean that 10% of the patients died while survivors were ventilated during a median of 14 days or that 40% died while survivors were ventilated during a median of 5 days. Changing the time horizon affected the Student t test but not the WRS result. The Gray test was more relevant than both the Student t test and WRS test in identifying differences in groups showing highly different MV duration, despite equal median VFD. This approach was also illustrated using real data.

Conclusion

Use of VFD as an outcome appears to have many drawbacks. Suitable methods of analyzing time to extubation should be preferred.

Introduction

Lung injury complicates serious medical or surgical conditions or involves a direct pulmonary insult. Acute lung injury (ALI) and acute respiratory distress syndrome (ARDS) are the most severe forms of lung injury with acute refractory hypoxemia (1). Patients with ALI and ARDS usually require invasive mechanical ventilation (MV). The mean duration of MV is 10 to 14 days, with some patients remaining ventilator-dependent for more than 3 weeks (2–5). Invasive MV duration is associated with increased morbidity, such as ventilator-associated pneumonia (6, 7). Therefore, although mortality remains the most clinically relevant outcome with ALI or ARDS, duration of invasive MV is also clinically relevant.

As a consequence, rather than considering mortality as a primary outcome, which requires a very large number of patients (typically a few thousands), some authors advised conducting smaller randomized clinical trials to assess physiologic endpoints such as those related to duration of invasive MV (8). Indeed, morbidity outcomes are expected to be more sensitive than mortality, and thus to require smaller numbers of patients to demonstrate that an intervention has a statistically significant impact. However, accounting for duration of invasive MV is not simple because many patients die before extubation. Death is a competing event; that is, if the patient died under MV, extubation can no longer be observed (9) and the time-to-extubation is unknown. Thus, the failure-free days concept was proposed, initially in 1992 in a randomized controlled trial concerning ARDS (10). These authors further detailed the ventilator-free days (VFD) concept in 1997(4):

VFD score is defined as *“one point [for] each day during the measurement period that [patients] are both alive and free of mechanical ventilation. A patient who is extubated on Day 2 of the study and remains alive and free of the ventilator for the*

remainder of the 28-day study period would receive a VFD score of 26, whereas the patient who is ventilated until death on Day 2 would receive a score of zero” (4). So, this outcome combines both mortality and duration of invasive MV in a unique quantitative score. Because VFD is a composite outcome, one has to study the metrological properties of such outcomes, and indeed, Rubenfeld *et al* (8), in 1999, recommended performing *“studies of the statistical performance and clinical significance of the [VFD] [...] before it can be recommended as a primary endpoint”*. As well, in 2002, Schoenfeld and Bernard (11) reminded that before seeking a difference in VFDs between two groups, one has to ensure that mortality rates are equivalent. However, these indications of caution seem to have been largely ignored, and today, VFD is considered a usual primary outcome, interpreted on its own, without prior consideration of mortality incidences (12, 13).

Our objective was to alert physicians that using VFD as an endpoint can lead to misleading interpretations. We also aimed to illustrate that frequently used statistical methods for this outcome, such as the Student’s *t* test and the Wilcoxon rank-sum (WRS) test, are inappropriate. We first present a hypothetical example that will be further used to illustrate our purpose, then we draw attention to the possible pitfalls when interpreting VFD values. Second, we draw attention to the impact of the choice of time horizon. Third, we draw attention to the inadequacy of usual statistical analysis methods to compare two groups. Fourth, we illustrate with real data how statistical analysis should better be performed, and results interpreted. Finally, we summarize our results and place them in the context of designing and interpreting a trial that compares two treatments in ARDS.

Hypothetical example

Let us consider a two-group parallel-randomized trial of 500 patients per group (See online data supplement 1). In the control group, the 28-day mortality rate is 34.2%, with 8.8% of the patients still intubated at this date, that is, rates close to those observed in the Bellamy *et al* study (5). In the experimental group, the 28-day mortality rate is much lower (21.6%), but more patients are still intubated at day 28 (18.8%). In this group, death is less frequent but at the price of a longer MV duration. Whatever the group, all events (death or extubation) occurred before day 60. Then, considering this dataset, we calculated the individual VFD at different time horizons (days 28, 60, 90, 180). Results are displayed in Table 1.

How to interpret a VFD value?

VFD is a quantitative variable that summarizes two pieces of information: a binary variable (whether the patient died or not during the period of interest) and a continuous variable (how long the patient has been under MV). Therefore, the presentation of a median VFD does not allow for determining the contribution of the two embedded outcomes. For instance, as illustrated in our example, a median VFD of 10 days can mean that 1) the mortality rate is 21.6% and the median duration of invasive MV among survivors is 10 days or 2) the mortality rate is 34.2% and the median duration of invasive MV among survivors is 7 days. No one would consider these two situations clinically equivalent, even though they have a similar median VFD value. There are actually an infinite number of combinations that lead to a common median VFD value, which allows for defining an iso-VFD curve (see online data supplement 2), as illustrated in Figure 1. Obviously, this situation hampers any interpretation of a median VFD value in itself.

In 2002, Schoenfeld et al. (11) reminded that when using VFD outcomes, *“ignoring mortality could result in misleading conclusions”*. Nevertheless, in practice, VFD is often interpreted without considering the mortality which is sometimes even reported considering a different time horizon than that selected for the VFD outcome [e.g., mortality at day 60 and VFD at day 28 in the study proposed by Truwit et al. (14)].

Besides the sample-level interpretation of a median VFD, an individual VFD value may be difficult to interpret on its own. Thus, for both a patient who dies before day 28 and one who is still alive but intubated at day 28, the VFD will be 0 at day 28. Obviously, these two situations are different. Furthermore, when considering two patients extubated at days 9 and 10, for instance, the associated VFDs at day 28 differ by one unit. This one unit between patient difference is the same as that observed when considering a patient who dies at day 5, for example, and a patient extubated at day 27. Considering these two very different situations similarly is obviously problematic, which illustrates the lack of interval property of VFD outcomes.

Which time horizon to select?

Schoenfeld et al. explained that to calculate VFDs *“the 28-day [time horizon] is suggested both because interventional trials in acute lung injury typically involve a 28-day treatment/follow-up period after the patient enrolls in the trial and because most, but not all, patients with acute lung injury have either died or been successfully weaned from mechanical ventilation by day 28”* (4). However, the proportion of patients under invasive MV at day 28 can be non-negligible: 8.8% in the Bellany *et al.* study (5), as in our fictive example. Thus, some authors also reported VFD at days 60, 90 or 180 (16).

Choosing the 28-day time horizon to calculate VFD even though there are still extubations occurring after this date is confusing. Indeed, we thus have two kinds of patients with a 0 VFD value: those who died and those alive but still intubated at day 28 (5). Bellani et al. (5) distinguished these two kinds of patients in their VFD histograms. Therefore, the 28-day time horizon is often too short. The time horizon should be established in light of the medical context (i.e., when one can reasonably expect that most patients are extubated or dead).

Which statistical approach should be used to compare two groups?

Because of the definition, the VFD outcome distribution is extremely asymmetric, with a peak in 0 associated with dead patients and those still under MV at the end of the study duration (5). Although Bridge et al. reminded that *“the Wilcoxon Rank-Sum should be the test of choice when distributions consist of [...] extreme skews”* (17), the Student *t* test is still frequently used (18, 19) because with groups containing more than 30 patients, the Student *t* test is adequate for comparing means.

Nevertheless, both the Student *t* test and WRS test are inappropriate. Considering our fictive example, we compared the two groups by both tests (Table 1). We also estimated cumulative incidence functions of extubation and used the Gray test (20), which compares time to extubation, taking into account that death is a competing event to extubation, so a patient who dies before extubation can no longer be extubated.

Several comments can be made. First, while all the events (death and extubation) occurred before day 60, the p-value associated with the Student *t* test decreased with increasing time horizon (0.08, 0.004 and 0.0002 at days 60, 90 and 180, respectively). This result was expected: by increasing the time horizon, the VFD

distribution was even more skewed, and the Student t test is sensitive to extreme values. Second, whatever the time horizon, the median VFD values were equal in the two groups. This situation illustrates a greater robustness of medians as compared to means and invites use of a non-parametric approach to compare the two groups. However, the WRS test result was non-significant, so this test failed to identify that the two groups are different. Finally, the Gray test result was significant ($p = 0.017$), so in accounting for mortality, time to extubation (and therefore MV duration) was significantly higher in the experimental than control group (Figure 2). Such an example, although fictive, is not unrealistic. Indeed, as observed by Willson et al. (21) and noted by Spragg et al. (22), with a significant difference in mortality rates (36% vs 19%, $p = 0.03$), the VFD means cannot differ (11.5 ± 10.5 days in the placebo group vs 13.2 ± 10.0 in the calfactant group, $p = 0.21$) These results advocate using time-dependent statistical approaches that account for a competitive risk situation.

Example based on real data: evidence-based extubation readiness bundle in brain-injury patients

To illustrate these statistical issues, we re-analysed the data proposed by Roquilly *et al.* (23) who assessed the effectiveness of an extubation readiness bundle to decrease ventilator time in 499 patients with brain injury. Although this study is a before-after study and not a randomized trial, this does not change anything on the statistical approach to be used when one is interested in comparing MV exposition between two groups. As previously explained, death is a competing event when considering extubation. We consider cumulative incidence functions (CIF) for extubation which correspond to the probability that extubation occurs before a specified time, taking into account that some patients may die before been

extubated. These CIF are estimated in both groups for each time point and compared using a Gray test. Results are displayed in Figure 3. At Day 28, 68% of patients are free of ventilation and still alive in the intervention group, as compared to 61% in the control group, and here is the right way to read such a Figure. We observe a significant difference between the two curves ($P = 0.036$). These CIF differ from classical survival curves obtained using the Kaplan-Meier (KM) approach. Thus, CIF asymptotes actually correspond to the proportion of extubated patients. On the contrary, when using KM estimates (and plotting one minus KM estimates) asymptotes equal 1. Such an approach has already been used in intensive care studies, but still marginally. Thus, Mekontso Dessap *et al* (24) used it to investigate whether fluid management guided by daily Brain Natriuretic Peptide plasma concentrations improves weaning outcomes compared with empirical therapy dictated by clinical acumen in a randomized trial of 304 patients. Alberti *et al* also used this approach to analyse ICU acquired-infections in a cohort of 8,353 patients (25). In this latter example, the event of interest was no longer extubation, but ICU acquired-infection, thus illustrating that a competing event approach should be used for other intensive care outcomes than extubation.

Conclusion

VFD is a challenging endpoint. The VFD definition is not consensual, and some authors called for an international expert consensus and standardized outcome definitions (16, 26–30), which agrees with the Core Outcome Measures in Effectiveness Trials (COMET) initiative (31). In this paper, we additionally demonstrated that 1) the VFD should not be interpreted in itself, 2) the choice of the time horizon for assessing the VFD outcome should be guided by clinical considerations, and 3) this outcome should not be analysed by classical statistical

methods such as the Student *t* test and WRS test but rather tests that consider the time-dependent and competing nature of death and extubation. Such findings question the widespread use of the VFD outcome, especially as a primary outcome (32).

