Statistical Analysis Plan-SPLIT ENZ

Overview

This Statistical Analysis Plan (SAP) outlines the general descriptive and analytical approach to the SPLIT ENZ study. This is a multi-centre observational (non-experimental) cohort study about survivorship outcomes after Intensive Care Unit (ICU) treatment in New Zealand. This study will be the first to describe disability and other outcomes for survivors of critical illness in a New Zealand cohort.

There are seven aims for the quantitative data collection for the study. A qualitative phase of the study has been analysed and reported separately. In summary these aims are:

1. To describe disability in survivors of critical illness needing ICU treatment in New Zealand.
2. To explore potential predictors of disability in survivors of critical illness needing ICU treatment after 6-months.
3. To explore the pattern of change with time for a set of key outcomes in survivors of critical illness needing ICU treatment.
4. To explore whether the key measurement of disability used in this study, the WHODAS, captures domains associated with the Post-Intensive Care Syndrome (PICS).
5. To explore return-to-work for survivors of critical illness needing ICU treatment.
6. To explore health care utilisation after discharge for survivors of critical illness needing ICU treatment.
7. To explore cognitive status and its potential predictors for survivors of critical illness needing ICU treatment.

Study Aims

1. Disability

The primary patient-related outcome in this study is disability. This is measured by the WHODAS instrument which was assessed at three times: Between 4 and 6 weeks after discharge, to capture baseline disability after discharge, 6-months after discharge, and one year after discharge.

The 12-item WHODAS 2.0 score range of minimum and maximum values is from 12 to 60 (12 being no difficulty in any of the domains and 60 being extreme difficulty in all domains). For the purposes of this study, it will also be rescaled to 0 to 100 by subtracting 12 from the score and multiplying by 100/48; a conversion factor of 2.083. This will be done so that disability can also be defined as mild or moderate to severe. This will be based on a rescaled score of between 0 and 24 (original score 12 to 24) to define mild, and moderate to severe is a rescaled score of 25 to 100 (original score 25+). The reason for using this rescaling and dichotomous classification is so that disability can be more directly compared with other research which has expressed disability in this way (Hodgson et al., 2017; Shulman et al., 2015).

The WHODAS will be described both on the continuous scales using data descriptors mean and standard deviation (SD), median and 25^{th} to 75^{th} percentiles (as the inter-quartile range), and minimum to maximum values (range) supplemented by boxplots and frequency histograms. The proportion of participants, summarised by numerators and denominators and expressed as a percentage, will be reported for each time of measurement.

Confidence intervals for these proportions will be estimated by an exact and conservative approach such as the Clopper Pearson method.

A small proportion of ICU survivors may have died during the follow up year. These proportions will be described as above and by a sensitivity analysis where those who are died are classified in the group having moderate to severe disability.

2. Disability predictors after six months

The selection of potential predictors of disability in survivors of critical illness treated in intensive care units is based mainly on reports from international studies about risk factors or predictors for Post-Intensive Care Syndrome (PICS). (Hodgson et al., 2017; Jackson et al., 2014; Marra et al., 2018) The potential predictors, classified by data type supplemented by a brief description is shown in the Table:

Table One: Potential predictors of disability

Scale or ordinal variables	**Description**
Age	At ICU admission, measured in years
Duration of Ventilation in ICU	Measured in hours
Duration of delirium	Measured in hours
ICU length of stay	Measured in days
Total Hospital length of stay	Measured in days
Best ICU mobility score at ICU discharge	The ICU Mobility Scale provides an 11-point ordinal scale, ranging from nothing (lying/passive exercises in bed, score of 0) to independent ambulation (score of 0-10)
Clinical frailty scores (CFS)	The CFS is an assessment tool that evaluates comorbidity, function, and cognition. The CFS (range 1–8) categorises patients as non-frail (1 = very fit; 2 = well; 3 = managing well; 4 = vulnerable) or frail (5 = mildly frail; 6 = moderately frail; 7 = severely frail; 8 = very severely frail
SOFA score	The SOFA score objectively describes organ (dys) function during the first 24 hours of admission, by applying a count of six different scores, one each for the respiratory, cardiovascular, hepatic, coagulation, renal and neurological systems from 0 to a maximum score of 24
Charlson Comorbidity Index	This score is a standard measurement of morbidity that can also be used to predict mortality based on weighted scores of age and comorbidities and are reported as scores ranging between 0 and 37.
APACHE II & III scores	APACHE II is a widely used ICU mortality prediction score and the APACHE III provides initial risk stratification for severely ill hospitalized patients and an estimate for hospital mortality. Scores are reported ranging between 0 and 71 (APACHE II) and between 0-299 (APACHE III).
Categorical variables	
Biological Sex	Classified as male or female
Ethnicity	Classified according to the entered ethnicity on admission in the ICU database* Full list here: Ethnicity classification SPLIT ENZ.docx
Recruitment Centre	Site of ICU treatment
History of depression or anxiety or post-traumatic stress disorder (PTSD)	
Diagnostic Category	Classified according to the ICD-10 codes entered into the ICU database on admission.
Treatment category	Operative (further categorised into surgical, vascular) Cardiothoracic (elective, emergency) and non-operative (medical) patient categories

