

Appendix I

DATA MANAGEMENT

APPENDIX 1: DATA MANAGEMENT**Introduction**

Data integrity of a register begins from the data source, data collection tools, data verification and data entry process. Registry data is never as perfect as clinical trial data. Caution should be used when interpreting the results.

Data source

The initial phase of the data collected in the Malaysian Dialysis and Transplant Registry (MDTR) covered all Renal Replacement Therapy (RRT) patients in the Ministry of Health program since its inception in the early 1970s. The Register subsequently received the data from other sectors of RRT providers like the private, non-government organization (NGO), armed forces and the universities.

MDTR continues to actively ascertain new RRT centres in the country. The mechanism of ascertainment is through feedback from the dialysis related companies, current Source Data Provider (SDP) and public propagandas. This will gradually and eventually result in a complete RRT centre database. The identified RRT centre is invited to participate in data collection.

Participation in the MDTR which was entirely voluntary prior to 2006 is now made compulsory by the Private Health Care Facilities and Services Act 1998 and its Regulations 2006 which was implemented on 1st May 2006. This however only applies to private and NGO centres and data submission from centres managed by the Ministry of Health, Ministry of Defence or the Universities is still voluntary. RRT centres which have expressed interest in participating will be recruited as SDP.

In 2016, there are 33 new HD, 4 new PD and 2 new Tx follow-up centres. Eight HD centre had ceased operation. Centre contributed data are shown below:

Table I: Data submission, 2016

	Known centres		Contributed data*		Contributed annual return	
	n	n	n	%	n	%
Haemodialysis	743	716	716	96.2	675	90.8
Chronic PD	45	44	44	97.8	37	82.2
Transplant	49	40	40	81.6	25	51.0
All modality	837	799	799	95.5	737	88.1

* data contributed – patient notification ± annual return forms

Table II: Participation in annual treatment return, 2016

	CRF received	Est. number	Submission rate (%)
HD patient data submission	38073	42075	90.5
HD centre data submission more than 80%	532#	716	74.3
PD patient data submission	4701	5057	93.0
PD centre data submission more than 80%	32*	44	72.7
Tx patient data submission	1717	1961	87.6
Tx centre data submission more than 80%	35	55	63.6

177 HD centres has 100% submission of patient's data

* 11 PD centres has 100% submission of patient's data

~ 26 Tx centres 100%

Data collection

MDTR is a paper base data submission. The case reporting forms are designed to facilitate the data transcription and the information required are readily available in the patient's case note. All the SDPs are provided with instructions on data collection and submission to the Register. The standard data collection forms are colour coded by modality and case report form (CRF) types. The notification forms are submitted periodically or whenever there is an incident. Annual return forms for the assessment year should reach the NRR coordinating office not later than January the following year. The CRFs are:

- Patient notification form
- Outcome notification form
- PD Infection
- HD annual return form
- PD annual return form
- Transplant annual return form
- Work related rehabilitation and quality of life assessment form – annual assessment

MDTR collects patients' demographic details, clinical data, dialysis treatment data, transplant data, peritonitis data and outcome data. MDTR holds individual patient's identifiable data that allow complete follow-up despite patient transfers from one centre to another or change of modality which are especially common among the RRT patients. These patients are monitored and tracked through from the time they were registered until their death. For those patients who were lost to follow-up, MDTR will verify their final outcome with the National Vital Registration System. Patient profiles are submitted to the Register throughout the year. The identity of patients in the database is not released publicly or in the registry reports.

Centre-specific reports are generated and forwarded to SDP on a quarterly basis. This has generated increased feedback from SDP and improved the patient ascertainment rate and the accuracy of the data transmittal in the registry.

MDTR also conducts an annual centre survey on the staffing and facility profile. The survey questionnaire provides summary information about the number of patients on various treatments. This acts as the basis to calculate the patient ascertainment rate.

Database System

The Register initial database was created in DBASE IV in a single computer environment. It was then upgraded to Microsoft Access as a client server application. Currently the NRR data system is a Pentium Xeon 2.33GHz with dual processors, with a total of 8GB RAM memory and 800GB of RAID-5 (Redundant Array of Independent Disks, level 5). In view of high volume of data accumulated throughout these years, capacity ability, performance and security issues of Microsoft Access, it was subsequently migrated to Microsoft SQL Server in the year 2004.

Data management personnel

The data management personnel in the Register office are trained base on the standard operating procedures (SOP). The data entry process is also designed to enhance data quality. Quality assurance procedures are in place at all stages to ensure the quality of data.

Visual review, Data entry and de-duplication verification, Data Editing

On receiving the case report form (CRF) submitted by SDP, visual review is performed to check for obvious error or missing data in the compulsory fields. Data entry will not be performed if a critical variable on the CRF is missing or ambiguous. The CRF is returned to the SDP for verification.

After passing the duplicate check, the data is then entered and coded where required. Edit checks are performed against pre-specified validation rules to detect missing values, out of range values or inconsistent values. Any data discrepancy found is verified against the source CRF and resolved within the Register office where possible. Otherwise the specific data query report will be generated and forwarded to the SDP to clarify and resolve the data discrepancy.

Data coding, data cleaning / data analysis

Most of the data fields have auto data coding. Those data in text fields will be manually coded by the Register manager. A final edit check run is performed to ensure that data is clean. All queries are resolved before dataset is locked and exported to the statistician for analysis

Limitation:

NRR data submission is still paper base. The majority of the RRT centres do not have electronic patient information system. Computer literacy among staff is still low.

The data submission to the Register is still mainly on voluntary basis using the standard data collection forms. Some SDP choose not to participate in data collection on the patient treatment data for various reasons.

Data release and publication policy

One of the primary objectives of the Registry is to make data available to the renal community. There are published data in the registry's annual report in the website: <http://www.msn.org.my>. This report is copyrighted. However it may be freely reproduced without the permission of the National Renal Registry. Acknowledgment would be appreciated. Suggested citation is Wong HS and Goh BL (Eds). Twenty Forth Report of the Malaysian Dialysis and Transplant Registry 2016, Kuala Lumpur 2018

A distinction is made between use of NRR results (as presented in NRR published report) and use of NRR data in a publication. The former is ordinary citation of published work. NRR, of course encourages such citation whether in the form of presentation or other write-ups. The latter constitutes original research publication. NRR position is as follows:

The NRR does not envisage independent individual publication based entirely on NRR published results, without further analyses or additional data collection.

NRR however agrees that investigator shall have the right to publish any information or material arising in part out of NRR work. In other words, there must be additional original contribution by the investigator in the work intended for publication.

NRR encourages the use of its data for research purpose. Any proposed publication or presentation (e.g. manuscript, abstract or poster) for submission to journal or scientific meeting that is based in part or entirely on NRR data should be sent to the NRR prior to submission. NRR will undertake to comment on such documents within 4 weeks. Acknowledgement of the source of the data would also be appreciated.

Any formal publication of a research based in part or entirely on NRR data in which the input of NRR exceeded that of conventional data management and provision will be considered as a joint publication by investigator and the appropriate NRR personnel.

Any party who wish to request data for a specific purpose that requires computer-run should make such requests in writing (by e-mail, fax, or classic mail) accompanied by a Data Release Application Form and signed Data Release Agreement Form. Such request will require approval by the Advisory Board before the data can be released.