Appendix 1 (as supplied by the authors): Survey design

Experimental survey design refers to the way that the alternatives' levels are set and structured into each choice set. A balance needs to be found between including enough attributes and attribute levels to adequately describe alternatives and facilitate trade-offs, and minimising the complexity of the task for respondents. Efficient survey designs are those that optimise the information to be gained from each choice set, without having to present all possible combinations of attributes and levels to respondents.

Statistically efficient designs enable the estimation of model parameters with as low as possible standard errors for a given sample size, or give the same standard errors but at a smaller sample size. The standard errors can be predicted by determining the asymptotic variance-covariance (AVC) matrix or Fisher Information matrix based upon some prior information about the parameter estimates. Such 'prior information' may come from the literature or from a pilot study with the relevant individuals. Efficient designs are measured with a D-error statistic, which is a measure of the design 'inefficiency' and therefore, the lower the D-error, the more efficient the design.[1] S-efficiency, a relatively new measure of efficiency developed by Bliemer and Rose[2] calculates the minimum DCE sample size required to achieve statistical significance (t-ratio > 1.96, p<0.05) for all attributes or β parameters in a model.

A sample size of 100 patients was based on an *s-estimate,* with a d-error of 0.0885. The minimum number of choice sets is a function of the number of parameters (attributes) and the number of alternatives. 12 choice sets was the minimum needed for an efficient design in which we could achieve an acceptable d-error and was a feasible number for patients with chronic kidney disease without creating undue respondent burden. Conditional statements that linked two or more levels in an alternative were created to ensure plausibility of alternatives. For example, when the 'treatment hours' level in the Dialysis A alternative was 10 hours, the 'time of day' level was fixed to overnight, simulating a plausible alternative of automated peritoneal dialysis.

Analysis

Probabilistic choice models are characterised by the following equation:

$U = V + \varepsilon$,

where U = utility (or satisfaction), V = the observed component of choice between alternatives, and ε = the unobserved component or error term. In the mixed logit model, one or more of the parameter estimates (i.e. random parameters) are represented as:

$\beta nk = \beta k + \eta k\ znsjk$,

where ***βk*** is the mean marginal utility in the sampled population and **η** is the deviation of the mean marginal utility held by patient ***n*** for characteristic ***k*** belonging to alternative ***j*** in choice set ***s***. ***znsjk*** represents an underlying distribution such as ***znsjk*** ~N(0,1). In matrix notation the utility function of the mixed logit model is represented like this:[1]

$$U_{nsj} = \sum_{k}^{K=1} \beta_{nk} x_{nsjk} + \varepsilon_{nsj}$$

We included all treatment characteristics and relevant socio-demographic variables that best explained patient choice.

Therefore our utility function was:
U(dialysis) = $\beta_{survival}$*survival + β_{visits} *visits + β_{travel} *travel + β_{hours} *hours + $\beta_{TOD(day/evening)}$ *time of day(day/evening)+ $\beta_{TOD(night)}$*time of day(night)+ $\beta_{transport}$ *transport + $\beta_{flexibility}$ *flexibility + $\beta_{agegroup}$*agegroup + β_{sex} *sex + $\beta_{distance}$ *distance +ε,

(with all attributes treated as normally distributed random parameters, and socio-demographic variables for age group, sex and distance as fixed parameters).

All model parameters were initially specified as random. When random parameters were found to have an insignificant standard deviation for their distribution, they were included as a fixed parameter and the model re-estimated. Model fit statistics, including Log likelihood, were assessed after each re-specification, and fixed parameters that were non-significant were dropped if their removal did not significantly compromise model fit. The model was estimated using 1,000 Halton draws.

Note: The McFadden pseudo-R^2 (or ρ^2) statistic is a model fit statistic for non-linear choice models calculated using the following equation:[3]

$$\rho^2 = 1 - \frac{\text{Log likelihood estimated model}}{\text{Log likelihood base model}}$$

The McFadden ρ^2 is not exactly analogous to the R^2 statistic of linear regression models; however a direct empirical relationship may exist (Domencich and Mc Fadden 1975). We know a higher ρ^2 is preferable to a lower one, but how it translates to an equivalence of R^2 is an empirical question related to each specific data set.

Estimation of trade-offs that individuals made between attributes (also known as marginal rates of substitution) were calculated by dividing the mean parameter estimate of one attribute by the mean parameter estimate of a second attribute to obtain its ratio. 95% confidence intervals were calculated to present the measure of precision around benefit/harm trade-offs.

References

1. Rose JM, Bliemer MCJ. Constructing Efficient Stated Choice Experimental Designs *Transport Reviews* 2009;29(5):587-617.
2. Bliemer MCJ, Rose JM. Construction of Experimental Designs for Mixed Logit Models Allowing for Correlation Across Choice Observations. *Transportation Research Part B: Methodological* 2010;46(3):720-34.
3. Hensher DA, Rose JM, Greene WH. *Applied choice analysis: a Primer*. Cambridge, UK: Cambridge University Press, 2005.
4. Domencich T, Mc Fadden D. *Urban travel demand: a behavioral analysis.* North-Holland Publishing Co., Amsterdam. 1975.

Appendix to: Morton RL, Snelling P, Webster AC, et al. Factors influencing patient choice of dialysis versus conservative care to treat end-stage kidney disease. *CMAJ* 2012. DOI:10.1503/cmaj111355.