Available online at www.sciencedirect.com

ScienceDirect

IFAC Papers Online CONFERENCE PAPER ARCHIVE

IFAC-PapersOnLine 48-20 (2015) 273–278

# Reducing the Length of Mechanical Ventilation with Significance: A Case Study of Sample Size Estimation Trial Design Using Monte-Carlo Simulation

**Yeong Shiong Chiew*, Christopher Pretty*, Elena Moltchanova**, Carl Scarrott**, Daniel Redmond*, Geoffrey M. Shaw***, J. Geoffrey Chase***

**Department of Mechanical Engineering, University of Canterbury, Christchurch, New Zealand (e-mail: yeongshiong.chiew@canterbury.ac.nz).*

***School of Mathematics and Statistics, University of Canterbury, Christchurch, New Zealand (email: elena.moltchanova@canterbury.ac.nz)*

**** Intensive Care Unit, Christchurch Hospital, Christchurch, New Zealand, (e-mail: geoff.shaw@cdhb.govt.nz)*

**Abstract:** The power of a study can be established with estimation of total effective sample size ($N_{total}$). In this study, the impact of the length of mechanical ventilation (LoMV) distribution shape in intensive care on the estimated $N_{total}$ is investigated. This study provides an overview on the study design involving LoMV, the resulting potential limitations, and the criteria for a 'successful' design. Data from mechanical ventilated patients in a single-center intensive care unit were used in this study. The $N_{total}$ was estimated using two methods: 1) Model-based Altman's nomogram (a standard); and 2) Monte-Carlo simulation. Using Monte-Carlo simulation, a patient selection criteria is imposed to estimate $N_{total}$ from 'realistic' patient cohorts. The Altman nomogram shows that the $N_{total}$ to detect a 25% change in LoMV (ΔLoMV) at power of 0.8 is ≥1000 patients. For the Monte-Carlo simulation, a $N_{total}$ ≥260 patients is needed to detect similar changes. It is important to consider the LoMV distribution shape and variability, particularly relative to target patient groups who might benefit from the intervention. Assessment of ΔLoMV in response to treatment should be carefully considered to avoid an under-powered studies. The Monte-Carlo simulation combined with objective patient selection provides better design of such studies.



*Keywords:* Length of Mechanical Ventilation, Outcome, Power Analysis, Sample Size.

## 1. INTRODUCTION

Mechanical ventilation (MV) is a primary therapy for patients with respiratory failure. MV is a straightforward therapy, but complex in execution with relatively little consensus in treatment practice. There are several randomised controlled trials (RCT) investigated the effect of MV strategies (The Acute Respiratory Distress Syndrome Network, 2000, Mercat et al., 2008, Brower et al., 2004, Meade et al., 2008), but only a few had shown significant results. Clinical significance was driven by outcome measurements that are surrogates for good outcome from patient-specific response to treatment (Blackwood et al., 2014). In MV, a primary target has always been reducing LoMV (Esteban et al., 2002), and its opposite ventilator free days (VFD). In particular, the mean incremental cost of MV in intensive care unit (ICU) patients was $1522 per day (Dasta et al., 2005). Thus, aside from improving the clinical outcomes, there is also a significant economic motivation to improve MV quality and LoMV.

However, the main issue is that patients are extremely variable (Dickson et al., 2014) and their response to treatment is highly individual and specific (Sundaresan and Chase, 2011). Thus, small reductions in LoMV for a specific study may not be detected if it is under-powered (Whitley and Ball, 2002, Pintado et al., 2013). Table 1 shows several RCTs where only three (The Acute Respiratory Distress Syndrome Network, 2000, Mercat et al., 2008, Strøm et al., 2010) were able to show statistical significance in reducing LoMV.

One possible reason for a non-significant result is an ineffective MV treatment protocol. However, the authors also hypothesize that inter- and intra- patient variability in LoMV response to MV are fundamental causes that affect the statistical power of a study (Altman and Bland, 1995, Van Der Lee et al., 2009). Thus, a clinical trial that aimed to reduce LoMV, or increase VFD, should be designed to target specific patient groups who are likely to benefit from the treatment and whose distribution of patient-specific LoMV is amended to seeing a change for reasonable sample size.