Local database that contributes to Australian and New Zealand Intensive Care Society (ANZICS) Centre for Outcome and Resource Evaluation (CORE) Adult Patient Database (APD)

Data descriptions of scale and ordinal variables will use as data descriptors mean and standard deviation (SD), median and 25th to 75th percentiles (as the inter-quartile range), and minimum to

maximum values (range) supplemented by boxplots and frequency histograms. Data distributions of categorical variables will be by counts (numerators and denominators) and proportions described as percentages.

Two approaches will be used to describe associations between disability and these potential predictors:

1. Firstly, disability defined as a dichotomous outcome (as described in Aim 1) will be analysed using the logistic regression to estimate the odds ratios for its association with potential predictors and the confidence intervals.
2. Secondly, disability will be used on its original continuous scale and linear regression will be used. With this analysis the regression coefficients will describe the change in WHODAS per unit change in explanatory variable. This will be supplemented by boxplots for categorical predictors and scatter plots for scale or ordinal predictors.

3. Change with time for dimensions of PICS

The Post-Intensive Care Syndrome is a term that encompasses a multi-dimensional concept of change for survivors of critical illness who have received ICU treatment. These dimensions, or domains include functional, cognitive, and psychological areas.

The table shows these dimensions together with a brief description.

Table two: Dimensions of PICS

Cognitive	
MOCA	The Montreal Cognitive Assessment – Blind (MOCA – blind) is a shortened tool redesigned from the original MOCA test, developed for rapid screening of cognitive issues remotely (or over the phone). The tool generates scores of 0-22. Whilst the score has not been validated in critically ill patients, the suggested cut off value of 18 and above has been proposed to indicate normal cognitive function.
Quality of life	
EuroQol 5D- 5L	This will use the data obtained from the EuroQol 5-dimension – 5 level (EQ-5D-5L) descriptive component and EQ Visual Analogue Scale (EQ VAS). The EQ5D-5L descriptive component including 5 domains, such as mobility, self-care, usual activity, pain/discomfort, and anxiety/depression. Each domain has a score of 1-5 which creates a unique code converted into a health state. The 5-digit health state profile represents the level of reported problems on each of the five dimensions of health (e.g.EQ-5D-5L health state 21143 represents a patient who indicates slight problems on the mobility dimension, no problems on the self-care and usual activities dimensions, severe pain or discomfort and moderate problems on the anxiety/depression dimension). These health states can be converted into a single index value using one of the standard EQ-5D-5L value sets. Index values reflects how good or bad the health state is according to the preferences of the general population of a country/region. The collection of index values for all possible EQ-5D states is called a value set. A value set for New Zealand is available (Sullivan et al., 2020; Sullivan et al., 2021) and will be used to analyse the HRQoL change over time https://www.1000minds.com/health/hrqol/eq-5d A second component of the overall EQ5D5L instrument, is an overall score scaled from 0-100 for the EQ VAS as a measure of overall self-rated health status.
Psychological	
HADS	The Hospital Anxiety and Depression Scale (HADS) subscale for anxiety and depression: For the 14 questions, a four-point Likert scale (range 0–3) gives a

	possible score of 0 (none) to 21 (severe) for each of the two subscales. Subscale scores 7 or less, indicates no symptoms, scores of 8-10 potential or borderline symptoms and scores >11 indicate clinically significant anxiety or depression symptoms.
IES-r	The Impact of Events Scale Revised (IES-r) and associated subscales: The IES-r has a scoring range of 0 to 88. Results consist of a total raw score, and raw scores for three subscales: the Avoidance Scale, Intrusion Scale, and the Hyperarousal Scale.