More research is needed on the best way to deal with data related to invasive MV, taking into account death, which, because of its high incidence, cannot be ignored. Methods have been developed (9), but they remain still under-used (33) especially for intensive care data. In addition, physicians need to become more familiar with these methods and their interpretations.

References

1. The ARDS Definition Task Force: Acute Respiratory Distress Syndrome. The Berlin Definition. The Berlin Definition of ARDS. JAMA: The Journal of the American Medical Association. 2012 May 21;1.
2. Afshari A, Brok J, Møller AM, et al: Inhaled nitric oxide for acute respiratory distress syndrome (ARDS) and acute lung injury in children and adults. Cochrane Database Syst Rev. 2010;(7):CD002787.
3. Meade MO, Cook DJ, Guyatt GH, et al: Ventilation strategy using low tidal volumes, recruitment maneuvers, and high positive end-expiratory pressure for acute lung injury and acute respiratory distress syndrome: a randomized controlled trial. JAMA. 2008 Feb 13;299(6):637–45.
4. Marini JJ, Evans TW: Acute Lung Injury. illustrated edition. Springer-Verlag Berlin and Heidelberg GmbH & Co. K; 1997. 400 p.
5. Bellani G, Laffey JG, Pham T, et al: Epidemiology, Patterns of Care, and Mortality for Patients With Acute Respiratory Distress Syndrome in Intensive Care Units in 50 Countries. JAMA. 2016 Feb 23;315(8):788–800.

6. Bouadma L, Sonnevile R, Garrouste-Orgeas M, et al: Ventilator-Associated Events: Prevalence, Outcome, and Relationship With Ventilator-Associated Pneumonia. *Crit Care Med*. 2015 Sep;43(9):1798–806.
7. Bassetti M, Taramasso L, Giacobbe DR, et al: Management of ventilator-associated pneumonia: epidemiology, diagnosis and antimicrobial therapy. *Expert Rev Anti Infect Ther*. 2012 May;10(5):585–96.
8. Rubenfeld GD, Angus DC, Pinsky MR, et al: Outcomes research in critical care: results of the American Thoracic Society Critical Care Assembly Workshop on Outcomes Research. The Members of the Outcomes Research Workshop. *Am J Respir Crit Care Med*. 1999 Jul;160(1):358–67.
9. Checkley W, Brower RG, Muñoz A: Inference for mutually exclusive competing events through a mixture of generalized gamma distributions. *Epidemiology*. 2010 Jul;21(4):557–65.
10. Bernard GR, Wheeler AP, Arons MM, et al: A trial of antioxidants N-acetylcysteine and procysteine in ARDS. The Antioxidant in ARDS Study Group. *Chest*. 1997 Jul;112(1):164–72.
11. Schoenfeld DA, Bernard GR: 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.
12. Bein T, Weber-Carstens S, Goldmann A, et al: Lower tidal volume strategy (≈ 3 ml/kg) combined with extracorporeal CO₂ 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.
13. McAuley DF, Laffey JG, O’Kane CM, et al: Simvastatin in the acute respiratory distress syndrome. *N Engl J Med*. 2014 Oct 30;371(18):1695–703.

14. National Heart, Lung, and Blood Institute ARDS Clinical Trials Network, Truitt JD, Bernard GR, Steingrub J, et al: Rosuvastatin for sepsis-associated acute respiratory distress syndrome. *N Engl J Med*. 2014 Jun 5;370(23):2191–200.
15. Marini JJ, Evans TW: Round table conference: acute lung injury, 15th-17th March 1997 Brussels, Belgium. *Intensive Care Med*. 1998 Aug;24(8):878–83.
16. 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.
17. Bridge PD, Sawilowsky SS: Increasing physicians' awareness of the impact of statistics on research outcomes: comparative power of the t-test and Wilcoxon Rank-Sum test in small samples applied research. *J Clin Epidemiol*. 1999 Mar;52(3):229–35.
18. Trouillet J-L, Luyt C-E, Guiguet M, 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.
19. Richard C, Warszawski J, Anguel N, et al: Early use of the pulmonary artery catheter and outcomes in patients with shock and acute respiratory distress syndrome: a randomized controlled trial. *JAMA*. 2003 Nov 26;290(20):2713–20.
20. Gray RJ: A Class of Sample Tests for Comparing the Cumulative Incidence of a Competing Risk. *Ann Statist*. 1988 Sep;16(3):1141–54.
21. Willson DF, Thomas NJ, Markovitz BP, et al: Effect of exogenous surfactant (calfactant) in pediatric acute lung injury: a randomized controlled trial. *JAMA*. 2005 Jan 26;293(4):470–6.
22. Spragg RG, Bernard GR, Checkley W, et al: Beyond mortality: future clinical research in acute lung injury. *Am J Respir Crit Care Med*. 2010 May

15;181(10):1121–7.

23. Roquilly A, Cinotti R, Jaber S, et al: Implementation of an evidence-based extubation readiness bundle in 499 brain-injured patients. a before-after evaluation of a quality improvement project. *Am J Respir Crit Care Med*. 2013 Oct 15;188(8):958–66.
24. Mekontso Dessap A, Roche-Campo F, Kouatchet A, et al: Natriuretic peptide-driven fluid management during ventilator weaning: a randomized controlled trial. *Am J Respir Crit Care Med*. 2012 Dec 15;186(12):1256–63.
25. Alberti C, Métivier F, Landais P, et al: Improving estimates of event incidence over time in populations exposed to other events: application to three large databases. *J Clin Epidemiol*. 2003 Jun;56(6):536–45.
26. Blackwood B, Clarke M, McAuley DF, et al: How outcomes are defined in clinical trials of mechanically ventilated adults and children. *Am J Respir Crit Care Med*. 2014 Apr 15;189(8):886–93.
27. Blackwood B, Marshall J, Rose L: Progress on core outcome sets for critical care research. *Curr Opin Crit Care*. 2015 Oct;21(5):439–44.
28. Blackwood B, Ringrow S, Clarke M, et al: Core Outcomes in Ventilation Trials (COVenT): protocol for a core outcome set using a Delphi survey with a nested randomised trial and observational cohort study. *Trials*. 2015;16:368.
29. Needham DM: Understanding and improving clinical trial outcome measures in acute respiratory failure. *Am J Respir Crit Care Med*. 2014 Apr 15;189(8):875–7.
30. Harhay MO, Wagner J, Ratcliffe SJ, 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.
31. Standardizing reporting of core outcome measures in ventilation studies : Core

Outcome Measures in Effectiveness Trials Initiative (COMET) [Internet]. [cited 2016 Feb 15]. Available from: <http://www.comet-initiative.org/studies/details/292>

32. Simonis FD, Binnekade JM, Braber A, et al: PReVENT--protective ventilation in patients without ARDS at start of ventilation: study protocol for a randomized controlled trial. *Trials*. 2015;16:226.

33. Wolkewitz M, Cooper BS, Bonten MJM, et al: Interpreting and comparing risks in the presence of competing events. *BMJ*. 2014;349:g5060.

Table 1: Mean and median ventilator-free days (VFD) at days 28, 60, 90, and 180 per group

Figure 1: Iso-ventilator-free day curve

Figure 2: Cumulative incidence curve for extubation per group, hypothetical example

Figure 3: Cumulative incidence curve for extubation per group, real data set

Table 1: Mean and median ventilator-free days (VFD) at days 28, 60, 90, and 180 per group

VFD	Control Group (n=500)	Intervention Group (n=500)	p-value (Student t-test)
VFD day 28	10.8 (10.8)	10.4 (10.2)	0.47
VFD day 60	30.9 (24.5)	33.5 (21.6)	0.08
VFD day 90	50.7 (37.9)	57.0 (32.5)	0.004
VFD day 180	109.9 (79.9)	127.6 (68.2)	0.0002

Results are presented mean (standard deviation).

VFD	Control Group (n=500)	Intervention Group (n=500)	p-value (Wilcoxon Signed-Rank test)
VFD day 28	10.0 (0.0 – 22.0)	10.0 (0.0 – 20.0)	0.50
VFD day 60	42.0 (0.0 – 54.0)	42.0 (11.5 – 52.0)	0.52
VFD day 90	72.0 (0.0 – 84.0)	72.0 (41.5 – 82.0)	0.52
VFD day 180	162.0 (0.0 – 174.0)	162.0 (131.5 – 172.0)	0.52

Results are presented median (1st-3rd quartile).

Cumulative incidence for Extubation

	Control Group (n=500)	Intervention Group (n=500)
At 28-days of follow up	0.57	0.62