This study aims to provide a more in-depth understanding on how to estimate a total effective sample size ($N_{total}$) for detecting changes in LoMV (ΔLoMV). We compare two effective sample size calculation methods: 1) Model-based Altman sample size calculation; (Altman, 1980, Whitley and Ball, 2002); and a 2) Monte-Carlo Simulation approach using data derived from retrospective patient data from a single centre intensive care unit (ICU). The first method is a well-accepted model-based Altman nomogram (Altman, 1980, Whitley and Ball, 2002, Van Der Lee et al., 2009, Kim and Seo, 2013). The second is a Monte-Carlo simulation-based approach that is hypothesized to produce a more effective and realistic calculation (Oakley et al., 2010).


10.1016/j.ifacol.2015.10.151

**Table 1. LoMV or VFD for several RCT, and their significance value**

| Study | No. Patient | Outcome | Groups (Number of patient) LoMV or Vent Free Days (in mean ± standard deviation or median [interquartile range]) | | p-value |
|---|---|---|---|---|---|
| **ARDSNet (The Acute Respiratory Distress Syndrome Network, 2000)** | 861 | VFD* | Low Vt+ (432) 12±11 | High Vt (429) 10±11 | 0.0070 |
| **ALVEOLI (Brower et al., 2004)** | 549 | VFD | Lower PEEP# (273) 14.5±10.4 | Higher PEEP (276) 13.8±10.6 | 0.5000 |
| **EXPRESS (Mercat et al., 2008)** | 767 | VFD | Minimal distension (382) 3 [0-17] | Increased recruitment (385) 7 [0-19] | 0.0400 |
| **LOVS (Meade et al., 2008)** | 983 | LoMV | Control (507) 10 [6-16] | Lung open (475) 10 [6-17] | 0.9200 |
| **Meta-Analysis (Briel et al., 2010)** | 2299 | VFD | Lower PEEP (1136) 11 [0-21] | Higher PEEP (1163) 13 [0-22] | 0.1000 |
| **Individualised PEEP (Pintado et al., 2013)** | 70 | VFD | Control (36) 0 [0-15.75] | Intervention (34) 1 [0-18] | 0.1600 |
| **Sedation Study (Strøm et al., 2010)** | 113 | VFD | Control (58) 18.0 [0-24.1] | No Sedation (55) 6.9 [0-20.5] | 0.0191 |

*VFD - Number of ventilator free days, days 1 to 28, in which the patient had been breathing without assistance for at least 48 consecutive hours, +Vt - tidal volume, #PEEP - Positive end expiratory pressure

In particular, a patient selection method can be implemented in the Monte-Carlo simulation approach to simulate a realistic clinical trial conditions. It also provides a means of testing whether inter-patient variability in a recruited cohort will affect study outcome/ power. This approach is enabled by the increasing ease of which patient 'meta-data' (Age, Sex, LoMV, Acute Physiology and Chronic Health Evaluation (APACHE) II or III Scores) can be gathered. More importantly, Monte-Carlo simulation makes no assumption on distribution shapes that may not be realistic for some target cohorts.

This approach thus gives an overview of study design involving LoMV, potential limitations, and the criteria for designing a potentially 'successful' study.

## 2. METHODS

### *2.1 Patient Information: Length of Mechanical Ventilation*

Data from patients admitted to the Christchurch Hospital ICU from 2011 to 2013 were used. Of the 3907 patients admitted, 2921 patients (75%) required MV, and 2534 (65%) patients were invasively ventilated through intubation or tracheotomy. For all 2534 invasively ventilated patients, their APACHE III diagnostic code and LoMV during invasive ventilation are included. The 2534 invasively ventilated patients are set as Cohort A. The mean (± SD) LoMV for Cohort A is 3.23 ± 7.03 days (median = 0.72 days [IQR: 0.24-2.62]).

Realistic potential trial exclusion criteria were imposed to Cohort A using APACHE III diagnostic code to obtain a targeted retrospective cohort, Cohort B. These criteria have been used in prior studies and they are: (1) Patients who are likely to be discontinued from MV within 24 hours, (2) Patients with increase intracranial pressure, (3) Patients who have significant weakness from any neurological disease, (4) Patients who have asthma as the primary presenting condition or a history of significant chronic obstructive pulmonary disease and (5) Patients who are pregnant (The Acute Respiratory Distress Syndrome Network, 2000, Brower et al., 2004, Mercat et al., 2008, Meade et al., 2008, Pintado et al., 2013, Strøm et al., 2010).

