**Secondary outcome data collection for the Advanced Trauma Life Support stepped wedge trial**

**PI: Martin Gerdin Wärnberg**

**Jessica Kasza, 2024-10-22**

**Original trial design**: batched stepped wedge trial, with 30 clusters; 6 batches of:

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 0 | 0 | 0 | 0 |  | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 0 | 0 | 0 | 0 | 0 |  | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 0 | 0 | 0 | 0 | 0 | 0 |  | 1 | 1 | 1 | 1 | 1 | 1 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 |  | 1 | 1 | 1 | 1 | 1 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |  | 1 | 1 | 1 | 1 |

This trial has been powered to detect a change in the primary outcome of in-hospital mortality within 30 days of arrival at emergency department. Each period is one month in length, and it is assumed that 12 patients will have the primary outcome available in each cluster in each month. Further details are available in the protocol.

Some secondary outcomes are more burdensome to collect:

* Adherence to ATLS principles. This is defined as the % of items adhered to in a 14-item checklist. I would say that we can expect 50% adherence at baseline and would like to be able to detect an increase to 70% after training (https://doi.org/10.1007%2Fs00068-019-01181-7, https://doi.org/10.1007/s00268-016-3759-8)
* Quality of life, measured using the EQ5D health status score. We can expect a score of 70 (range 0-100) before training and would like to detect an increase to 75 after training (https://doi.org/10.1101/2024.09.28.24314529).
* Disability, measured using WHODAS 2.0 (range 0-100). An educated guess on the baseline value would be 25, and then we would ideally like to detect a decrease to 22.5 after training (<https://doi.org/10.51893/2021.1.OA10>).

Question: Can the data collection for these outcomes be limited? That is, can a staircase design be applied to collect data for these outcomes?

Notes:

* In the designs below, we will assume that different period terms are included for each of the 6 batches, implying that sample size calculations can be conducted assuming that there are 6 clusters randomised to each of the 5 sequences of the considered designs. For Design 1, there is no overlap in data collection for these outcomes across batches; for Designs 2 and 3, there is some overlap, so gains in study power could be made by assuming period effects are shared across batches.
* Effect size of 0.5 assumed for Quality of Life and Disability outcomes.
* Note that these calculations have not allowed for small sample corrections, as would be recommended for a design with 30 clusters (and are being applied for the primary outcome).

**Design 1: 1 pre-switch and 1 post-switch data collection period:**

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  | 0 |  | 1 |  |  |  |  |  |  |  |
|  |  |  |  | 0 |  | 1 |  |  |  |  |  |  |
|  |  |  |  |  | 0 |  | 1 |  |  |  |  |  |
|  |  |  |  |  |  | 0 |  | 1 |  |  |  |  |
|  |  |  |  |  |  |  | 0 |  | 1 |  |  |  |

* Quality of Life and Disability outcomes:
  + With 30 clusters in total (6 per sequence), for a discrete time decay correlation structure, with ICCs of 0.01 to 0.15 and a CAC of 0.8, there is >80% power to detect an effect size of 0.5 with 5 measurements in each cluster in each period. (Total number of outcomes available for analysis: 30\*2\*5 = 300)
* Adherence:
  + An increase from 50% to 70% can be detected with >80% power with 6 measurements in each cluster in each period, with power maintained for ICCs in the range 0.01 to 0.05, and a CAC of 0.8 (for a discrete time decay correlation structure). (Total number of outcomes available for analysis: 30\*2\*6 = 360)
* Further considerations: this design may be particularly vulnerable to cluster drop-out, and it is unclear to me (at this stage) how cluster drop-out and the batched nature of the design might interact. So although this design does appear to provide sufficient power to detect effects of interest, a design with a greater number of pre- and post-switch measurement periods may be more robust to cluster and participant drop-out.

**Design 2: 2 pre-switch and 2 post-switch data collection periods:**

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | 0 | 0 |  | 1 | 1 |  |  |  |  |  |  |
|  |  |  | 0 | 0 |  | 1 | 1 |  |  |  |  |  |
|  |  |  |  | 0 | 0 |  | 1 | 1 |  |  |  |  |
|  |  |  |  |  | 0 | 0 |  | 1 | 1 |  |  |  |
|  |  |  |  |  |  | 0 | 0 |  | 1 | 1 |  |  |

* Quality of Life and Disability outcomes:
  + With 30 clusters in total (6 per sequence), for a discrete time decay correlation structure, with ICCs of 0.01 to 0.15 and a CAC of 0.8, there is >80% power to detect an effect size of 0.5 with 4 measurements in each cluster in each period. (Total number of outcomes available for analysis: 30\*4\*4 = 480)
* Adherence:
  + An increase from 50% to 70% can be detected with >80% power with 5 measurements in each cluster in each period, with power maintained for ICCs in the range 0.01 to 0.05, and a CAC of 0.8 (for a discrete time decay correlation structure). (Total number of outcomes available for analysis: 30\*4\*5 = 600)

**Design 3: 3 pre-switch and 3 post-switch data collection periods:**

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | 0 | 0 | 0 |  | 1 | 1 | 1 |  |  |  |  |  |
|  |  | 0 | 0 | 0 |  | 1 | 1 | 1 |  |  |  |  |
|  |  |  | 0 | 0 | 0 |  | 1 | 1 | 1 |  |  |  |
|  |  |  |  | 0 | 0 | 0 |  | 1 | 1 | 1 |  |  |
|  |  |  |  |  | 0 | 0 | 0 |  | 1 | 1 | 1 |  |

* Quality of Life and Disability outcomes:
  + With 30 clusters in total (6 per sequence), for a discrete time decay correlation structure, with ICCs of 0.01 to 0.15 and a CAC of 0.8, there is >80% power to detect an effect size of 0.5 with 4 measurements in each cluster in each period. (Total number of outcomes available for analysis: 30\*6\*4 = 720)
* Adherence:
  + An increase from 50% to 70% can be detected with >80% power with 4 measurements in each cluster in each period, with power maintained for ICCs in the range 0.01 to 0.05, and a CAC of 0.8 (for a discrete time decay correlation structure). (Total number of outcomes available for analysis: 30\*6\*4 = 720)

**Comments:**

* I recommend Design 2 or 3 (due to unknown impact of cluster drop-out in batched staircase designs), collecting outcomes for as many participants as possible, with minimums as indicated above.
* Given the high levels of power obtained for moderate numbers of participants in each cluster in each period, could consider reducing the effect sizes of interest.
* Designs 1, 2 and 3 above are a starting point for discussions; we can consider variations on these, alternative choices for correlation parameters, and alternative effect sizes.