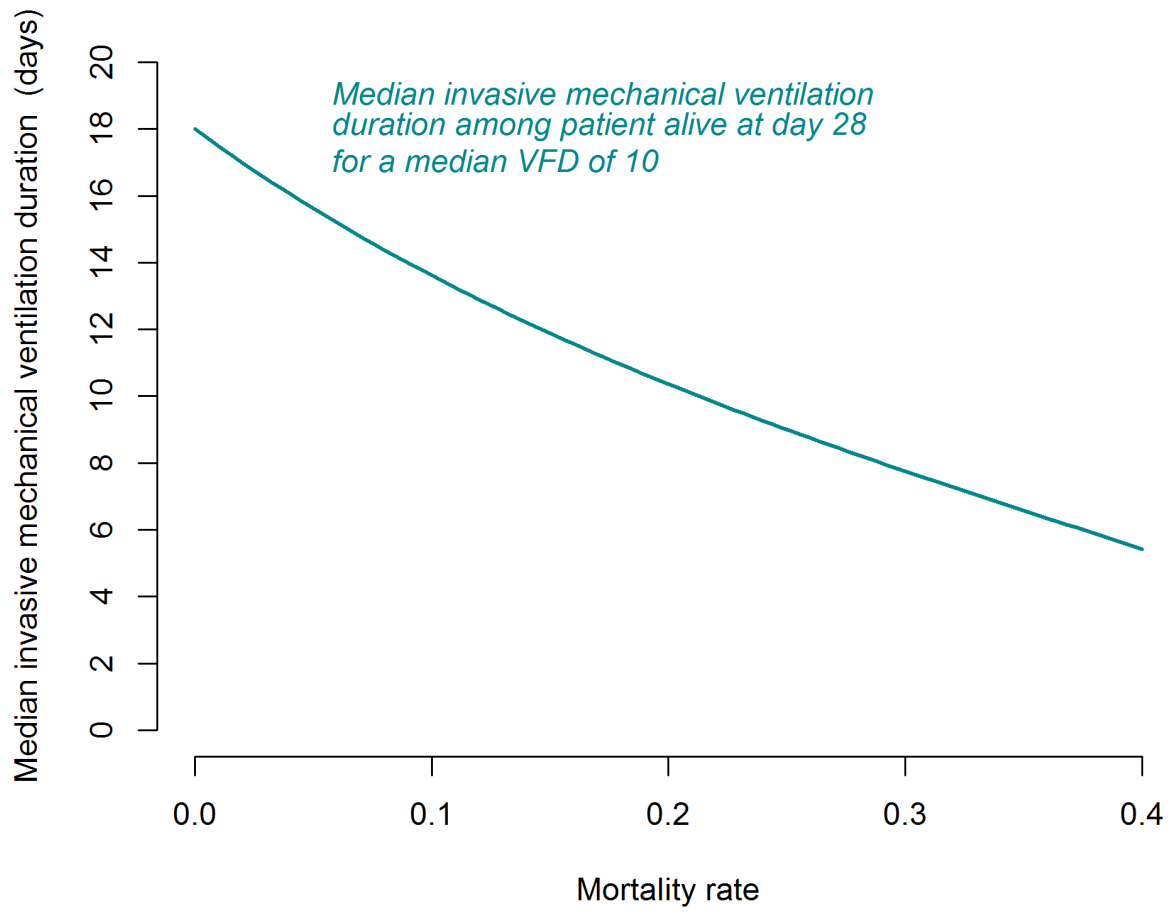
Figure 1: Iso-ventilator-free day curve

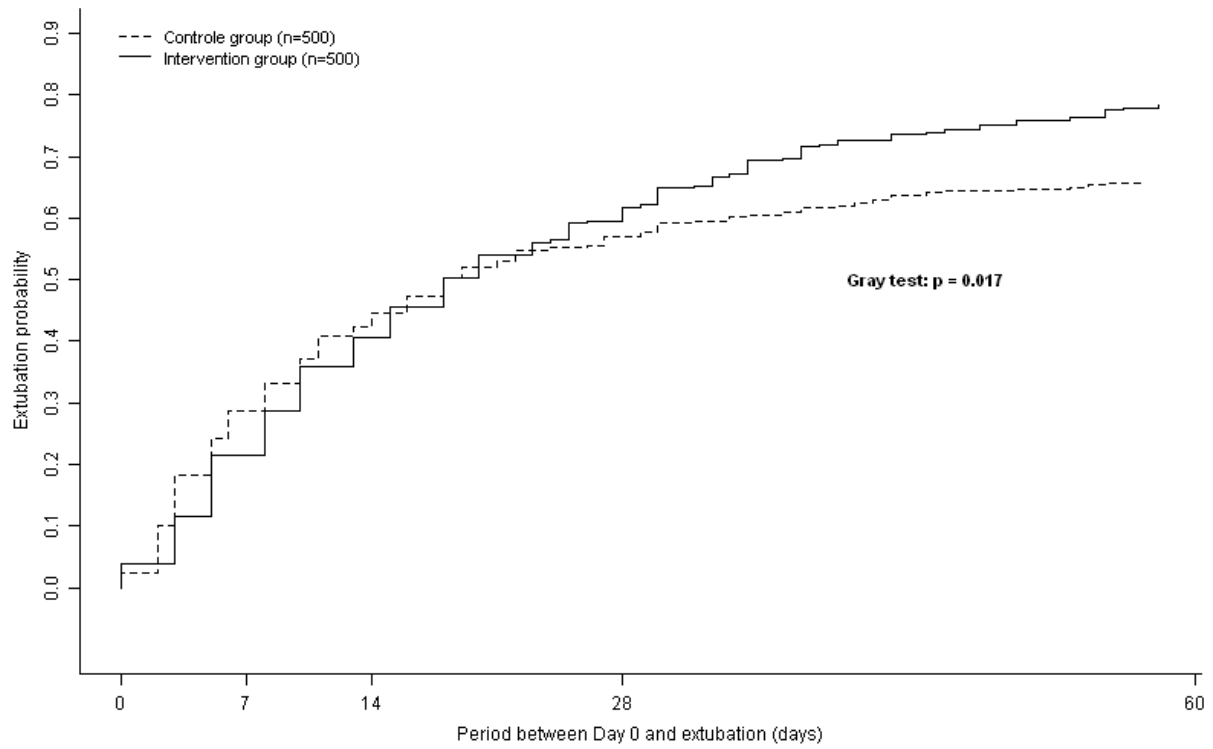
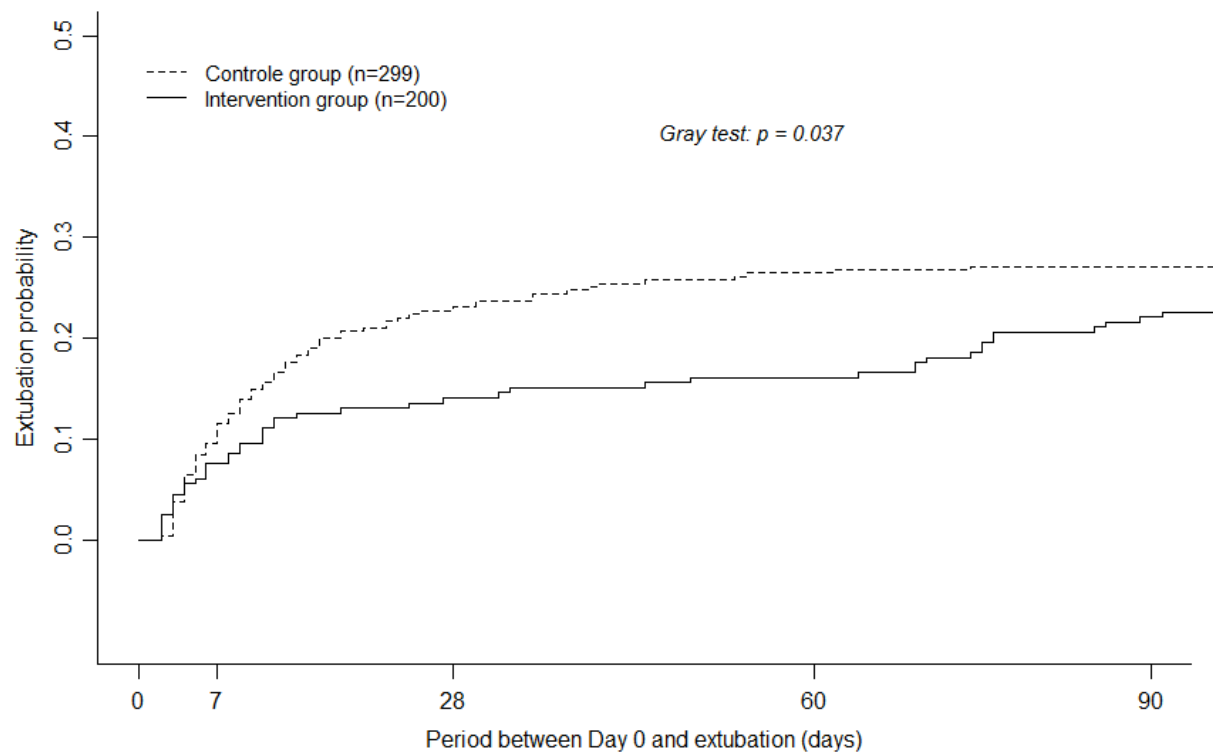
Figure 2: Cumulative incidence curve for extubation per group, hypothetical example

Figure 3: Cumulative incidence curve for extubation per group, real data set

Online data supplement

Ventilator-free day outcomes can be misleading

Bodet-Contentin L, Frasca D, Tavernier E, Feuillet F, Roquilly A

Foucher Y, Giraudeau B

Online data supplement 1: data for the hypothetical example

Let us consider a two-arm randomized trial:

- 500 patients per group
- mortality rate of 34.2% in the control group
- mortality rate of 21.6% in the intervention group
- median invasive mechanical ventilation (MV) duration of 7 days for survivors in the control group
- median invasive MV duration of 10 days for survivors in the experimental group

In this example, all events (death and extubation) are observed before day 60.

The full data set is displayed below, using the following coding:

- Group:
 - 0 = Control group
 - 1 = Intervention group
- Event:
 - 1 = Extubation
 - 2 = Death

Patient	Group	Event	Delay before the event (days)
1	0	2	0
2	0	2	19
3	0	1	14
4	0	2	19
5	0	1	10
6	0	2	6
7	0	2	2
8	0	1	3
9	0	2	22
10	0	1	6
11	0	1	5
12	0	2	0
13	0	2	10
14	0	2	38
15	0	1	8
16	0	2	0
17	0	1	10
18	0	1	0
19	0	1	3
20	0	1	3
21	0	1	0
22	0	1	3
23	0	1	14
24	0	1	6
25	0	1	11
26	0	1	2
27	0	1	6
28	0	1	5
29	0	1	0
30	0	1	27
31	0	2	16
32	0	1	38
33	0	1	8
34	0	2	6
35	0	1	18
36	0	1	2
37	0	1	2
38	0	1	3
39	0	1	13
40	0	2	3
41	0	2	45
42	0	2	3
43	0	2	2
44	0	1	6
45	0	2	42
46	0	2	29
47	0	1	5
48	0	1	2
49	0	1	3

Patient	Group	Event	Delay before the event (days)
50	0	1	5
51	0	2	32
52	0	2	0
53	0	1	24
54	0	2	0
55	0	2	34
56	0	2	3
57	0	2	3
58	0	2	5
59	0	2	6
60	0	2	38
61	0	2	11
62	0	2	21
63	0	2	0
64	0	2	38
65	0	2	21
66	0	2	30
67	0	2	22
68	0	2	34
69	0	2	49
70	0	2	3
71	0	2	6
72	0	2	14
73	0	2	16
74	0	1	10
75	0	1	6
76	0	1	11
77	0	1	3
78	0	1	10
79	0	1	0
80	0	1	2
81	0	2	10
82	0	1	27
83	0	1	50
84	0	1	41
85	0	1	6
86	0	2	3
87	0	1	42
88	0	1	8
89	0	1	21
90	0	2	10
91	0	2	8
92	0	1	2
93	0	1	0
94	0	1	54
95	0	1	18
96	0	1	10
97	0	1	5
98	0	1	18

Patient	Group	Event	Delay before the event (days)
99	0	2	2
100	0	1	18
101	0	1	8
102	0	1	18
103	0	2	16
104	0	1	2
105	0	1	5
106	0	1	19
107	0	1	30
108	0	1	3
109	0	2	2
110	0	1	32
111	0	2	8
112	0	1	2
113	0	1	2
114	0	2	6
115	0	1	8
116	0	1	16
117	0	2	2
118	0	1	30
119	0	2	2
120	0	1	3
121	0	1	24
122	0	1	6
123	0	1	3
124	0	1	5
125	0	1	14
126	0	2	11
127	0	2	2
128	0	1	6
129	0	1	2
130	0	1	6
131	0	1	27
132	0	2	3
133	0	1	18
134	0	1	19
135	0	1	8
136	0	1	3
137	0	1	3
138	0	1	8
139	0	2	2
140	0	2	10
141	0	2	2
142	0	1	3
143	0	1	2
144	0	2	2
145	0	1	5
146	0	1	14
147	0	2	0

Patient	Group	Event	Delay before the event (days)
148	0	1	22
149	0	1	27
150	0	1	16
151	0	1	5
152	0	1	29
153	0	1	16
154	0	1	18
155	0	1	5
156	0	1	2
157	0	1	3
158	0	1	16
159	0	1	10
160	0	2	10
161	0	1	3
162	0	1	6
163	0	1	2
164	0	1	18
165	0	2	10
166	0	1	8
167	0	1	8
168	0	1	3
169	0	1	22
170	0	1	34
171	0	1	6
172	0	1	5
173	0	1	10
174	0	1	11
175	0	1	2
176	0	1	8
177	0	1	22
178	0	1	8
179	0	1	16
180	0	2	2
181	0	1	3
182	0	2	45
183	0	2	3
184	0	1	11
185	0	1	3
186	0	1	40
187	0	1	3
188	0	1	2
189	0	1	0
190	0	2	2
191	0	1	2
192	0	1	13
193	0	1	2
194	0	1	8
195	0	1	22
196	0	1	3

