Appendix I

ANALYSIS CRITERIA AND STATISTICAL METHODOLOGY

ANALYSIS SETS

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

1. All biopsies from 2005-2012
The analysis set in Chapter 1 includes all patients who underwent native and graft kidney biopsies from 1st January 2005 to 31st December 2012.

The analysis set consists of biopsy number where it is defined as number of episodes of distribution of renal biopsy in patients. Biopsy number was taken for the highest episode only for each patient. This analysis set was used for analysis in Chapter 1.

2. All native renal biopsy, 2005-2012
The analysis set in Chapter 1 includes patients who underwent native renal biopsy from 1st January 2005 to 31st December 2012.

3. Primary glomerulonephritis patients
Patients described in Chapter 2 are those whose age are $\geq$ 15 years old at the time of biopsy with primary glomerulonephritis on renal biopsies performed in 2005-2012.

4. Lupus nephritis patients
Patients described in Chapter 3 are those whose age are are $\geq$ 15 years old at the time of biopsy, were ticked YES on SLE and were diagnosed lupus nephritis on renal biopsies from 1st January 2005 to 31st December 2012.

5. Paediatric native renal biopsy, 2005-2012
Patients described in Chapter 4 are those whose age are < 15 years old at the time of biopsy were performed from 1st January 2005 to 31st December 2012.

6. Renal Allograft biopsy
The analysis set is confined to all graft biopsies from 1st January 2005 to 31st December 2012.

2. STATISTICAL METHOD

Patient's characteristics
These sections included the patient's age at biopsy, gender and ethnic group in every chapter of this report. In statistics, imputation is the substitution of some value for a missing data point. Therefore, missing of patient's age has been considered to replace with technique imputation for chapter 1, 2 and 3. Then we used the imputation values for the analysis set. For ethnic group other than Malay, Chinese or Indian will be classified as Others. Patient's centre state was used to describe the reported renal biopsy by state and is used for the analysis in chapter 1.

Clinical presentation
These sections described clinical presentation at the time of biopsy. Apart from clinical presentation, chapter 2 and chapter 3 also report on prevalence of hypertension and degree of renal function.

Biopsy data
The biopsy data and outcome on complications are reported in chapter 4 and chapter 5.

Laboratory data
Few variables in this dataset were missing. Those variables are GFR, urine protein, 24hrs urine protein and urine RBC. Therefore, imputation was done to these variables.

Histological diagnosis
In this section, analysis was confined to available data only and no imputation was done.

Centre survey data
Centre survey data were used to determine the ascertainment rate for each SDP's and it is reported in chapter 1.

Hazard ratio
The hazard ratio in survival analysis is the effect of an explanatory variable on the hazard or risk of an event. The hazard ratio compares groups differing in risk factors. If the hazard ratio is 2.0, then the rate failure in one group is twice the rate in other group. This was used for analysis in Chapter 4.

Risk ratio
The risk ratio is the risk of an event (diagnosis) relative to exposure. The risk ratio takes on values between zero and infinity. One is the neutral value and means that there is no difference between the groups compared. This was used for analysis in Chapter 4.

Survival analysis
The unadjusted survival probabilities were calculated using the Kaplan-Meier method. Survival analysis involves the modelling of time to event data, in this context, death is considered an event.
Survival rate is a part of survival analysis, indicating the percentage of people in this group who are alive for a given period of time after diagnosis with the minimal change disease and focal segmental glomerulosclerosis. This was used for analysis in Chapter 4.

Renal allograft biopsy rates
Renal transplant biopsy rate is calculated by the ratio of the count of number of patients in a given year (according to its age group) to the mid-year population of Malaysia in that year, and expressed as in per million populations. This was used for analysis in Chapter 5.

American Rheumatological Association (ARA) Criteria
An ARA criterion is defined as YES on SLE clinical presentation and SLE lab data. Eleven criteria have been considered for ARA. Nine criteria are from SLE clinical presentation with presentation of malar rash, discoid rash, photosensitivity, oral ulcers, arthritis, serositis, renal, cerebral, and hematological. However, the other two are from SLE lab data where patient should have positive on ANF, and at least 1 is Positive on dsDNA, ssDNA, Anti-cardiolipin antibody, Anti-phospholipid antibody, Histone, Nucleo, Ro, La or Sm. This was used for analysis in Chapter 3.

Extra renal involvement criteria
Patients who have at least one of the followings: malar rash, discoid rash, photosensitivity or oral ulcers will be grouped as Muco-cutaneous for other organ involvement criteria. This was used for analysis in Chapter 3.

Density of histogram
Density scales the height of the bars so that the sum of their areas equals 1. The density scale is calculated by the probability of the patients in the interval that concerned and divides with that interval. This figure was considered in Chapter 2.