APPENDIX 2: ANALYSIS SETS AND STATISTICAL METHODS

Analysis sets

This refer to the sets of cases whose data are to be included in the analysis.
Three analysis sets are defined:
1. Dialysis patients notified between 1980 and 2003
This analysis set consists of patients commencing dialysis between 1980 and 2002. This analysis set is used for the analysis in Chapter 1 and 2.

2. Dialysis patients between 1993 and 2002
Since 1993, the NRR conducted an annual survey on all dialysis patients to collect data on dialysis and drug treatments, clinical and laboratory measurements. All available data were used to describe the trends in these characteristics.
However, in the early years, these data collected from annual survey were relatively incomplete. Hence, for survival analysis in relation to these characteristics, we used only data from 1997 onwards when the data were more complete. Remaining missing data in this analysis set was imputed using first available observation carried backward or last observation carried forward.

3. Rehabilitation outcomes
Analysis is confined to the relevant population. Hence we exclude the following groups.
(i) Age less than or equal to 21 years
(ii) Age more than or equal to 55 years
(iii) Homemaker
(iv) Full time student
(v) Retired

Statistical methods

Population treatment rates (new treatment or prevalence rates)

Treatment rate is calculated by the ratio of the count of number of new patients or prevalent patients in a given year to the mid-year population of Malaysia in that year, and expressed in per million-population.
Results on distribution of treatment rates by state are also expressed in per million-population since states obviously vary in their population sizes.
Classification of level of provision in a state is based on dialysis treatment rate over period 2000-2002. High provision states are defined as those with rate > 100pmp, mid provision states 50-100pmp and low provision states <50pmp.

Death rate calculation

Annual death rates were calculated by dividing the number of deaths in a year by the estimated mid-year patient population.

Odds ratio

The odds of an event is the probability of having the event divided by the probability of not having it.
The odds ratio is used for comparing the odds of 2 groups. If the odds in group 1 is O1 and group 2 is O2, then odds ratio is O1/O2. Thus the odds ratio expresses the relative probability that an event will occur when 2 groups are compared.
With multiple factors, logistic regression model was used to estimate the independent effect of each factor, expressed as odds ratio, on the event of interest.

Cumulative odds ratio

For QOL outcome, which is measured on an ordinal scale, the cumulative odd ratios for a factor that affected the outcome expresses the relative cumulative probability for the QOL score. This is best explained by an example. The cumulative OR for QOL score for female dialysis patients compare to males is 0.77. This means the odds for higher QOL score are lower for female than male patients. In other words, the cumulative distribution for the QOL score for female patients is shifted to the left of male patients.
The cumulative odds ratio associated with a factor of interest is estimated using the proportional odds model. In this model, the cumulative probabilities for the ordinal dependent variable (QOL score), after suitable transformation (logit transform), is modelled as a linear function of all the factors of interest (covariates).

Survival analysis

The unadjusted survival probabilities (with 95% confidence intrevals) were calculated using the Kaplan-Meier method, in which the probability of surviving more than a given time can be estimated for members of a cohort of patients without accounting for the characteristics of the members of that cohort. Where centres are small or the survival probabilities are greater than 90%, the confidence intervals are only approximate.

In order to estimate the difference in survival of different subgroups of patients within the cohort, a proportional hazards model (Cox) was used where appropriate. The results from Cox model are interpreted using a hazard ratio. Adjusted survival probabilities are with age, gender, primary diagnosis and time on RRT used as adjusting risk factors. For diabetics compared with non-diabetics, for example, the hazard ratio is the ratio of the estimated hazards for diabetics relative to non-diabetics, where the hazard is the risk of dying at time t given that the individual has survival until this time. The underlying assumption of a proportional hazards model is that the ratio remains constant throughout the period under consideration.

Analysis of trend of intermediate results

For summarizing intermediate results like continuous laboratory data, we have calculated summary statistics like mean, standard deviation, median, lower quartile, upper quarter and interquartile range (IQR). For QOL and rehabilitation outcomes of dialysis patients, cumulative distribution plot shows a listing of the sample values of a variable on the X axis and the proportion of the observations less than or greater than each value on the Y axis. An accompanying table gives the Median (50% of values are above or below it), upper quartile (UQ, 25% of values above and 75% below it), lower quartile (LQ, 75% of values above and 25% below it) and other percentiles. The table also shows percent of observations above or below a target value, or within an interval of values; the target value or interval obviously vary with the type of laboratory data. For example, interval of values for prescribed KT/V is <1, 1-<1.2, 1.2-<1.4, 1.4-<1.6 and ≥1.6 and that for haemoglobin is <10, 10-≤12 and >12 g/l. The choice of target value is guided by published clinical practice guidelines, for example, the DOQI guideline; or otherwise they represent consensus of the local dialysis community.