Patient	Group	Event	Delay before the event (days)
197	0	2	11
198	0	2	51
199	0	1	5
200	0	1	3
201	0	1	3
202	0	2	38
203	0	1	5
204	0	1	3
205	0	1	11
206	0	1	6
207	0	1	3
208	0	2	37
209	0	1	18
210	0	1	3
211	0	1	6
212	0	1	2
213	0	2	14
214	0	2	6
215	0	1	32
216	0	1	16
217	0	1	8
218	0	1	2
219	0	1	6
220	0	1	11
221	0	1	5
222	0	1	10
223	0	1	14
224	0	2	0
225	0	1	5
226	0	1	13
227	0	1	29
228	0	1	3
229	0	1	3
230	0	2	53
231	0	1	5
232	0	1	30
233	0	1	18
234	0	2	8
235	0	1	53
236	0	2	24
237	0	2	3
238	0	1	30
239	0	1	14
240	0	1	11
241	0	2	22
242	0	1	5
243	0	1	19
244	0	1	2
245	0	1	37

Patient	Group	Event	Delay before the event (days)
246	0	1	6
247	0	2	5
248	0	2	2
249	0	2	13
250	0	2	2
251	0	1	8
252	0	2	13
253	0	1	13
254	0	1	3
255	0	1	2
256	0	1	18
257	0	2	5
258	0	1	10
259	0	1	11
260	0	1	13
261	0	1	55
262	0	1	5
263	0	1	10
264	0	1	3
265	0	1	2
266	0	2	2
267	0	1	2
268	0	1	54
269	0	1	16
270	0	2	5
271	0	1	14
272	0	1	5
273	0	1	10
274	0	1	2
275	0	1	10
276	0	1	13
277	0	1	27
278	0	2	16
279	0	1	10
280	0	2	5
281	0	1	18
282	0	1	8
283	0	1	16
284	0	1	11
285	0	1	26
286	0	1	11
287	0	2	19
288	0	1	21
289	0	1	0
290	0	1	5
291	0	1	2
292	0	1	6
293	0	1	0
294	0	1	3

Patient	Group	Event	Delay before the event (days)
295	0	1	22
296	0	1	10
297	0	2	10
298	0	1	13
299	0	2	6
300	0	2	2
301	0	2	16
302	0	2	3
303	0	1	13
304	0	1	45
305	0	1	11
306	0	1	11
307	0	2	6
308	0	1	40
309	0	1	2
310	0	1	0
311	0	1	22
312	0	1	43
313	0	2	8
314	0	1	2
315	0	1	2
316	0	1	19
317	0	2	5
318	0	1	30
319	0	1	11
320	0	1	34
321	0	1	5
322	0	1	2
323	0	1	3
324	0	1	11
325	0	1	29
326	0	1	2
327	0	1	14
328	0	1	8
329	0	2	6
330	0	1	35
331	0	2	50
332	0	1	2
333	0	1	8
334	0	2	19
335	0	1	2
336	0	1	3
337	0	1	5
338	0	1	5
339	0	1	6
340	0	1	3
341	0	1	18
342	0	1	5
343	0	1	26

Patient	Group	Event	Delay before the event (days)
344	0	2	5
345	0	1	8
346	0	1	2
347	0	1	6
348	0	2	14
349	0	2	6
350	0	1	10
351	0	1	5
352	0	1	11
353	0	1	5
354	0	1	0
355	0	1	2
356	0	1	3
357	0	1	2
358	0	2	45
359	0	1	0
360	0	1	10
361	0	1	14
362	0	1	30
363	0	1	5
364	0	1	2
365	0	2	3
366	0	2	19
367	0	2	8
368	0	2	35
369	0	2	2
370	0	2	13
371	0	2	42
372	0	2	0
373	0	2	3
374	0	2	37
375	0	2	2
376	0	2	5
377	0	2	22
378	0	2	26
379	0	2	14
380	0	2	24
381	0	2	3
382	0	2	27
383	0	1	21
384	0	1	46
385	0	1	30
386	0	1	16
387	0	2	2
388	0	1	50
389	0	1	22
390	0	1	16
391	0	1	6
392	0	1	43

Patient	Group	Event	Delay before the event (days)
393	0	1	16
394	0	1	37
395	0	2	6
396	0	1	41
397	0	2	0
398	0	1	18
399	0	2	19
400	0	1	8
401	0	1	6
402	0	1	43
403	0	1	27
404	0	1	3
405	0	1	22
406	0	2	0
407	0	2	54
408	0	1	21
409	0	1	19
410	0	1	38
411	0	1	29
412	0	1	11
413	0	1	57
414	0	1	2
415	0	2	26
416	0	2	40
417	0	1	27
418	0	1	16
419	0	1	10
420	0	1	16
421	0	1	14
422	0	1	34
423	0	1	19
424	0	1	5
425	0	2	2
426	0	1	0
427	0	2	26
428	0	1	8
429	0	1	3
430	0	2	42
431	0	1	11
432	0	2	0
433	0	1	5
434	0	1	42
435	0	1	10
436	0	1	45
437	0	1	22
438	0	1	8
439	0	1	6
440	0	2	6
441	0	1	37

Patient	Group	Event	Delay before the event (days)
442	0	1	42
443	0	1	3
444	0	1	43
445	0	2	35
446	0	2	13
447	0	2	19
448	0	2	5
449	0	2	6
450	0	2	8
451	0	2	2
452	0	2	57
453	0	2	26
454	0	2	8
455	0	2	0
456	0	2	14
457	0	2	32
458	0	2	6
459	0	2	2
460	0	1	2
461	0	1	8
462	0	1	19
463	0	1	3
464	0	1	8
465	0	1	38
466	0	1	3
467	0	1	10
468	0	1	19
469	0	1	5
470	0	1	19
471	0	2	2
472	0	2	47
473	0	1	6
474	0	2	3
475	0	1	2
476	0	2	10
477	0	2	38
478	0	2	0
479	0	2	10
480	0	1	3
481	0	2	8
482	0	2	6
483	0	2	0
484	0	2	5
485	0	2	10
486	0	2	6
487	0	2	5
488	0	2	19
489	0	2	6
490	0	2	37

Patient	Group	Event	Delay before the event (days)
491	0	2	5
492	0	2	3
493	0	2	16
494	0	2	40
495	0	1	21
496	0	1	14
497	0	1	18
498	0	1	11
499	0	2	2
500	0	1	10
501	1	1	38
502	1	1	10
503	1	1	28
504	1	1	0
505	1	2	53
506	1	1	8
507	1	1	38
508	1	1	5
509	1	2	5
510	1	1	3
511	1	1	23
512	1	1	10
513	1	1	40
514	1	1	26
515	1	1	13
516	1	1	5
517	1	2	3
518	1	1	30
519	1	2	8
520	1	1	25
521	1	1	8
522	1	2	3
523	1	1	8
524	1	2	56
525	1	1	10
526	1	1	3
527	1	1	55
528	1	2	18
529	1	1	28
530	1	1	8
531	1	2	37
532	1	1	0
533	1	1	5
534	1	2	8
535	1	1	3
536	1	1	18
537	1	1	5
538	1	2	5
539	1	1	20

Patient	Group	Event	Delay before the event (days)
540	1	1	33
541	1	1	3
542	1	1	0
543	1	1	5
544	1	1	3
545	1	1	3
546	1	1	30
547	1	1	5
548	1	1	43
549	1	1	15
550	1	1	3
551	1	2	3
552	1	1	58
553	1	1	10
554	1	1	25
555	1	1	18
556	1	1	35
557	1	1	46
558	1	1	13
559	1	1	35
560	1	1	5
561	1	2	15
562	1	2	18
563	1	1	25
564	1	1	20
565	1	1	25
566	1	1	28
567	1	1	5
568	1	2	18
569	1	1	8
570	1	1	0
571	1	1	15
572	1	1	13
573	1	1	13
574	1	1	8
575	1	1	13
576	1	1	23
577	1	1	48
578	1	1	53
579	1	2	45
580	1	1	25
581	1	1	3
582	1	1	25
583	1	1	24
584	1	2	3
585	1	1	20
586	1	2	28
587	1	1	30
588	1	1	28

Patient	Group	Event	Delay before the event (days)
589	1	1	25
590	1	1	50
591	1	1	38
592	1	2	10
593	1	1	23
594	1	2	3
595	1	1	20
596	1	1	23
597	1	1	10
598	1	1	38
599	1	2	3
600	1	1	33
601	1	2	3
602	1	1	20
603	1	1	30
604	1	2	3
605	1	1	15
606	1	1	3
607	1	1	15
608	1	2	53
609	1	1	5
610	1	2	26
611	1	1	5
612	1	1	10
613	1	2	31
614	1	2	27
615	1	1	5
616	1	2	18
617	1	2	3
618	1	1	8
619	1	1	15
620	1	2	0
621	1	1	25
622	1	1	34
623	1	2	3
624	1	1	10
625	1	1	32
626	1	1	10
627	1	2	3
628	1	1	5
629	1	1	15
630	1	1	23
631	1	1	3
632	1	2	31
633	1	1	15
634	1	1	5
635	1	2	3
636	1	1	55
637	1	1	0

Patient	Group	Event	Delay before the event (days)
638	1	1	10
639	1	1	55
640	1	1	13
641	1	1	15
642	1	1	10
643	1	1	10
644	1	1	0
645	1	1	8
646	1	1	3
647	1	1	30
648	1	1	0
649	1	1	40
650	1	1	10
651	1	2	3
652	1	1	33
653	1	1	53
654	1	1	20
655	1	1	5
656	1	1	10
657	1	1	5
658	1	1	10
659	1	1	3
660	1	1	3
661	1	1	20
662	1	2	0
663	1	1	37
664	1	1	8
665	1	2	5
666	1	1	56
667	1	2	28
668	1	1	35
669	1	2	35
670	1	1	28
671	1	1	15
672	1	1	8
673	1	1	0
674	1	1	3
675	1	1	35
676	1	1	13
677	1	2	3
678	1	2	8
679	1	1	8
680	1	1	3
681	1	1	58
682	1	1	20
683	1	1	50
684	1	1	0
685	1	2	3
686	1	1	43

Patient	Group	Event	Delay before the event (days)
687	1	1	3
688	1	1	35
689	1	1	10
690	1	1	0
691	1	1	10
692	1	2	3
693	1	1	3
694	1	1	18
695	1	1	10
696	1	1	8
697	1	1	25
698	1	1	33
699	1	1	18
700	1	1	38
701	1	1	5
702	1	2	3
703	1	1	20
704	1	1	10
705	1	1	15
706	1	1	28
707	1	1	13
708	1	1	10
709	1	1	13
710	1	1	10
711	1	1	20
712	1	2	43
713	1	1	35
714	1	1	5
715	1	1	15
716	1	1	3
717	1	1	30
718	1	1	5
719	1	1	8
720	1	1	0
721	1	2	53
722	1	2	25
723	1	1	10
724	1	1	13
725	1	1	28
726	1	2	10
727	1	1	18
728	1	1	8
729	1	1	29
730	1	1	23
731	1	2	25
732	1	2	23
733	1	2	0
734	1	1	8
735	1	1	33