Specifically, patients with following APACHE III diagnostic code (ANZICS, 2013) were excluded: Non-operative neurological (400), post-operative neurological (1500), chronic obstructive pulmonary disease (206), asthma (209), head trauma with or without multi trauma (601), multi trauma with spinal injury (604), isolated cervical spine injury (605), post operation patients: head trauma with or without multi trauma (1601), post operation patients: multi trauma with spinal injury (1604), post operation patients: isolated cervical spine injury (1605) and pregnancy-related disorder (1802). Patients with LoMV less than 1 day and more than 30 days were also excluded. These exclusion criteria were chosen based the clinical implication that these patient may not benefit from a MV intervention, or could be harmed in some cases. It is also easy to implement in simulation and in a clinical trial.

After imposing the exclusion criteria, the number of patients who might benefit and be eligible for the study was reduced to 744 (19% of total patients admitted to ICU or 29% of patients requiring invasive MV). The LoMV for Cohort B was mean 5.81 ± 6.30 days (median = 2.92 days [IQR: 1.67-7.38]), and is significantly different from Cohort A ($p < 0.05$ using Student *t*-test and Wilcoxon ranksum test and Kolmogorov-Smirnov test).

### *2.2 Sample Size Estimation Methods*

#### *2.2.1 Model-based Altman Nomogram*

The Altman nomogram (Altman, 1980, Whitley and Ball, 2002) uses an estimation of standardised difference of a clinical surrogate and desired statistical. The standardised

differences in this study were obtained using the mean (± SD) of Cohorts A and B. The $N_{total}$ is determined for several standardised differences for 0.8 power. It should be noted that the normality assumption in this approach matches those of many other methods (Van Der Lee et al., 2009, Kim and Seo, 2013, van der Tweel et al., 2012, Lachin, 1981). Thus, these methods may not be suitable to estimate $N_{total}$ for clinical surrogates that are severely skewed and heavy upper-tailed, like LoMV and many others despite their frequent use.

### 2.2.2 *Monte-Carlo Simulation Sample Size Estimation*

Monte-Carlo simulation is a simple standardised method and provides a more realistic estimate of the sample size needed to achieve the required changes in LoMV. We have used it to simulate ΔLoMV by percentage or any other method, instead of standardised difference. Thus, in response to an effective treatment, patients with longer LoMV may experience higher ΔLoMV, whereas patients with lower LoMV may see a lesser or zero absolute change as might be expected clinically. Thus, such a relative change is comparatively more clinically relevant, and likely to match actual trial conditions.

Monte-Carlo simulation is applied to both Cohorts A and B to determine the power for different values of $N_{total}$. Each power analysis is carried out in five hierarchical steps shown in Table 2, with 10,000 iterations. All simulations were performed using MATLAB (R2014a, The Mathworks, Natick, Massachusetts, USA). A one-sided Student *t*-test (*t*-test), a one-sided Wilcoxon ranksum (RS-test) and a Kolmogorov-Smirnov test (KS-test) were used for significance testing of the difference in mean and other distribution characteristic (median and variability). In addition, a separate Student t-test was applied following log transformation of highly skewed LoMV data prior to the Student *t*-test to account for those studies which properly treat skewed data, to make the assumptions of approximate normality of the population more realistic.

## 3. RESULTS

### 3.1 *Altman Nomogram Sample Size Estimation*

For Cohort A shown in Fig 1, the minimum estimated $N_{total}$ was 170 for a 0.43 standardised difference ΔLoMV (~3 days or ~94% difference from the mean). The $N_{total}$ increases to 5000 patients for a 0.1 standardised difference in LoMV (~1 day or 33% difference from the mean).

Fig. 2 shows the $N_{total}$ for the targeted cohort, Cohort B. The minimum $N_{total}$ estimated was 140 for 0.47 standardised difference in LoMV (~3 days or ~52% difference from the Cohort B mean). The $N_{total}$ increases to 4000 patients for a 0.08 standardised difference in LoMV (~1 day or 17% difference from the Cohort B mean). It is clear that the $N_{total}$ increases exponentially for every reduction of standardised difference.