Data description for the scale variables will use as data descriptors mean and standard deviation (SD), median and 25th to 75th percentiles (as the inter-quartile range), and minimum to maximum values (range) supplemented by boxplots and frequency histograms. These will be shown for each relevant time point. The EuroQoL will also be described as a categorical variable for the frequency of each of the patterns of the response.

The change in mean values with time will be described by plots including a LOESS plot and a mixed linear model (repeated measures analysis) to establish if there is evidence of a linear change in time (from baseline (4-6 weeks), six months, and twelve months) considering repeated measurements on the same individuals.

4. WHODAS in relation to PICS

A standalone tool that can capture and evaluate the multidimensional, overlapping issues of PICS (psychosocial, physical, cognitive) would be extremely useful to both researchers and clinicians. This part of the study aims to broadly explore the utility of the WHODAS as a predictor for functional, cognitive, and psychological issues, measured using the tools described in table two. The aim will be to describe the association between the WHODAS, and the four variables described in table two.

Associations will be described using boxplots where the WHODAS is described by categories and by scatter plots with WHODAS treated as a scale variable. For WHODAS treated as a dichotomous variable, associations will be described by t-tests. For WHODAS treated as a scale variable, associations will be described by correlation coefficients and regression models.

5. Return-to-work

Return-to-work is an important determinant of recovery because of its effects on financial, personal, and social well-being. Hodgson et al, explored return to work as a sub study from their main study of disability and recovery at six months (Hodgson et al., 2018). In that study risk factors and associations with return-to-work were explored using multivariate logistic regression. In this study we are limited by a much smaller sample size.

Return-to-work will be described as numerators and denominators and proportions expressed as percentages at each time-point. Other descriptors of return-to work included: whether the participant was in paid employment before the illness, working part-time of full-time, including a reduction in hours compared to before the illness.

6. Health care utilisation

The count of presentations per participant to the Emergency Department or Hospital will be described in relation to the length of time of observation, for up to a year after discharge.

The association between these counts and the dichotomous measurement of disability based on the WHODAS will be estimated by Poisson regression, with an offset for the time of observation, and described as estimation of relative rates comparing moderate to severe with none to mild.

Survival analysis will be used to describe the mean number of days to first healthcare presentations for mild, moderate-severe groups.

7. Cognitive status and potential predictors

Associations between cognitive function measured by the MOCA-blind and the potential predictors will be visually summarised by scatter plots, a correlation matrix, and estimation of associations by linear regression.

Several predictor variables have been chosen to model the association with cognitive dysfunction after critical illness and are described below. Delirium and duration of mechanical ventilation are directly associated with worse cognitive dysfunction after critical illness and are thus included predictor/explanatory variables (Bassi et al., 2021; Salluh et al., 2015). The remaining variables of age, operative status (especially those undergone cardiopulmonary bypass for cardiac surgery) and higher clinical frailty scores are also associated with cognitive dysfunction in non ICU cohorts but may also be potential confounding variables (Li et al., 2023; Vu & Smith, 2022).

Other variables such as depth of sedation (measured by daily Richmond sedation scores) and increased doses of sedatives, opiates, benzodiazepines, antipsychotics, and steroids may influence cognitive dysfunction through delirium. However, this association remains unclear (Long et al., 2020). Sedation depth (deep sedation), benzodiazepines, steroids, antipsychotics, and doses of analgosedation are typically used ICU treatments. They may also be independently associated with increased severity of illness, longer ICU length of stays and increased mortality. They are also considered management of agitated delirium. There are thus several pathways to explain cognitive dysfunction. These variables may therefore be considered mediator variables.

All these variables will be explored in terms of their relationship with cognitive outcomes measured by the MOCA scores.

