APPENDIX - II

ANALYSIS SETS, STATISTICAL METHODS AND DEFINITIONS

APPENDIX II: ANALYSIS SETS, STATISTICAL METHODS AND DEFINITIONS

Analysis sets

This refers to the sets of cases whose data are to be included in the analysis.

Seven analysis sets were defined:

1. Dialysis patients notification between 2004 and 2013

This analysis set consists of patients commencing dialysis between 2004 and 2013. This analysis set was used for the analysis in Chapter 1, 2 and 3.

Patients who were less than 20 years old at the start of dialysis between 2004 and 2013 were used for the analysis in Chapter 5.

Since 1993, the MDTR conducted an annual survey on all dialysis patients to collect data on dialysis and drug treatment, clinical and laboratory measurements. All available data were used to describe the trends in these characteristics. For this analysis in relation to these characteristics, only data from 2004 onwards were used. Remaining missing data in this analysis set was imputed using first available observation carried backward or last observation carried forward. This analysis set was used for the analysis in Chapters 6 to 12. However, the generated variable that has been imputed is prescribed Kt/V for HD patients. Prescribed Kt/V was generated using the formula below:

Kt/V = kdx x hd_time x 60/(0.58 x post weight x 1000)
where
kdx =[1 – exp(-ex)] x HD flow rate x 500/[500 – HD flow rate x exp(-ex)]
and
ex = (500 – HD flow rate) X ka/(500 x HD flow rate).

This variable is considered in Chapter 11.

2. New Dialysis Patients

The number of new dialysis patients was based on the first dialysis treatment of the patients. Patients who convert from one dialysis modality to another (from HD to PD or vice versa) are not counted as new patients. If transplant is the 1st modality and patient's kidney transplant failed and he received dialysis, then for RRT count, the patients will be counted twice. However, if the patients receive transplant between the dialysis, then the dialysis after transplant will be counted if the transplant last for more than 90 days while if it is less than or equal to 90 days, then the dialysis after the transplant will not be counted. This analysis set definition was used in chapters 1, 2 and 5.

3. Rehabilitation outcomes

Analysis is confined to the living patients as at 31st December 2013. Hence we exclude the following groups.

- Age less than or equal to 21 years
- Age more than or equal to 55 years
- Homemaker
- Full time student
- Retired

This analysis set was used for the analysis in Chapter 4.

4. Centre Survey data

Section 2.2 in the report was based on annual centre survey data between 2004 to 2013 rather than individual patient data reported to the Registry.

5. Peritonitis data

Analysis was confined to chronic PD patients who were on peritoneal dialysis from 2004 to 2013. This analysis set was used for the analysis in Section 12.4.

6. Renal transplant data

This analysis set was confined to patients who had undergone renal transplantation from 2004 to 2013. This analysis set was used for the analysis in Chapter 13.

7. Sero-conversion patients

The number of sero-conversion patients was based on the first dialysis treatment of patient with sero-negative from 2004 to 2013. Patients with sero-positive at entry of dialysis treatment were excluded from analysis cohort. The analysis cohort also excluded patients who convert from one dialysis modality to another (from HD to PD or vice versa). Patients with sero-negative at the beginning and subsequently converted to positive will be considered as sero-converted patients and duration of conversion was calculated based on year of conversion minus year of entry. This analysis set was used in chapter 10.

8. Diabetes Mellitus

Patients are considered to have diabetes mellitus (DM) as the cause of ESRD if the primary cause of ESRD is notified as DM; or as unknown but the comorbid is DM. This is applicable to chapter 2, 3 and 13.

Statistical methods

1. 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 in the state since states obviously vary in their population sizes.

2. Adjusted Mortality of dialysis patients

Cox proportional hazards model was considered for mortality of the patients adjusted with demographic and laboratory variables. This analysis was used in Chapter 3 and 12.

3. 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 quartile and the cumulative frequency distribution graph is plotted by year. 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). Other percentiles can be read directly off the cumulative distribution plot. The table also shows percent of observations above or below a target value, or with 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 >1.3 and that for haemoglobin is <10, 10-11 and >11 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. This analysis was used in Chapter 4, 6, 7, 8, 9, 11 & 12.

4. Centre survey data

In contrast to other results reported in this report, Section 2.2 in chapter 2 was based on centre survey data rather than individual patient data reported to the Registry. This is to provide up-to-date information on patient and centre census in the country and thus overcome the inevitable time lag between processing individual patient data and subsequent reporting of results. The survey was conducted in the month of December 2013. Centre response rate to survey was almost 100%. Standard error estimates are not reported because no sample was taken. Results on distribution by state are also expressed in per million state-population since states obviously vary in their population sizes. State population data are based on 2012 census projection. It is very difficult to estimate the amount of cross boundary patient flow; this source of error is therefore not accounted for in computing states estimates. However, we minimize the bias by combining states (eg Kedah and Perlis) based on geographical considerations. HD treatment capacity is derived by assuming on average patients underwent 3 HD sessions per week and a centre can maximally operate 2.5 shifts per day. A single HD machine can therefore support 5 patients' treatment. Obviously HD treatment capacity is calculated only for centre HD. The ratio of the number of centre HD capacity to number of centre HD patient is a useful measure of utilization of available capacity.