### 3.2 *Monte-Carlo Simulation Sample Size Estimation*

Fig. 3 shows the statistical power for each sample size and reasonable ΔLoMV ranging from 5-30% using Monte-Carlo analysis for Cohort A. The KS-test compares the shape of the ΔLoMV distribution for both control and intervention group, and it requires smaller sample size to see the difference compared to other significance tests. The RS-test comparing the median and *t*-test on the log scale (comparing the mean) overlapped in sample size estimation, as might be expected. In this study, the *t*-test on the original (un-logged) scale had the lowest power of all the statistical tests. Fig. 4 shows the statistical power for each $N_{total}$, at various ΔLoMV for Cohort B, where the $N_{total}$ required are much smaller compared to Cohort A.

**Table 2. Sequence of the power analysis for each Monte-Carlo simulation**

| | Step | Description | Tests |
|---|---|---|---|
| 1. | **Patient Cohorts** | • Select a patient cohort:<br>1) Cohort A includes all invasively ventilated patients.<br>2) Cohort B is created from Cohort A by imposing exclusion criteria. | • Cohort A<br>• Cohort B |
| 2. | **Sample Size Selection** | • Randomly select patients from a patient cohort (Cohort A or Cohort B) and assign each patient to a treatment group. 1) Control group or 2) Intervention group.<br>• The cohort patient is selected with replacement.<br>• Various sample sizes of each treatment group is tested.<br>• Total sample size = Control sample size + Intervention sample size | • Total sample size = 200, 300, 400… 5000 patients.<br>• Or other total sample sizes with higher resolution such as 200, 220, 240… 600 patients. |
| 3. | **Difference in LoMV** | • Impose a difference in LoMV between two groups.<br>• The LoMV in Intervention group is reduced by the chosen percentage. LoMV intervention = LoMV patient × (100% -Percentage reduction)<br>• The difference in LoMV ranges from 5 to 30% of total LoMV. | • Difference of LoMV = 5, 10, 15, … 30% |
| 4. | **Statistical Test** | • Perform statistical test comparing the LoMV between two groups.<br>• Using parametric and non-parametric tests.<br>• A value of $p < 0.05$ is indicates that LoMV for intervention group is significantly different from control group. | • Student *t*-test<br>• Student *t*-test (log scale)<br>• Wilcoxon Ranksum test<br>• Kolmogorov-Smirnov test |
| 5. | **Power Analysis** | • Each Monte-Carlo simulation iteration will generate a p-value for each statistical test.<br>• For a given sample size and significance level α, statistical power is evaluated as the proportion of iterations for which the $p < \alpha$. | • E.g. for 10000 Monte-Carlo iterations, if $p < \alpha$ for 84% (8400 iterations), Power = 0.84. |

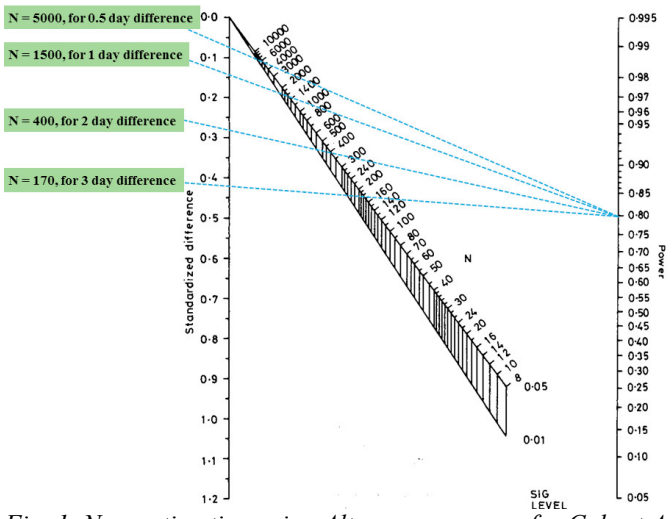


*Fig. 1. $N_{total}$ estimation using Altman nomogram for Cohort A at 0.8 power (Whitley and Ball, 2002).*

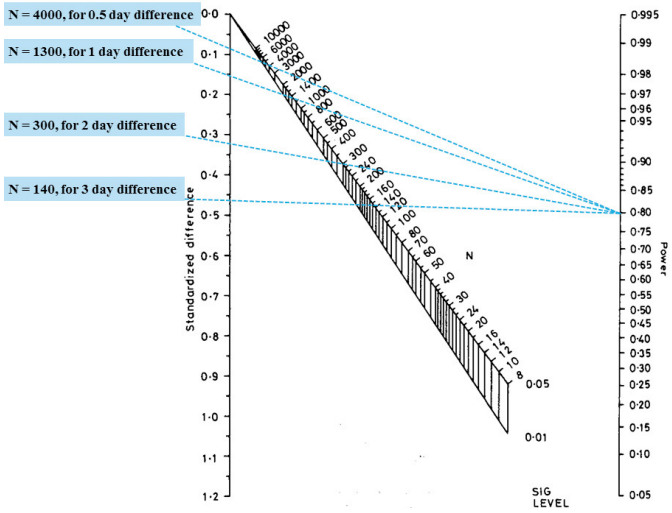


*Fig. 2. $N_{total}$ estimation using Altman nomogram for Cohort B at 0.8 power (Whitley and Ball, 2002) .*