Patient	Group	Event	Delay before the event (days)
736	1	1	18
737	1	1	20
738	1	2	54
739	1	1	20
740	1	1	13
741	1	2	3
742	1	2	38
743	1	2	13
744	1	1	15
745	1	1	38
746	1	1	3
747	1	1	15
748	1	1	3
749	1	1	3
750	1	1	3
751	1	1	5
752	1	2	25
753	1	1	5
754	1	1	43
755	1	1	8
756	1	2	35
757	1	1	13
758	1	1	5
759	1	2	3
760	1	1	25
761	1	1	13
762	1	2	0
763	1	1	26
764	1	2	53
765	1	1	0
766	1	1	18
767	1	1	38
768	1	1	43
769	1	1	5
770	1	1	18
771	1	1	3
772	1	2	5
773	1	1	23
774	1	1	3
775	1	1	8
776	1	1	20
777	1	1	3
778	1	1	33
779	1	1	8
780	1	1	10
781	1	1	3
782	1	1	3
783	1	2	25
784	1	1	35

Patient	Group	Event	Delay before the event (days)
785	1	1	0
786	1	1	58
787	1	1	3
788	1	1	13
789	1	1	18
790	1	1	13
791	1	1	5
792	1	1	18
793	1	2	38
794	1	1	8
795	1	1	13
796	1	1	0
797	1	1	5
798	1	1	15
799	1	1	5
800	1	1	39
801	1	1	30
802	1	1	3
803	1	1	3
804	1	1	8
805	1	1	10
806	1	1	50
807	1	1	10
808	1	2	28
809	1	1	3
810	1	1	18
811	1	1	38
812	1	2	3
813	1	1	30
814	1	2	13
815	1	2	43
816	1	1	23
817	1	1	38
818	1	1	30
819	1	2	8
820	1	1	20
821	1	1	55
822	1	1	8
823	1	1	5
824	1	1	46
825	1	1	35
826	1	2	24
827	1	1	20
828	1	1	13
829	1	1	0
830	1	1	18
831	1	1	0
832	1	1	5
833	1	1	5

Patient	Group	Event	Delay before the event (days)
834	1	1	5
835	1	1	8
836	1	1	50
837	1	1	5
838	1	2	28
839	1	1	10
840	1	1	48
841	1	2	3
842	1	1	8
843	1	1	13
844	1	1	10
845	1	2	15
846	1	1	8
847	1	1	3
848	1	1	8
849	1	1	0
850	1	1	5
851	1	2	45
852	1	2	28
853	1	1	13
854	1	1	15
855	1	2	13
856	1	1	5
857	1	1	5
858	1	1	8
859	1	1	15
860	1	1	30
861	1	1	0
862	1	1	18
863	1	1	3
864	1	2	15
865	1	1	5
866	1	1	5
867	1	1	18
868	1	1	15
869	1	1	5
870	1	1	5
871	1	1	8
872	1	1	55
873	1	1	43
874	1	2	28
875	1	1	10
876	1	1	0
877	1	1	10
878	1	1	18
879	1	1	3
880	1	1	10
881	1	1	5
882	1	1	25

Patient	Group	Event	Delay before the event (days)
883	1	2	43
884	1	2	53
885	1	1	20
886	1	1	28
887	1	1	30
888	1	1	8
889	1	2	43
890	1	1	23
891	1	1	10
892	1	2	10
893	1	1	13
894	1	1	5
895	1	1	5
896	1	1	8
897	1	1	5
898	1	1	5
899	1	1	18
900	1	1	13
901	1	1	25
902	1	1	35
903	1	1	25
904	1	1	8
905	1	2	23
906	1	2	13
907	1	1	30
908	1	1	18
909	1	1	5
910	1	2	13
911	1	2	10
912	1	1	10
913	1	1	18
914	1	1	40
915	1	1	3
916	1	2	3
917	1	1	5
918	1	1	8
919	1	1	8
920	1	1	28
921	1	1	30
922	1	1	15
923	1	1	15
924	1	1	3
925	1	1	33
926	1	1	10
927	1	2	57
928	1	1	5
929	1	1	15
930	1	1	33
931	1	1	5

Patient	Group	Event	Delay before the event (days)
932	1	2	5
933	1	1	15
934	1	2	5
935	1	1	38
936	1	1	8
937	1	1	5
938	1	2	30
939	1	1	10
940	1	1	8
941	1	1	18
942	1	1	13
943	1	1	20
944	1	1	15
945	1	1	3
946	1	1	5
947	1	1	10
948	1	2	20
949	1	1	23
950	1	2	58
951	1	1	28
952	1	2	48
953	1	1	28
954	1	1	55
955	1	1	30
956	1	1	18
957	1	1	18
958	1	2	13
959	1	1	35
960	1	1	8
961	1	1	18
962	1	1	3
963	1	1	34
964	1	1	10
965	1	1	46
966	1	2	8
967	1	1	18
968	1	2	8
969	1	1	35
970	1	1	20
971	1	1	8
972	1	1	15
973	1	1	40
974	1	1	15
975	1	1	13
976	1	1	10
977	1	1	25
978	1	2	35
979	1	1	53
980	1	2	23

Patient	Group	Event	Delay before the event (days)
981	1	2	8
982	1	2	10
983	1	2	48
984	1	2	18
985	1	2	5
986	1	2	0
987	1	2	8
988	1	2	18
989	1	2	3
990	1	2	38
991	1	2	3
992	1	1	45
993	1	1	48
994	1	1	29
995	1	1	15
996	1	1	5
997	1	1	3
998	1	1	5
999	1	1	18
1000	1	1	24

Online data supplement 2: Iso-ventilator-free day curve execution (Fig. 1)

Let us consider ventilator-free day (VFD) at day 28 defined as follows:

- 0 for a patient who died before day 28
- $(28 - x)$, with x the number of days under invasive mechanical ventilation (MV), for a patient alive at day 28

Of note, if a patient is alive at day 28, but still mechanically ventilated, the VFD is set to 0.

With such a definition, VFD then combines two pieces of information, which, from a mathematical point of view, means that a VFD value is the solution of a single equation with two unknowns (the vital status and the duration of invasive MV).

Let us consider π , the mortality rate at day 28, and X , a random variable associated with duration of invasive MV for patients still alive at day 28. We suppose that the X random variable is exponentially distributed, with parameter λ , so the mean and median duration of MV (among people alive at day 28) are $1/\lambda$, and $\log(2)/\lambda$ respectively.

If we observe in a group of patients a median VFD of v , half of the individual VFD values are less than v . If we have a proportion π of patients who died before day 28 (and therefore with an individual VFD of 0), a proportion $(0.5-\pi)$ of patients are still alive at day 28 and have individual VFD values less than v . For those latter patients, their individual duration of MV is greater than $(28-v)$. Because we hypothesized an exponential distribution for duration of invasive MV, we have the following results:

$$(0.5-\pi) = \text{Prob}(X > 28-v) \text{ with } X \sim E(\lambda)$$

$$(0.5-\pi) = 1 - \text{Probability}(X < 28-v)$$

$$(0.5-\pi) = 1 - [1 - \exp(-\lambda(28-v))]$$

Which can be solved, leading to:

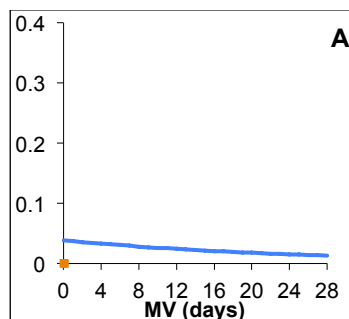
$$\lambda = -\log(0.5-\pi)/(28-\nu)$$

Therefore, as an example, a median VFD of 10 days corresponds to the following combinations of mortality rates and median invasive MV duration among survivors at day 28:

	π	λ	Median invasive MV duration among survivors at day 28 $\log(2)/\lambda$ (days)
A	0	0.0385	18.0
B	0.1	0.0509	13.6
C	0.2	0.0668	10.3
D	0.3	0.0890	7.8
E	0.4	0.1279	5.4

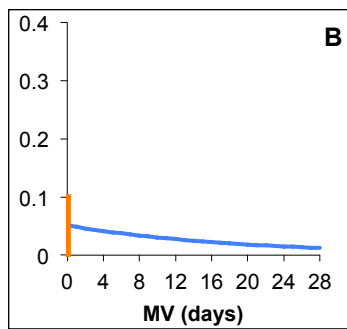
The combination A corresponds to a mortality rate of 0% and a median duration of invasive MV of 18 days. In the following Figure, we report both the mortality rate (orange line), which is 0 for this first combination, denoted A, and the invasive MV probability density function (blue line) for patients alive at day 28 (i.e., all patients, in this first situation, because the mortality rate is 0%).

Combination A: mortality rate 0%

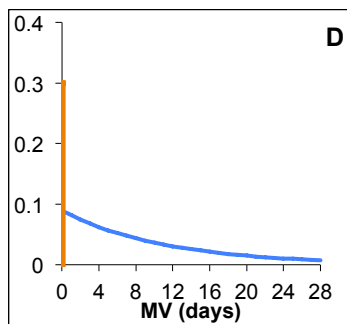


Such a figure has also been drawn for the other combinations B to E displayed in the Table, for the following four figures:

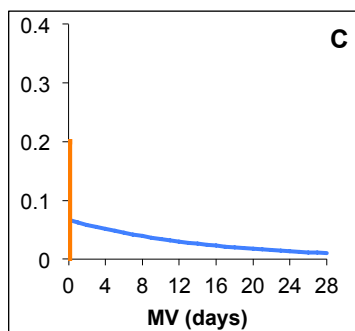
Combination B: mortality rate 10%



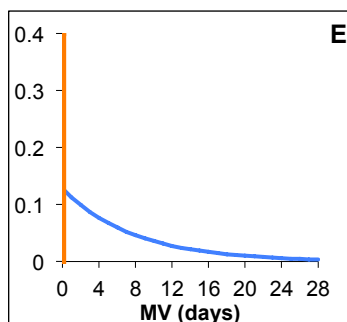
Combination D: mortality rate 30%



Combination C: mortality rate 20%



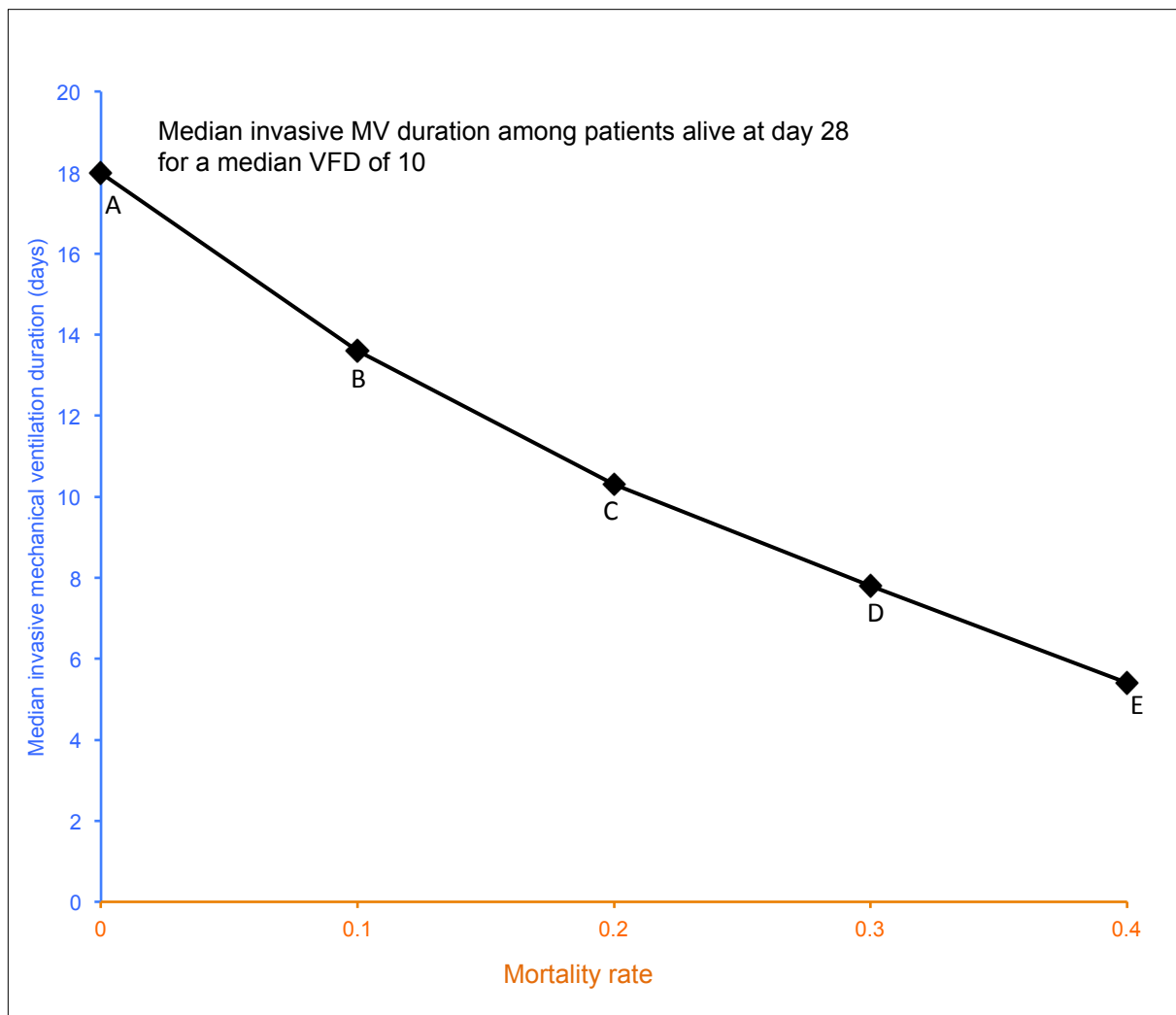
Combination E: mortality rate 40%



All these five combinations (A to E) lead to a common median VFD of 10 days, which allows for defining the concept of iso-VFD curve. The iso-VFD curve (black line in the following Figure) represents the different combinations of mortality rate (on the X-axis) and median duration of invasive MV among patients alive at day 28 (on the Y-axis). In this iso-VFD curve, we represent each possible combination (A to E).

Iso-VFD curve

(Dots A to E refer to the combinations cited in the Table)



4.3 Discussion et Perspectives

L'interprétation d'une valeur de ventilator-free days est difficile aussi bien en terme d'interprétation d'une valeur individuelle que lorsqu'on raisonne à l'échelle d'un groupe. Ceci s'explique par le fait que cette notion contient deux informations : la mortalité et le nombre de jours ventilés. Vouloir synthétiser cette double information en un seul chiffre conduit donc potentiellement à des erreurs d'interprétation aussi bien à l'échelon de l'essai lui-même qu'à l'échelon inter-essais (comparaison des essais entre eux, agrégation des résultats des essais dans le cadre d'une méta-analyse).

Par ailleurs, l'analyse statistique classique par un test Student ou un test de Wilcoxon n'est pas satisfaisante. Des alternatives existent pour comparer l'exposition à la ventilation invasive entre deux groupes de patients de réanimation qui ont une forte probabilité de décéder. Il faut pour cela prendre en compte le risque compétitif de décéder chez les patients ventilés. Les méthodes de risque compétitif ont déjà été utilisées quelques fois en réanimation aussi bien pour la ventilation invasive (61) que pour la survenue d'infections en réanimation (62). Nous les présentons dans les deux paragraphes qui suivent.

En 2003, Alberti et al. (62) ont repris trois cohortes de patients dont l'une de réanimation ayant 8353 patients inclus. Ils ont comparé les méthodes : Kaplan Meier et incidence cumulée (Fig. 20).

Figure 20 : Représentation des différentes analyses statistiques en présence de risque compétitif, extrait de Alberti et al. (62)

Ils ont montré qu'il existe un risque de sur-estimation de 3 à 30% (en fonction du taux de mortalité : plus la mortalité est importante plus le risque de sur-estimation de l'incidence de l'événement est importante) lors de l'utilisation d'une méthode de Kaplan Meier (Fig. 20). Ils présentent les courbes d'incidence cumulée en fonction d'un score de gravité.

Figure 21 : Présentation du risque de surestimation de l'événement en présence de risque compétitif lors d'une analyse par Kaplan Meier par rapport à une analyse en incidence cumulée, extrait de Alberti et al. (62)

En 2012, Mekontso Dessap et al. (61) ont utilisé une analyse en risques compétitifs pour l'analyse d'un essai de 306 patients consistant à comparer l'intérêt du dosage d'un biomarqueur dans la période de sevrage de la ventilation invasive. L'événement d'intérêt était le succès de l'extubation. Le risque compétitif pris en compte n'était pas la mortalité mais la nécessité d'une sédation continue. Ils ont construit les courbes d'incidence cumulée pour chacun des deux groupes et les ont comparées avec un modèle de Fine and Gray (Fig. 21) (61).

Figure 22 : Exemple d'utilisation d'une méthode d'analyse statistique en risque compétitif, extrait de Mekontso Dessap et al. (61)

DISCUSSION ET PERSPECTIVES

5 Discussion

5.1 Principaux résultats

Ce travail de thèse avait pour objectif de sensibiliser la communauté médicale de médecine intensive aux particularités des critères de jugement de type « durée de ventilation invasive » et « nombre de jours sans ventilation invasive (ventilator-free days, VFDs) » dans les essais randomisés de réanimation.

Dans un premier temps, nous avons identifié que :

- 1/ Dans les articles rapportant les résultats d'essais dans lesquels ces critères étaient évalués, peu donnaient une définition précise de comment ils étaient mesurés.
- 2/ Lorsqu'une définition était donnée, elle était, dans la majorité des cas, partielle.
- 3/ Les définitions retrouvées étaient hétérogènes.
- 4/ Lorsqu'on appliquait les différentes définitions à un patient fictif, cela conduisait à une variabilité des résultats.

Dans un deuxième temps, nous avons montré que :

- 1/ Pour une valeur de ventilator-free days, plusieurs interprétations en terme de mortalité et durée de ventilation invasive étaient possibles (cf. courbe iso-VFDs).
- 2/ L'utilisation du test de Student (analyse paramétrique) et du Wilcoxon (analyse non paramétrique) n'était pas adéquate.
- 3/ L'utilisation d'une approche en risques compétitifs était souhaitable (e.g. courbe d'incidence cumulée et test de Gray).

Au total, ces résultats montrent que la définition des critères de jugement concernant la ventilation invasive (durée de ventilation mécanique et ventilator-free days) n'est pas

clairement établie, impliquant une hétérogénéité et donc des difficultés d'analyse des essais randomisés. De plus, l'analyse statistique classiquement réalisée pour ces critères de jugement ne prend que rarement en compte le risque compétitif qu'est la mortalité, quand on s'intéresse à l'extubation.

5.2 Aspect de définition des critères de jugement

Dès 1997, le groupe de travail sur les critères de jugement en réanimation (45) a interpellé les cliniciens sur les critères de jugement à utiliser en réanimation avec les limites de chacun. Cependant à ce jour, les critères de jugement utilisés en réanimation sont nombreux, mal définis, et hétérogènes.

Concernant les critères de jugement en rapport avec la ventilation invasive, la difficulté est liée à:

- la définition de la fin de la ventilation qui est actuellement une nébuleuse (extubation, ventilation non invasive ou oxygénothérapie à haut débit selon certains paramètres réglés),
- la prise en compte ou non prise en compte des différentes périodes de ventilation,
- la période sur laquelle on observe l'événement d'intérêt qu'est l'extubation, et qui est variable (bien souvent 28, 60 ou 90 jours),
- ce que l'on considère comme de la ventilation « invasive » : initialement il s'agissait de la ventilation mécanique invasive via une sonde endotrachéale ou une trachéotomie, alors que pour certains, par exemple, une ventilation non invasive en continue sur 24 heures avec certains réglages du respirateur peut être considéré comme de la ventilation invasive,

- la date du décès est impactée par les décisions d'arrêt et de limitation des thérapeutiques actives, notamment par celles concernant la poursuite ou non de la ventilation invasive.

En 2014, Harhay et al. (50), dans la conclusion de leur revue systématique, ont appelé à définir les critères de jugement en spécifiant clairement le moment de leur évaluation. Ils ont en effet montré que les critères de jugement sont mal définis ou analysés.

Un travail sur la standardisation des critères de jugement dans les essais concernant la ventilation (COVent) est en cours dans le cadre de la COMET initiative. Il est mené par le Dr Bronagh Blackwood (44). La méthode retenue pour définir un « core outcome set » est une procédure de Delphi suivie d'un consensus, avec dans le même temps le suivi des études publiées sur une période de 5 à 10 ans pour évaluer l'utilisation de ce « core outcome set » et l'impact de la publication de ce set. Cependant, à ce jour, le « core outcome set » n'est pas encore publié.

5.3 Aspect de gestion de l'analyse des données

La méthode d'analyse statistique utilisée pour comparer deux groupes sur un critère de jugement en rapport avec la ventilation invasive est bien souvent non précisée (51). Lorsqu'elle l'est, on observe que la plupart du temps on recourt à une approche « classique » : une analyse paramétrique de type test de Student ou non paramétrique de type test de Wilcoxon. Cependant, par définition, la mortalité est un événement en compétition avec l'extubation, élément qui permet de définir les critères de jugement en rapport avec la ventilation invasive. Or, s'agissant de réanimation, la mortalité est importante, et la négliger devient donc inacceptable. L'analyse du critère de jugement concernant l'exposition à la ventilation invasive doit donc prendre en compte cet événement en compétition qu'est la mortalité. Pour cela, différents types d'analyses statistiques sont envisageables tels une approche en risques compétitifs ou un modèle de mélange (63).

Le concept de risque compétitif a été introduit par Bernouilli au XVIIIème siècle. Il s'agit d'une situation où un individu est exposé au risque de survenue de plusieurs événements, la survenue de l'un modifiant la probabilité de survenue de l'autre. Ajoutons enfin que, comme nous l'avons vu dans l'exemple du parcours ventilatoire d'un patient de réanimation (Fig. 12), un patient peut passer d'un état intubé à un état extubé à plusieurs reprises, situation que nous n'avons pas envisagée dans ce travail, dans un souci de simplicité.