Table three: Potential predictors of cognitive function

Variable	Description
Age	At ICU admission, measured in years
Non operative versus operative status (further defined as cardiothoracic, surgical, or vascular)	Classified on admission in the ICU database
Clinical Frailty Scores (CFS)	The CFS (range 1–8) categorises patients as non-frail (1 = very fit; 2 = well; 3 = managing well; 4 = vulnerable) or frail (5 = mildly frail; 6 = moderately frail; 7 = severely frail; 8 = very severely frail
Duration of delirium (hours)	Measured in hours
Duration of mechanical ventilation	Measured in hours
Depth of sedation reported as daily mean Richmond -Agitation – Sedation Scores (RASS) scores. These are reported on a scale from minus 5 to +4 (minus 5 being comatose, to +4 rampant/ dangerous agitation).	All Daily RASS scores during mechanical ventilation will be collected and will be converted into total RASS scores per each level of the scale for the entire ICU stay.
Total drug dosages of Intravenous (IV) and oral sedation, IV analgosedation, antipsychotics, steroids, hypnotics etc).	Daily drug doses of opiates, benzodiazepines, sedation, alpha agonists and IV analgesia, steroids have been collected and will be converted into total doses given across the ICU stay.

Anticipated Limitations

The anticipated limitations for the study are:

1. Lack of power to detect important associations (type II error).
2. Selection bias in relation to the potential population of all survivors of Intensive Care Unit treatment.
3. Inability to truly measure pre-illness disability.
4. Multiple statistical testing leading to type I error inflation.

Type I and Type II errors

Generally, we acknowledge there will be limitations in the statistical analyses and conclusions we can generate in this study. Firstly, this relates to underpowering with a small sample size and the potential for a type I and type II error to occur. During the study planning, we defined a priori an estimated 100 participants data at six months would be needed to adequately describe the proportion and range of disability (and effect size) at six months. We remained reflexive to the recruitment of participants to oversample and recruit as many participants as possible by broadening the inclusion criteria (of duration of mechanical ventilation from 72 hours to 48 hours) and approaching two further centres in New Zealand (Waikato and Christchurch). To ensure an adequate sample size at six months, exhaustive strategies (koha, mailouts, text reminders etc.) by the principle investigator to retain participants and keep loss to follow up rate low has also been employed.

It is acknowledged that the statistical analysis plan for this study is ambitious with multiple statistical testing planned, included variables, end points and time points. This will increase the risk of associated statistically significant p values being found, when infact they are there by chance (and thus type I error). All though the development of the SAP and the study methodology, advice from senior supervisors and senior statistical analysts has been sought. Conclusions from the statistical analysis will consider these errors and any statistically significant findings will be evaluated in the context of study size, reported with confidence intervals and standard deviations (where appropriate) and other measures of accuracy and precision.

Confounders, mediators, and selection bias

It is acknowledged the inclusion of participants from other centres may introduce potential confounders with difference in ICU setting, care, and treatment between centres. However, all centres used in the study are similar in relation to cohort admissions, ICU type (i.e. tertiary, closed units) and adherence to ANZICS care guidelines. Irrespective, there may be hidden confounders that change the outcome for patients depending on the centre they were treated in. Differences in recovery service provision for participants once home may be one of the potential differences related to outcome. The centre the patient was treated in will be captured in this study data set.

Other potential confounders for this study relate to the limited ability to measure pre-existing disability and its impact on post illness disability. It is widely understood the biggest predictor of post illness disability is pre-ICU disability (Denehy & Hough, 2017). However, this is challenging data to collect after critical illness with recall bias being a limitation. Socioeconomic status, and social support systems are also other potential contributory factors that impact on recovery and disability. These are known limitations of survivorship research, difficult areas to collect meaningful valid and responsive data elements. They are however, acknowledged as potentially notable limitations in this study. Other potential confounders such as age, working status before and after critical illness, pre-existing clinical frailty scale and co-morbidities (measured using the Charlson comorbidity index), and COVID illness during recovery are included variables.

Additionally, there are acknowledged potential for bias to be present affecting the sample characteristics especially participant selection bias. There is a large proportion of eligible participants who decline to participate and potential reasons for this may be worse disability, poor cognitive

function or those who are struggling with recovery (exactly the cohort we are trying to capture in this study). Every attempt to contact and recruit every available patient has been used.

