CHERISH analysis plan

Design: pragmatic cluster randomised controlled trial

Primary comparison: The primary comparison will use post-implementation period data on the 4 wards (general medicine, respiratory medicine, orthopaedic and general surgery) implementing Eat Walk Engage compared to 4 wards (matched for hospital: two general medicine, specialty medicine and general surgery) not implementing the program, controlling for age, gender, Charlson comorbidity score, admission ADL status and admission cognitive status (SPMSQ score) and adjusting for clustering by ward.
Secondary comparison 1: will include data from the pre-implementation cohorts as additional controls, controlling for time period. This analysis will provide greater precision to the estimates from the post-intervention comparison. This approach may induce a bias because of unrelated temporal trends in length of stay over time due to other organisational factors, which could be wrongly attributed to the program.
Sensitivity analysis: will use the primary analysis in the post-implementation cohort but include a time since intervention variable to identify whether there is an increasing effect on outcomes over the 6 month post-implementation period as the "dose" of intervention may have been increasing as the model matured. The change over time may be non-linear and therefore we will involve a range of non-linear shapes using the fractional polynomial approach[1]. The best fitting change over time will be estimated using the deviance information criterion (DIC).

Pre-specified subgroup comparisons: will be examined using interaction terms within the primary outcome models and are:

- age under 75 years versus age 75 and older;
- frailty subgroups (less than 0.25 non-frail, 0.25-0.40 mildly frail, 0.40 and above moderately-severely frail) based on a deficit accumulation frailty index;
- the four hospitals.

Primary outcomes:

1. Length of stay (treating unit): described as median time to discharge and median differences between groups; analysed using Bayesian parametric survival analysis
2. Composite outcome of any "hospital associated complication of older people" (HAC-OP) which will consist of:
    a. Hospital-associated delirium (delirium documented either by assessment or chart review, first recorded more than 1 day after admission)
    b. Hospital-associated functional decline (increase in count of ADL requiring human assistance at discharge compared to 2 weeks prior to admission, by patient self-report; or in-hospital death or new residential care)
    c. Hospital-associated incontinence (urinary or faecal incontinence present at discharge which was not present 2 weeks prior to admission, by patient self-report)

[1] P Royston, G Ambler, W Sauerbrei; The use of fractional polynomials to model continuous risk variables in epidemiology., *International Journal of Epidemiology*, Volume 28, Issue 5, 1 October 1999, Pages 964–974

d. Hospital-associated pressure ulcer (identified by patient report or chart documentation, not present at admission assessment)
e. Hospital-associated fall (identified by patient report or chart documentation after admission)

This outcome will be modelled as a dichotomous outcome using logistic regression. Each of the five syndromes will be modelled in the same logistic regression model using a mixed model with a random intercept per participant to adjust for correlated data from the same participant. This regression model will estimate the effect of the intervention on the overall syndrome. In a sensitivity analysis we will add an interaction between the intervention and each syndrome to examine whether the intervention had a stronger effect on some syndromes. The models will control for age, gender, comorbidity score, admission ADL status and admission cognitive status (SPMSQ score) and adjust for clustering by ward.

Secondary outcomes:

1. Individual HAC-OP as defined above
2. Death or discharge to institutional care (new residential care, continuing acute, rehabilitation or convalescent care) versus discharge home
3. 30 day functional recovery (return to baseline ADL and IADL status)
4. 30 day all-cause hospital readmission
5. 30 day all-cause mortality
3. Quality of life (EQOL5D) at 30 days
4. 6 month all-cause hospital readmission
5. 6 month all-cause mortality

A "scrambled" analysis (based on simulated intervention groups) will be undertaken and shared with the investigator group for final refinement of methods before commencing full analysis. This aims to reduce the bias of making changes after the full results are available.

Missing data:

The small amount of missing data for ADL and SPMSQ at baseline will be imputed using Multivariate Imputation by Chained Equations (MICE). [2] This is to ensure that the maximum amount of available data are used and to help avoid selection biases caused by participants with partially missing data (e.g., sicker patients being excluded). The variables used by MICE to impute ADL and SPMSQ will be age, IADL at baseline and Charlson comorbidity index.

We will use logistic regression to examine the missing outcome data and see what variables predict missing using treatment group, ward, age, gender, Charlson comorbidity score, admission ADL status and admission cognitive status (SPMSQ score). If strong associations exist we will use inverse-probability weighting to adjust the primary and secondary outcomes to compensate for the non-random missingness. This will be an additional sensitivity analysis.

[2] S van Buuren, K Groothuis-Oudshoorn (2011). mice: Multivariate Imputation by Chained Equations in R. Journal of Statistical Software, 45(3), 1-67.