## 4. DISCUSSION

### *4.1 Sample Size Estimation using Altman Nomogram*

Results in Fig. 1 suggest that a relatively large $N_{total}$ is required to detect a reasonable 0.05-0.15 standardised differences (0.5 to 1 day or 10-33%) in LoMV for Cohort A (an average of ≥1500 patients). $N_{total}$ of 1300 patients are required to detect 0.05-0.15 standardised difference (0.5 to 1 day or 10-33%) with 0.8 power for Cohort B.

Cohort A has highly skewed LoMV. Thus, the nomogram, which assumes a normal distribution for LoMV may not be suitable for $N_{total}$ estimation even using log transformed data. Even for Cohort B, where it is much less skewed, $N_{total}$ decreases only slightly. Thus, there is a need to specify the $N_{total}$ using different method, as well as the ability to include patient selection criteria to target a more specific patient and LoMV distribution (Van Der Lee et al., 2009). For that type of analysis, a Monte-Carlo based simulation is better to assess power.

### *4.2 Sample Size Estimation using Monte-Carlo Simulation*

In this study, the Monte-Carlo simulation performed using Cohort A showed that a very large sample size is required to detect a reasonable ΔLoMV. It was impossible to detect 10% changes or lower in LoMV in this cohort with power of more than 0.8 within the sample sizes considered, as shown in Fig 3. A power of 0.8 can only be achieved when the ΔLoMV is increased to 15% or more. This result clearly shows that a statistically significant study to detect ΔLoMV must be conducted in a more targeted patient distribution that is still easily segregated from the full cohort.

One of the main causes of being under-powered is the skewed population (Van Der Lee et al., 2009), which has high variability and a heavy upper tail. Thus, it is very important to impose strict inclusion and exclusion criteria to filter outlying patients that also may not benefit from a new treatment. For example, surgical patients that required very short LoMV (≤1 day) or patients with very long LoMV due to clinical reasons other than respiratory failure (≥30 days, e.g. due to head or spinal injuries). However, if the treatment affects all patients in the same multiplicative way, then it should be noted that statistical power for tests on absolute difference can be lost by removing patients with longest LoMV (as the absolute treatment effect is larger for them).

One of the ideal inclusion criteria in other studies is to focus on patients with a more severe form of respiratory failure, such as the acute respiratory distress syndrome (ARDS). And, as observed in this study, when other patients were excluded, the $N_{total}$ was significantly smaller in Cohort B when compared to the all patients Cohort A.

There are various guides on how to calculate an effective sample size (Van Der Lee et al., 2009, Kim and Seo, 2013, van der Tweel et al., 2012, Lachin, 1981). These studies emphasise the need for normally distributed population and use a standardised difference. However, in this study, we observed that LoMV, like many outcome measurements cannot be characterised by a normal distribution, and that performing data transformations (log-transformation considered here) may not completely solve the problem. Monte-Carlo simulation provides the opportunity to estimate $N_{total}$ without such as strong population assumption and specific, objectively determined cohorts. Estimation of $N_{total}$ for these cohorts can be performed using retrospective cohort data and Monte-Carlo simulation.

The use of clinically realistic cohorts and objective exclusion criteria, such as APACHE II or APACHE III diagnostics codes (ANZICS, 2013), is easily replicated in a real trial. Equally, the ability to use a percentage ΔLoMV or similar, such as a more clinically realistic percent LoMV change, is another advantage in making the $N_{total}$ estimation better match what could occur in reality. Thus, all these advantages provide an estimate that is more realistic and more likely to be matched in trial conditions, providing greater certainty of achieving the necessary power.