6 Perspectives

6.1 Les autres critères de jugement en réanimation

6.1.1 Les autres « free days »

Dans le premier essai qui utilise le ventilator-free days, le concept de « failure free days » est également utilisé pour d'autres défaillances d'organes que le poumon (Fig. 22).

Figure 23 : Les « failure free days », extrait de Bernard et al. (55)

Depuis lors, ce concept est largement utilisé en réanimation pour les défaillances d'organes. Cependant la même problématique de définition et d'analyse statistique de ce concept se pose : quelle que soit la défaillance d'organe suppléée ou la prise en charge thérapeutique réalisée, elle est toujours en compétition avec le risque de décéder. A chaque fois, les analyses prenant en compte le risque compétitif sont à privilégier, comme l'indiquent Timsit et al. (64).

6.1.2 L'hétérogénéité des critères de jugement

Ce travail de thèse a montré l'hétérogénéité dans la définition des critères de jugement en rapport avec la ventilation invasive (65). D'autres auteurs ont également montré l'hétérogénéité dans les critères de jugement de réanimation que ce soit pour les critères en rapport avec la ventilation (52), pour les critères en rapport avec les pneumopathies liées aux soins (66) ou l'ensemble des critères de jugement utilisés en réanimation (48). Ces auteurs appellent également à standardiser les critères de jugement afin d'améliorer la recherche. Dès 2000, Hayes et al. (67) font le constat que les critères de jugement sont hétérogènes et principalement centrés sur les critères de jugement en rapport avec la

ventilation invasive. Ils indiquent que cela est inadapté pour les patients sortant vivant de réanimation et que d'autres critères centrés sur le patient doivent être étudiés. En 2017, Pham et al. qui réalisent une mise au point sur l'épidémiologie du syndrome de détresse respiratoire aiguë indiquent que l'un des challenges est le choix des critères de jugement et leurs définitions (68).

6.1.3 Les critères de jugement important pour le patient

Dès 2009, Spragg et al. (47) indiquent que les critères de jugement en rapport avec les défaillances d'organes ne sont importants pour le patient que s'ils s'accompagnent d'une baisse de la mortalité ou bien d'une amélioration de la qualité de vie. En 2000, Hayes et al. (67) ont indiqué les cinq items importants pour les patients en sortie de réanimation :

- retour au travail,
- retour à la maison,
- degré de récupération,
- productivité,
- statut de santé.

En 2017, Gaudry et al. (48) considèrent dans leur classification des critères de jugement de réanimation, comme critères de jugement importants pour le patient : la mortalité quelle que soit le temps auquel on l'évalue, la qualité de vie, le statut fonctionnel, cognitif ou neurologique. Parmi leurs 112 articles analysés, 27 (24%) rapportent les résultats d'essais pour lesquels le critère de jugement principal est important pour le patient. Parmi ces 27 articles, 21 concernent la mortalité (78%) et 6 (22%) la qualité de vie.

Le travail en cours de la COMET initiative sur les critères de jugement pour évaluer le suivi des patients après la réanimation : « development of a core outcomes set for trials of rehabilitation following critical illness » mené par le Docteur Connolly à Londres, s'intègre dans la dynamique actuelle de prise en charge des séquelles des patients qui sortent vivants de réanimation.

6.2 Sensibiliser et former les réanimateurs

A ce jour, les réanimateurs doivent être sensibilisés à la nécessité de mettre en place des « core outcome set » en fonction des pathologies rencontrées en réanimation pour les utiliser dans les essais afin d'améliorer la recherche, comme cela existe pour les pathologies de rhumatologie (OMERACT), pour les douleurs chroniques (IMMPACT) ou pour les cancers (WHO handbook of guidelines on the minimal requirements for data collection in cancer trials) (69). Ce travail de définition des critères de jugement (« core outcome set ») doit être fait en collaboration avec les patients et leurs proches, car les critères de jugement importants pour le patient sont essentiels. Par la suite, les « core outcome set » doivent être utilisés afin de pouvoir comparer les essais entre eux et pouvoir agréger leurs résultats lors de la réalisation de méta-analyses.

Lors de la définition d'un critère de jugement dans le cadre d'un « core outcome set », le groupe de travail doit être aidé de biostatisticiens afin de définir ensemble la meilleure méthode d'analyse statistique à utiliser pour analyser ce critère de jugement.

Plusieurs voies d'information peuvent être utilisées pour informer et sensibiliser les cliniciens sur ces points importants concernant les critères de jugement (leurs définitions et leurs analyses statistiques) :

- les publications dans des revues de réanimation
- les présentations au congrès de réanimation
- les sociétés savantes de réanimation (site internet, réseaux sociaux professionnels et non professionnels)

- les réseaux de recherche en réanimation tels que ARDSnetwork, CRICS-TRIGGERSEP..
- les autorités de santé tels que la Food and Drug Administration (FDA), l'agence européenne du médicament..
- les éditeurs des revues

A noter que ces quatre derniers acteurs ont également un rôle dans la mise en place de ces « core outcome set » comme cela a été le cas pour le « reporting » des articles dans les revues avec le CONSORT.

6.3 Réfléchir aux essais de réanimation de demain

La mise en place de « core outcome set » utilisant des critères de jugement bien définis avec leurs analyses statistiques spécifiques et adaptées doit s'intégrer dans une réflexion plus vaste sur la manière de conduire les essais en réanimation afin d'améliorer leur efficacité. Harhay et al. (50) ont listé les différents points à prendre en compte et ont proposé des recommandations afin d'améliorer les essais de réanimation (Tableau 10). Ils ont également appelé à réaliser des essais plus réalistes. François et al. (70) proposent d'utiliser d'autres types d'essais, tels que des essais adaptatifs ou en cluster.

Tableau 10 : Recommandations pour la rédaction des essais de réanimation, extrait de Harhay et al. (50)

CONCLUSION

La ventilation mécanique invasive est intimement liée à la naissance de la réanimation. Il s'agit de la suppléance d'organes pour le poumon, nécessaire à la survie du patient dans de nombreuses pathologies. La poliomyélite est l'exemple historique. Aujourd'hui, la ventilation mécanique est incontournable dans le syndrome de détresse respiratoire aiguë, qui est un enjeu de santé publique.

Les essais randomisés en réanimation ont des critères de jugement hétérogènes et multiples. Les critères de jugement en rapport avec la ventilation mécanique invasive en sont un exemple. Les définitions sont multiples aussi bien pour la durée de ventilation mécanique que pour le nombre de jours sans ventilation invasive « ventilator-free days ». Les analyses statistiques actuellement utilisées pour ces critères de jugement ne sont pas adaptées car elles ne prennent pas en compte le risque compétitif qu'est la mortalité. Un travail de réflexion doit donc être réalisé par la communauté des réanimateurs en collaboration avec des méthodologistes et des biostatisticiens afin de définir des « core outcome set » adaptés aux pathologies de réanimation. Dans le même temps, une réflexion doit être menée avec les patients de réanimation et leurs proches sur les critères de jugement importants pour le patient « patient reported outcome ». Ces réflexions doivent déboucher sur des recommandations concernant les critères de jugement à utiliser en réanimation avec :

- un nombre limité de critères,

- des critères bien définis et pertinents,
- l'analyse statistique qui doit être utilisée pour chacun d'entre eux.

Ce travail doit être encouragé par les sociétés savantes de réanimation, les éditeurs des revues médicales et les autorités de santé.

Références

1. OMS | Poliomyélite [Internet]. WHO. [cited 2017 Jun 17]. Available from: <http://www.who.int/mediacentre/factsheets/fs114/fr/>
2. Kelly FE, Fong K, Hirsch N, Nolan JP. Intensive care medicine is 60 years old: the history and future of the intensive care unit. *Clin Med Lond Engl*. 2014 Aug;14(4):376–9.
3. Chopin C. L'histoire de la ventilation mécanique: des machines et des hommes. *Réanimation*. 2007 Feb 1;16(1):4–12.
4. Bernard GR. Acute respiratory distress syndrome: a historical perspective. *Am J Respir Crit Care Med*. 2005 Oct 1;172(7):798–806.
5. Kolff WJ. UPDATE ON ARTIFICIAL ORGANS *. *Cardiovasc Dis*. 1975;2(3):273–84.
6. Makdisi G, Wang I. Extra Corporeal Membrane Oxygenation (ECMO) review of a lifesaving technology. *J Thorac Dis*. 2015 Jul;7(7):E166–76.
7. Repères et situations éthiques en médecine [Internet]. [cited 2017 Jun 30]. Available from: http://www.unitheque.com/Livre/ellipses/Sciences_humaines_en_medecine/Reperes_et_situations_ethiques_en_medecine-759.html
8. Réanimation médicale - CNERM [Internet]. [cited 2017 Jul 1]. Available from: <http://www.livres-medicaux.com/reanimation-medicale-cnerm.html>
9. Mollaret P, Goulon M. [The depassed coma (preliminary memoir)]. *Rev Neurol (Paris)*. 1959 Jul;101:3–15.
10. La SRLF - Missions et Historique [Internet]. SRLF. [cited 2017 Jun 12]. Available from: <https://www.srlf.org/mission-et-historique/>
11. SCCM | Critical Care Statistics [Internet]. [cited 2017 May 10]. Available from: <http://www.sccm.org/Communications/Pages/CriticalCareStats.aspx>
12. RFE Intubation Extubation [Internet]. SRLF. [cited 2017 Jun 19]. Available from: <https://www.srlf.org/referentiels/respiratoire/rfe-intubation-extubation/>
13. Trachéotomie en Réanimation [Internet]. SRLF. [cited 2017 Jun 19]. Available from: <https://www.srlf.org/referentiels/respiratoire/tracheotomie-en-reanimation/>
14. 1201-Reanimation-vol21-N1-p96_p105.pdf [Internet]. [cited 2017 Jun 30]. Available

from: https://www.srlf.org/wp-content/uploads/2015/11/1201-Reanimation-vol21-N1-p96_p105.pdf

15. Needham DM, Davidson J, Cohen H, Hopkins RO, Weinert C, Wunsch H, et al. Improving long-term outcomes after discharge from intensive care unit: report from a stakeholders' conference. *Crit Care Med.* 2012 Feb;40(2):502–9.
16. Elliott D, Davidson JE, Harvey MA, Bemis-Dougherty A, Hopkins RO, Iwashyna TJ, et al. Exploring the scope of post-intensive care syndrome therapy and care: engagement of non-critical care providers and survivors in a second stakeholders meeting. *Crit Care Med.* 2014 Dec;42(12):2518–26.
17. Alp E, Voss A. Ventilator associated pneumonia and infection control. *Ann Clin Microbiol Antimicrob.* 2006 Apr 6;5:7.
18. Rotstein C, Evans G, Born A, Grossman R, Light RB, Magder S, et al. Clinical practice guidelines for hospital-acquired pneumonia and ventilator-associated pneumonia in adults. *Can J Infect Dis Med Microbiol.* 2008 Jan;19(1):19–53.
19. Sprung CL, Cohen SL, Sjøkvist P, Baras M, Bulow H-H, Hovilehto S, et al. End-of-life practices in European intensive care units: the Ethicus Study. *JAMA.* 2003 Aug 13;290(6):790–7.
20. Cottureau A, Robert R, le Gouge A, Adda M, Audibert J, Barbier F, et al. ICU physicians' and nurses' perceptions of terminal extubation and terminal weaning: a self-questionnaire study. *Intensive Care Med.* 2016 Aug;42(8):1248–57.
21. Ferrand E, Robert R, Ingrand P, Lemaire F, French LATAREA Group. Withholding and withdrawal of life support in intensive-care units in France: a prospective survey. French LATAREA Group. *Lancet Lond Engl.* 2001 Jan 6;357(9249):9–14.
22. Lesieur O, Leloup M, Gonzalez F, Mamzer M-F, EPILAT study group. Withholding or withdrawal of treatment under French rules: a study performed in 43 intensive care units. *Ann Intensive Care.* 2015 Dec;5(1):56.
23. Le Syndrome de détresse respiratoire aigu de l'adulte - sdra.pdf [Internet]. [cited 2017 Jun 25]. Available from: <http://www.cnerea.fr/UserFiles/File/national/desc-des/livre-masson-2015/respiratoire/sdra.pdf>
24. Petty TL, Ashbaugh DG. The adult respiratory distress syndrome. Clinical features, factors influencing prognosis and principles of management. *Chest.* 1971 Sep;60(3):233–9.