5. Centre variation

To compare the variation of the intermediate results between centres, graphs describing intermediate results in each centre are presented. The 95% confidence intervals have been calculated using the normal approximation of the Poisson to show the variation of proportion in centres. Lower quartile and upper quartile are instead plotted in comparison of variation in median among centres. An accompanying table gives the summary statistics like minimum, 5th percentile, lower quartile, median, upper quartile, 95th percentile and maximum value among centres over year.

Centres with intermediate results for <10 patients were combined into one composite centre. This analytical method was used in Chapter 6, 7, 8, 9, 10 11 & 12.

6. Death rate calculation

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

7. Incidence rate ratio

The incidence rate is determined by dividing the number of new cases of a disease or condition in a specific population over a given period of time by the total population. Therefore incidence rate ratio is the comparison of two groups in terms of incidence rate. Poisson regression model was considered to estimate the independent effect of each factor, expressed as incidence rate ratio. An incidence rate ratio of 3 means that group 2 have the rate 3 times higher than group 1 when group 1 is the reference group.

8. 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 1 and group 2 is 2, then odds ratio is 1/2. Thus the odds ratio expresses the relative probability that an event will occur when 2 groups are compared.

With multiple factors such as dialysis center, age, sex, modality, albumin, hemoglobin, calcium, cardiovascular and cholesterol, logistic regression model was used to estimate the independent effect of each factor, expressed as odds ratio, on the event of interest and the variation is odds ratio. This method was used in Chapter 3.

9. Standardized mortality rate

The cohort considered for this analysis was patients who were on dialysis in 2012 and new patients in 2012 by modality. SMR is a ratio between the observed number of death with the expected, based on the age group, diabetic, serum album group, diastolic blood pressure group and hemoglobin group rates in a standard population and the age group, diabetic, serum album group, diastolic blood pressure group and hemoglobin group distribution of the study population. If the ratio observed : expected death is greater the 1.0, we conclude that there is “excess death” in the study population. SMR was generated using the following formula:

SMR = observed death / expected death

10. Risk adjusted mortality rate (RAMR)

When the mortality rate are risk adjusted, the information becomes more comparable among the hospitals because the data is adjusted to take into account variations in patients’ severity of renal disease and their risk of mortality. RAMR was generated using the following formula:

RAMR = SMR x AvMR where AvMR is the average of the overall observed mortality rate

11. Risk ratio

Risk ratio is the relative measure of the difference in risk between the exposed and unexposed populations in a cohort study. The relative risk is defined as the rate of disease among the exposed divided by the rate of the disease among the unexposed. A relative risk of 2, means that the exposed group has twice the disease risk as the unexposed group.

12. Survival analysis

The unadjusted survival probabilities 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.

In order to estimate the difference in survival of different subgroups of patients within the cohort, a stratified proportional hazards model (Cox) was used where appropriate. The results from Cox model are interpreted using a hazard ratio. Adjusted survival probabilities are adjusted for age, gender, primary diagnosis and time on RRT. 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 survived until this time. The underlying assumption of a proportional hazards model is that the ratio remains constant throughout the period under consideration.

Technique failure is defined as occurrence of death or transfer to another modality of dialysis. Similarly, graft failure is defined as occurrence of death or returned to dialysis.

13. Patient survival was considered in two ways:

Survival censored for change of modality based on first modality. Duration survival for patients will be calculated from the date commencing the first modality till first modality outcome. Hence duration after the change modality or transplant will not be considered. Death occurring during the first modality will be considered in the analysis since patients will be censored for change of modality before death.

Survival not censored for change of modality based on first modality. Duration survival for patients will be calculated from the date commencing the first modality till 31 Dec 2013 for patients who were still on RRT. For patients who died, duration of survival will be calculated from date commencing the first modality till date of final outcome which is death. All death outcomes whether occurring during first modality or after change in modality will be considered for this analysis.

14. Survival of incident patients by centre

1-year survival

The cohort considered for this analysis was considered from 2004-2012. Many patients commencing dialysis in 2012 would still not have completed one year.

5-year survival

The cohort considered for this analysis was considered from 2004-2008. This is due to those commence from 2008 onwards still not able to have 5 year survivals analysis.

15. Funnel plot

This analysis was confined to new dialysis patients from year 2004-2012. The figure is included to assess whether survival probability adjusted to age and diabetes of each centre is likely to be different from the national average. This plot was used in Chapter 3.

16. Peritonitis rate

The occurrence of peritonitis is expressed as number of episode per patient-month of observation; peritonitis rate in short. Relapse peritonitis is defined as peritonitis caused by the same organism occurring within 6 weeks of diagnosis of previous peritonitis.

17. Cumulative Risk

Cumulative risk of sero-conversion is the cumulative incidence rate of patient being converted from sero-negative to sero-positive over a period of time. It was calculated by the number of cases during a period divided by number of subjects at risk i.e. sero-nega tive patients at the beginning of time. This analysis was used in chapter 10.