MJA
The Medical Journal of Australia

Appendix 3

This appendix was part of the submitted manuscript and has been peer reviewed.
It is posted as supplied by the authors.

Appendix to: Dyer SM, Liu E, Gnanamanickam ES, et al. Clustered domestic residential aged care in Australia: fewer hospitalisations and better quality of life. *Med J Aust* 2018; 208: 433-438. doi: 10.5694/mja17.00861.

Supplement 3 Detailed statistical methods for the multi-level models

The multi-level models we used to analyse this cross-sectional study are two-level multilevel models which involve several level-1 explanatory variables (such as age, gender et al) and several level-2 explanatory variable (such as care models, staff training et al). We used multilevel models to analyse the data because we think the independent and identically distribution (i.i.d) assumption of the sample may not hold for individuals who live in the same aged care facility.

The formulas of the two-level multilevel models for the data analysis are:

$$y_{ij} = \alpha_{0j} + \beta_1 x_{1ij} + \beta_2 x_{2ij} + e_{ij} \tag{1}$$

$$\alpha_{0j} = \gamma_{00} + \gamma_{01} w_{1j} + u_{0j} \tag{2}$$

where y_{ij} stands for level-1 outcome measurement (EQ5D5L utility score) of i^{th} individual in the j^{th} level-2 unit(e.g, the aged care facility);$i = 1, 2, ..., 541$, 541 is the total sample size of this study.$j = 1, 2, ..., 17$, 17 is the number of facilities (level-2 units) for this study.

Equation (1) is the level-1 equation that models the variation of y_{ij}, Noted the level-1 intercept α_{0j} has subscript "j" indicating that level-1 intercept varies across level-2 units (facilities here) i.e for this analysis we only consider random intercepts, all other coefficients are fixed. x_1 and x_2 are level-1 explanatory variables, β_1 and β_2 are corresponding fixed coefficients.e_{ij} represents individual variations within the tactility j

Equation (2) is the level-2 model that model the variation of random coefficients in the level-1 model. For this study, the level-2 equation will only model the α_{0j}, the intercepts of the level-1 model. In this equation γ_{00} represents the grand mean of y_{ij} and u_{0j} is the variations between different level-2 facilities. w_1 is level-2 explanatory variable, for simplicity we only show one level-2 explanatory variable, γ_{01} is the corresponding fixed coefficient.

Now we can substitute equation (2) into equation (1) we get

$$y_{ij} = \gamma_{00} + \gamma_{01} w_{1j} + \alpha_{0j} + \beta_1 x_{1ij} + \beta_2 x_{2ij} + e_{ij} + (u_{0j} + e_{ij}) \tag{3}$$

Equation (3) looks like an ordinary linear regression model,however, the combined model has a composite error structure.

For equation (3), we make the following assumptions:

$$e_{ij} \sim N(0, \sigma^2)$$
$$u_{0j} \sim N(0, \sigma^2_{u_0})$$
$$Cov(e_{ij}, u_{0j}) = 0$$

By matrix notation we express the multilevel model error terms as

$$E \begin{pmatrix} u \\ e \end{pmatrix} = 0$$

$$Var\begin{pmatrix} u \\ e \end{pmatrix} = \begin{pmatrix} G & 0 \\ 0 & R \end{pmatrix}$$

where G is the variance/co-variance matrix for level-2 residual and R is variance/co-variance matrix for level-1 residual.The variance/co-variance matrix can have different structures such as variance components, autoregressive(1), compound symmetry, etc. For this data analysis we used variance components structure which is the default structure for the SAS mixed model procedure.

Both the maximum likelihood method (MLE) and REML (residual/restricted maximum likelihood) can be used to estimate the parameters in the multilevel model when the sample size is big. The default setting for the SAS software is REML which is usually considered unbiased. We used this method to estimate our coefficients and variances.

The followings are the main SAS codes we used to perform the multi-level analysis. The program corresponds to the two-level multilevel models with the random intercept and variance components structure we described above.

```
proc mixed data=EQ5D5L_cost2 method=reml covtest;
class facilityID
ageCat(ref='under 65 years')
gender(ref='Male')
PASCOGScoreNewCat(ref='PAS-Cog 0-4 (no cognitive impairment)')
Cottage_Category_Res(ref='No')
bedsCat(ref='low (<80)')
clsocties3
DirectCareHours(ref='High direct care hours')
StaffTraining(ref='Higher staff training')
regional;
model eq5d5l_Index=
Cottage_Category_Res
regional bedsCat
StaffTraining
DirectCareHours
ageCat
gender
PASCOGScoreNewCat
Barthel_Total_Score_060716
clsocties3
NoCoMorb/ cl solution;
random intercept/subject=facilityID;
lsmeans Cottage_Category_Res/diff cl;
run;
```