References:

Bassi, T. G., Rohrs, E. C., & Reynolds, S. C. (2021). Systematic review of cognitive impairment and brain insult after mechanical ventilation. *Critical Care*, *25*(1), 99. https://doi.org/10.1186/s13054-021-03521-9

Denehy, L., & Hough, C. L. (2017). Critical illness, disability, and the road home. *Intensive Care Med*, *43*(12), 1881-1883. https://doi.org/10.1007/s00134-017-4942-6

Hodgson, C. L., Haines, K. J., Bailey, M., Barrett, J., Bellomo, R., Bucknall, T., Gabbe, B. J., Higgins, A. M., Iwashyna, T. J., Hunt-Smith, J., Murray, L. J., Myles, P. S., Ponsford, J., Pilcher, D., Udy, A. A., Walker, C., Young, M., Jamie Cooper, D. J., & Investigators, I. C.-R. (2018). Predictors of return to work in survivors of critical illness. *J Crit Care*, *48*, 21-25. https://doi.org/10.1016/j.jcrc.2018.08.005

Hodgson, C. L., Udy, A. A., Bailey, M., Barrett, J., Bellomo, R., Bucknall, T., Gabbe, B. J., Higgins, A. M., Iwashyna, T. J., Hunt-Smith, J., Murray, L. J., Myles, P. S., Ponsford, J., Pilcher, D., Walker, C., Young, M., & Cooper, D. J. (2017). The impact of disability in survivors of critical illness. *Intensive Care Med*, *43*(7), 992-1001. https://doi.org/10.1007/s00134-017-4830-0

Jackson, J. C., Pandharipande, P. P., Girard, T. D., Brummel, N. E., Thompson, J. L., Hughes, C. G., Pun, B. T., Vasilevskis, E. E., Morandi, A., Shintani, A. K., Hopkins, R. O., Bernard, G. R., Dittus, R. S., & Ely, E. W. (2014). Depression, post-traumatic stress disorder, and functional disability in survivors of critical illness in the BRAIN-ICU study: a longitudinal cohort study. *The Lancet Respiratory Medicine*, *2*(5), 369-379. https://doi.org/10.1016/s2213-2600(14)70051-7

Li, C., Ge, S., Yin, Y., Tian, C., Mei, Y., & Han, P. (2023). Frailty is associated with worse cognitive functioning in older adults. *Front Psychiatry*, *14*, 1108902. https://doi.org/10.3389/fpsyt.2023.1108902

Long, L., Ren, S., Gong, Y., Zhao, H., He, C., Shen, L., Zhao, H., & Ma, P. (2020). Different depths of sedation versus risk of delirium in adult mechanically ventilated patients: A systematic review and meta-analysis. *PLoS One*, *15*(7), e0236014. https://doi.org/10.1371/journal.pone.0236014

Marra, A., Pandharipande, P. P., Girard, T. D., Patel, M. B., Hughes, C. G., Jackson, J. C., Thompson, J. L., Chandrasekhar, R., Ely, E. W., & Brummel, N. E. (2018). Co-Occurrence of Post-Intensive Care Syndrome Problems Among 406 Survivors of Critical Illness. *Crit Care Med*, *46*(9), 1393-1401. https://doi.org/10.1097/CCM.0000000000003218

Salluh, J. I., Wang, H., Schneider, E. B., Nagaraja, N., Yenokyan, G., Damluji, A., Serafim, R. B., & Stevens, R. D. (2015). Outcome of delirium in critically ill patients: systematic review and meta-analysis. *Bmj*, *350*, h2538. https://doi.org/10.1136/bmj.h2538

Shulman, M. A., Myles, P. S., Chan, M. T., McIlroy, D. R., Wallace, S., & Ponsford, J. (2015). Measurement of disability-free survival after surgery. *Anesthesiology*, *122*(3), 524-536. https://doi.org/10.1097/ALN.0000000000000586

Sullivan, T., Hansen, P., Ombler, F., Derrett, S., & Devlin, N. (2020). A new tool for creating personal and social EQ-5D-5L value sets, including valuing 'dead'. *Soc Sci Med*, *246*, 112707. https://doi.org/10.1016/j.socscimed.2019.112707

Sullivan, T., Turner, R. M., Derrett, S., & Hansen, P. (2021). New Zealand Population Norms for the EQ-5D-5L Constructed From the Personal Value Sets of Participants in a National Survey. *Value Health*, *24*(9), 1308-1318. https://doi.org/10.1016/j.jval.2021.04.1280

Vu, T., & Smith, J. A. (2022). An Update on Postoperative Cognitive Dysfunction Following Cardiac Surgery. *Front Psychiatry*, *13*, 884907. https://doi.org/10.3389/fpsyt.2022.884907