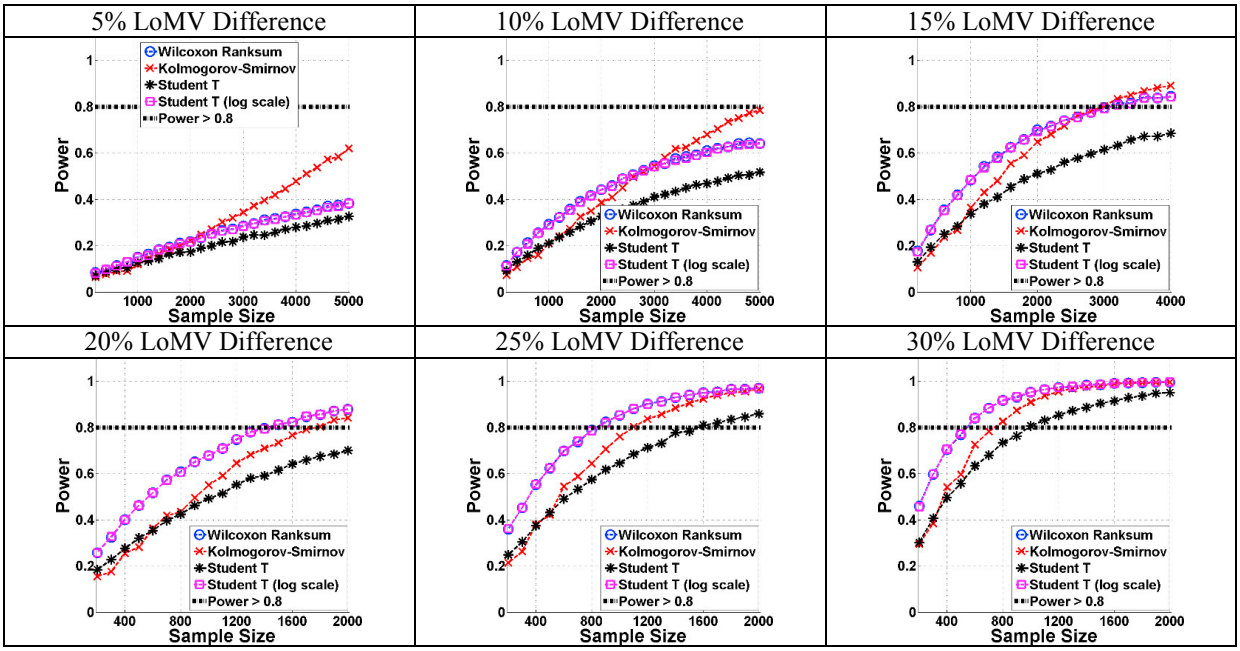


*Fig. 3. Sample size statistical power with different ΔLoMV using Cohort A. (Top row from left to right) 5% difference, 10% difference and 15% difference, (Bottom Row from left to right) 20% difference, 25% difference and 30% difference.*