25. Marini JJ, Evans TW. Round table conference: acute lung injury, 15th-17th March 1997 Brussels, Belgium. *Intensive Care Med.* 1998 Aug;24(8):878–83.
26. ARDS Definition Task Force, Ranieri VM, Rubenfeld GD, Thompson BT, Ferguson ND, Caldwell E, et al. Acute respiratory distress syndrome: the Berlin Definition. *JAMA.* 2012 Jun 20;307(23):2526–33.
27. 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.
28. Ferguson ND, Fan E, Camporota L, Antonelli M, Anzueto A, Beale R, et al. The Berlin definition of ARDS: an expanded rationale, justification, and supplementary material. *Intensive Care Med.* 2012 Oct;38(10):1573–82.
29. Acute Respiratory Distress Syndrome Network, Brower RG, Matthay MA, Morris A, Schoenfeld D, Thompson BT, et al. Ventilation with lower tidal volumes as compared with traditional tidal volumes for acute lung injury and the acute respiratory distress syndrome. *N Engl J Med.* 2000 04;342(18):1301–8.
30. Santa Cruz R, Rojas JI, Nervi R, Heredia R, Ciapponi A. High versus low positive end-expiratory pressure (PEEP) levels for mechanically ventilated adult patients with acute lung injury and acute respiratory distress syndrome. *Cochrane Database Syst Rev.* 2013 Jun 6;(6):CD009098.
31. Bellani G, Laffey JG, Pham T, Fan E, Brochard L, Esteban A, et al. Epidemiology, Patterns of Care, and Mortality for Patients With Acute Respiratory Distress Syndrome in Intensive Care Units in 50 Countries. *JAMA.* 2016 Feb 23;315(8):788–800.
32. Glossary | Cochrane Community [Internet]. [cited 2017 Jun 22]. Available from: <http://community.cochrane.org/glossary/o>
33. OMERACT 13 Glossary April 2016.docx - OMERACT_Glossary.pdf [Internet]. [cited 2017 Jun 25]. Available from: https://www.omeract.org/pdf/OMERACT_Glossary.pdf
34. Guyatt GH, Deyo RA, Charlson M, Levine MN, Mitchell A. Responsiveness and validity in health status measurement: A clarification. *J Clin Epidemiol.* 1989;42(5):403–8.
35. Boers M, Kirwan JR, Wells G, Beaton D, Gossec L, d'Agostino M-A, et al. Developing core outcome measurement sets for clinical trials: OMERACT filter 2.0. *J Clin Epidemiol.* 2014 Jul;67(7):745–53.

36. Acquadro C, Berzon R, Dubois D, Leidy NK, Marquis P, Revicki D, et al. Incorporating the patient's perspective into drug development and communication: an ad hoc task force report of the Patient-Reported Outcomes (PRO) Harmonization Group meeting at the Food and Drug Administration, February 16, 2001. *Value Health J Int Soc Pharmacoeconomics Outcomes Res.* 2003 Oct;6(5):522–31.
37. WHO | About Registries [Internet]. [cited 2013 Jan 30]. Available from: <http://www.who.int/ictrp/network/trds/en/index.html>
38. Zarin DA, Tse T, Ide NC. Trial Registration at ClinicalTrials.gov between May and October 2005. *N Engl J Med.* 2005;353(26):2779–87.
39. Zarin DA, Tse T, Williams RJ, Califf RM, Ide NC. The ClinicalTrials.gov Results Database — Update and Key Issues. *N Engl J Med.* 2011;364(9):852–60.
40. The EQUATOR Network | Enhancing the QUAlity and Transparency Of Health Research [Internet]. [cited 2017 Jul 2]. Available from: <http://www.equator-network.org/>
41. Moher D, Hopewell S, Schulz KF, Montori V, Gøtzsche PC, Devereaux PJ, et al. CONSORT 2010 Explanation and Elaboration: Updated guidelines for reporting parallel group randomised trials. *J Clin Epidemiol.* 2010 Aug;63(8):e1-37.
42. The CONSORT Group > CONSORT Statement > 3-12 - Methods > 6a - Outcomes [Internet]. [cited 2013 Jan 30]. Available from: http://www.consort-statement.org/consort-statement/3-12---methods/item6a_outcomes/
43. Home : Core Outcome Measures in Effectiveness Trials Initiative (COMET) [Internet]. [cited 2017 Jul 2]. Available from: <http://www.comet-initiative.org/>
44. Blackwood B, Ringrow S, Clarke M, Marshall J, Rose L, Williamson P, et al. Core Outcomes in Ventilation Trials (COVenT): protocol for a core outcome set using a Delphi survey with a nested randomised trial and observational cohort study. *Trials.* 2015;16:368.
45. Rubenfeld GD, Angus DC, Pinsky MR, Curtis JR, Connors AF Jr, Bernard GR. Outcomes research in critical care: results of the American Thoracic Society Critical Care Assembly Workshop on Outcomes Research. The Members of the Outcomes Research Workshop. *Am J Respir Crit Care Med.* 1999 Jul;160(1):358–67.
46. Wood KA, Huang D, Angus DC. Improving clinical trial design in acute lung injury. *Crit Care Med.* 2003 Apr;31(4 Suppl):S305-311.
47. 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.
48. Gaudry S, Messika J, Ricard J-D, Guillo S, Pasquet B, Dubief E, et al. Patient-important outcomes in randomized controlled trials in critically ill patients: a systematic review. *Ann Intensive Care*. 2017 Dec;7(1):28.
 49. Chevret S. Logistic or Cox model to identify risk factors of nosocomial infection: still a controversial issue. *Intensive Care Med*. 2001 Oct;27(10):1559–60.
 50. 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.
 51. Bodet-Contentin L, Giraudeau B. Revue systématique des critères de jugement utilisés dans les essais de réanimation. Données non publiées;
 52. Blackwood B, Clarke M, McAuley DF, McGuigan PJ, Marshall JC, Rose L. How outcomes are defined in clinical trials of mechanically ventilated adults and children. *Am J Respir Crit Care Med*. 2014 Apr 15;189(8):886–93.
 53. Dechartres A, Altman DG, Trinquart L, Boutron I, Ravaud P. Association between analytic strategy and estimates of treatment outcomes in meta-analyses. *JAMA*. 2014 Aug 13;312(6):623–30.
 54. Savović J, Jones HE, Altman DG, Harris RJ, Jüni P, Pildal J, et al. Influence of reported study design characteristics on intervention effect estimates from randomized, controlled trials. *Ann Intern Med*. 2012 Sep 18;157(6):429–38.
 55. Bernard GR, Wheeler AP, Arons MM, Morris PE, Paz HL, Russell JA, et al. A trial of antioxidants N-acetylcysteine and procysteine in ARDS. The Antioxidant in ARDS Study Group. *Chest*. 1997 Jul;112(1):164–72.
 56. Marini JJ, Evans TW. *Acute Lung Injury. illustrated edition*. Springer-Verlag Berlin and Heidelberg GmbH & Co. K; 1997. 400 p.
 57. Schoenfeld DA, Bernard GR. 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.
 58. Young P, Hodgson C, Dulhunty J, Saxena M, Bailey M, Bellomo R, et al. End points for phase II trials in intensive care: recommendations from the Australian and New Zealand

Clinical Trials Group consensus panel meeting. *Crit Care Resusc J Australas Acad Crit Care Med*. 2012 Sep;14(3):211–5.

59. Needham DM. Understanding and improving clinical trial outcome measures in acute respiratory failure. *Am J Respir Crit Care Med*. 2014 Apr 15;189(8):875–7.

60. Roquilly A, Cinotti R, Jaber S, Vourc'h M, Pengam F, Mahe PJ, et al. Implementation of an evidence-based extubation readiness bundle in 499 brain-injured patients. a before-after evaluation of a quality improvement project. *Am J Respir Crit Care Med*. 2013 Oct 15;188(8):958–66.

61. Mekontso Dessap A, Roche-Campo F, Kouatchet A, Tomicic V, Beduneau G, Sonnevile R, et al. Natriuretic peptide-driven fluid management during ventilator weaning: a randomized controlled trial. *Am J Respir Crit Care Med*. 2012 Dec 15;186(12):1256–63.

62. Alberti C, Métivier F, Landais P, Thervet E, Legendre C, Chevret S. Improving estimates of event incidence over time in populations exposed to other events: application to three large databases. *J Clin Epidemiol*. 2003 Jun;56(6):536–45.

63. Sevransky JE, Checkley W, Martin GS. Critical care trial design and interpretation: a primer. *Crit Care Med*. 2010 Sep;38(9):1882–9.

64. Timsit J-F, de Kraker MEA, Sommer H, Weiss E, Bettiol E, Wolkewitz M, et al. Appropriate endpoints for evaluation of new antibiotic therapies for severe infections: a perspective from COMBACTE's STAT-Net. *Intensive Care Med*. 2017 Jul;43(7):1002–12.

65. 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.

66. Weiss E, Essaïed W, Adrie C, Zahar J-R, Timsit J-F. Treatment of severe hospital-acquired and ventilator-associated pneumonia: a systematic review of inclusion and judgment criteria used in randomized controlled trials. *Crit Care Lond Engl*. 2017 Jun 27;21(1):162.

67. Hayes JA, Black NA, Jenkinson C, Young JD, Rowan KM, Daly K, et al. Outcome measures for adult critical care: a systematic review. *Health Technol Assess Winch Engl*. 2000;4(24):1–111.

68. Pham T, Rubenfeld GD. Fifty Years of Research in ARDS. The Epidemiology of Acute Respiratory Distress Syndrome. A 50th Birthday Review. *Am J Respir Crit Care Med*. 2017

Apr 1;195(7):860–70.

69. Clarke M. Standardising outcomes for clinical trials and systematic reviews. *Trials*. 2007 Nov 26;8(1):39.

70. François B, Clavel M, Vignon P, Laterre P-F. Perspective on optimizing clinical trials in critical care: how to puzzle out recurrent failures. *J Intensive Care*. 2016;4:67.