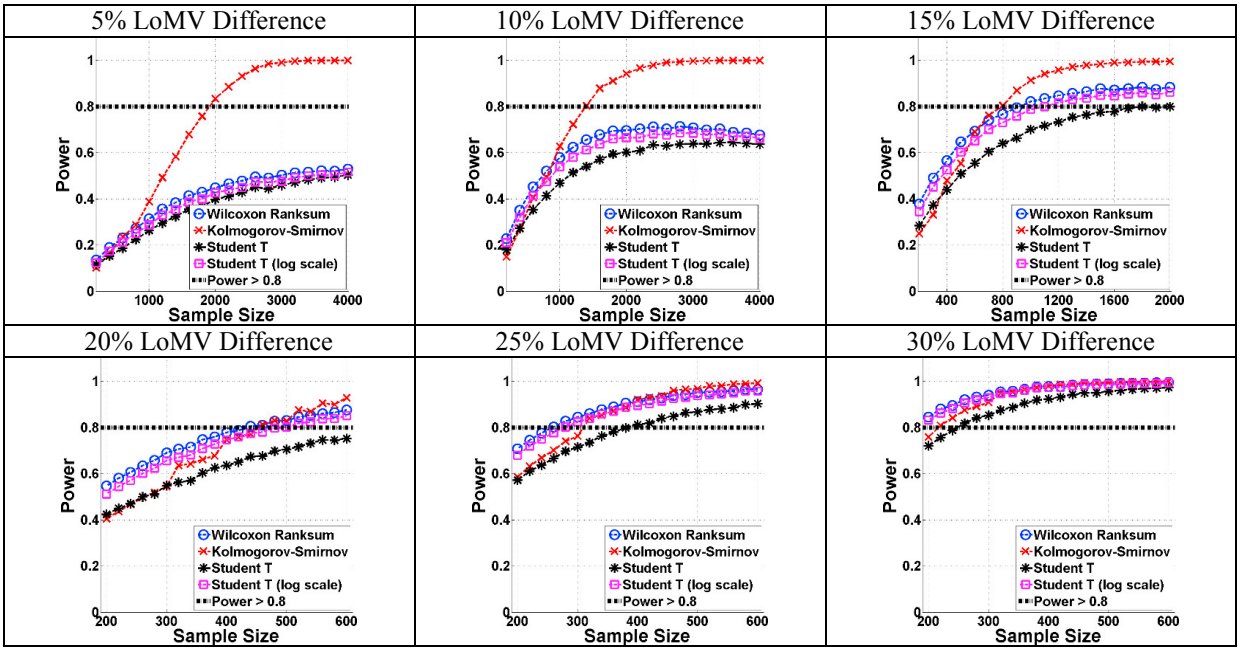


*Fig. 4. Sample size statistical power with different ΔLoMV using Cohort B. (Top row from left to right) 5% difference, 10% difference and 15% difference, (Bottom Row from left to right) 20% difference, 25% difference and 30% difference.*

### *4.3 Study Limitation and Recommendation*

One of the major concerns is that LoMV distributions may vary between centres (Van Der Lee et al., 2009). This variability in patient distribution means that the $N_{total}$ derived from this study may only be applicable to the participating centre or other regional centres that have similar characteristics (Van Der Lee et al., 2009). However, the solution and approach presented in this study is generalisable and easily repeated across multiple centres, separately or together. A reasonably small, center-specific pilot study may be used to obtain information on LoMV distribution if it is not available.

The changes of LoMV used in this study were arbitrarily chosen and may not represent the true possible LoMV change for any given intervention (Schulz and Grimes, 2005). In fact, this percentage or absolute changes in LoMV is not reported in most studies, as shown in Table 1, as it is dependent on the effect of the treatment and the specific target of the intervention. However, LoMV, and its surrogate VFD, are still the most and frequently used metrics. Thus, given the effective sample size and percentage difference in LoMV needed, the results suggest that the clinical outcome (LoMV)

requires a large sample size due to the high variability in the population.

## 5. CONCLUSIONS

In this study, $N_{total}$ to detect changes in LoMV are examined using a standard nomogram and through Monte-Carlo simulation. The nomogram is not suitable to estimate the sample size of a standardised difference from typically highly skewed and heavy tailed data. In these cases, the Monte-Carlo simulation can be used to determine the trial sample at a pre-set power. For an intervention study that aims to reduce LoMV, it is important to consider those who might benefit from the treatment, targeting patients who have moderate LoMV in a way that is objective and can be easily replicated in the study. Finally, even after strict inclusion and exclusion criteria, ΔLoMV is a difficult metric and may be under-powered without larger cohorts.

## REFERENCES


Altman, D. and Bland, J. (1995). Absence of evidence is not evidence of absence. *BMJ,* volume 311**,** 485.

Altman, D. G. (1980). Statistics and ethics in medical research: III How large a sample? *British Medical Journal,* volume 281**,** 1336-1338.

Anzics 2013. Centre for Outcome and Resource Evaluation (CORE) Adult Patient Database. Victoria, Australia.

Blackwood, B., Clarke, M., Mcauley, D. F., Mcguigan, P. J., Marshall, J. C. and Rose, L. (2014). How Outcomes Are Defined in Clinical Trials of Mechanically Ventilated Adults and Children. *American Journal of Respiratory and Critical Care Medicine,* volume 189**,** 886-893.

Briel, M., Meade, M., Mercat, A., Brower, R. G., Talmor, D., Walter, S. D., Slutsky, A. S., Pullenayegum, E., Zhou, Q., Cook, D., Brochard, L., Richard, J.-C. M., Lamontagne, F., Bhatnagar, N., Stewart, T. E. and Guyatt, G. (2010). Higher vs lower positive end-expiratory pressure in patients with acute lung injury and acute respiratory distress syndrome: Systematic review and meta-analysis. *JAMA,* volume 303**,** 865-873.

Brower, R. G., Lanken, P. N., Macintyre, N., Matthay, M. A., Morris, A., Ancukiewicz, M., Schoenfeld, D. and Thompson, B. T. (2004). Higher versus Lower Positive End-Expiratory Pressures in Patients with the ARDS. *N Engl J Med,* volume 351**,** 327-336.

Dasta, J. F., Mclaughlin, T. P., Mody, S. H. and Piech, C. T. (2005). Daily cost of an intensive care unit day: The contribution of mechanical ventilation. *Crit Care Med,* volume 33**,** 1266-1271.

Dickson, J., Gunn, C. and Chase, J. (2014). Humans are horribly variable. *International Journal of Clinical and Medical Imaging,* volume Volume 1(2)**,**, 1-page.

Esteban, A., Anzueto, A., Frutos, F., Alia, I., Brochard, L., Stewart, T. E., Benito, S., Epstein, S. K., Apezteguia, C., Nightingale, P., Arroliga, A. C. and Tobin, M. J. (2002). Characteristics and outcomes in adult patients receiving mechanical ventilation: A 28-day international study. *JAMA,* volume 287**,** 345-355.

Kim, J. and Seo, B. S. (2013). How to Calculate Sample Size and Why. *Clin Orthop Surg,* volume 5**,** 235-242.

Lachin, J. M. (1981). Introduction to sample size determination and power analysis for clinical trials. *Controlled Clinical Trials,* volume 2**,** 93-113.

Meade, M. O., Cook, D. J., Guyatt, G. H., Slutsky, A. S., Arabi, Y. M., Cooper, D. J., Davies, A. R., Hand, L. E., Zhou, Q., Thabane, L., Austin, P., Lapinsky, S., Baxter, A., Russell, J., Skrobik, Y., Ronco, J. J., Stewart, T. E. and For the Lung Open Ventilation Study, I. (2008). 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,* volume 299**,** 637-645.

Mercat, A., Richard, J.-C. M., Vielle, B., Jaber, S., Osman, D., Diehl, J.-L., Lefrant, J.-Y., Prat, G., Richecoeur, J., Nieszkowska, A., Gervais, C., Baudot, J., Bouadma, L., Brochard, L. and Group, E. P. E. S. (2008). Positive End-Expiratory Pressure Setting in Adults With ALI and ARDS: A Randomized Controlled Trial. *JAMA,* volume 299**,** 646-655.

Oakley, J. E., Brennan, A., Tappenden, P. and Chilcott, J. (2010). Simulation sample sizes for Monte Carlo partial EVPI calculations. *Journal of Health Economics,* volume 29**,** 468-477.

Pintado, M.-C., De Pablo, R., Trascasa, M., Milicua, J.-M., Rogero, S., Daguerre, M., Cambronero, J.-A., Arribas, I. and Sánchez-García, M. (2013). Individualized PEEP Setting in Subjects With ARDS: A Randomized Controlled Pilot Study. *Respiratory Care,* volume 58**,** 1416-1423.

Schulz, K. F. and Grimes, D. A. (2005). Sample size calculations in randomised trials: mandatory and mystical. *The Lancet,* volume 365**,** 1348-1353.

Strøm, T., Martinussen, T. and Toft, P. (2010). A protocol of no sedation for critically ill patients receiving mechanical ventilation: a randomised trial. *The Lancet,* volume 375**,** 475-480.

Sundaresan, A. and Chase, J. G. (2011). Positive end expiratory pressure in patients with acute respiratory distress syndrome - The past, present and future. *Biomedical Signal Processing and Control,* volume 7**,** 93-103.

The Acute Respiratory Distress Syndrome Network (2000). 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,* volume 342**,** 1301-1308.

Van Der Lee, J. H., Tanck, M. W. T., Wesseling, J. and Offringa, M. (2009). Pitfalls in the design and analysis of paediatric clinical trials: a case of a 'failed' multi-centre study, and potential solutions. *Acta Pædiatrica,* volume 98**,** 385-391.

Van Der Tweel, I., Askie, L., Vandermeer, B., Ellenberg, S., Fernandes, R. M., Saloojee, H., Bassler, D., Altman, D. G., Offringa, M. and Van Der Lee, J. H. (2012). Standard 4: Determining Adequate Sample Sizes. *Pediatrics,* volume 129**,** S138-S145.

Whitley, E. and Ball, J. (2002). Statistics review 4: Sample size calculations. *Critical Care,* volume 6**,** 335 - 341.