



## A Guide to Registry Design, Anil Mehta FRCP FRCPC University of Dundee, Scotland.

This short introduction describes the essential principles underpinning the design of a registry in order to define the natural history of any inherited disease. The guide is based on two decades of experience in developing registries in Scotland, the UK and Europe all focussed on a 'common rare inherited disease', cystic fibrosis (CF). There are three critical design issues to consider before embarking on any registry project.

### Critical Issue 1 – Principles of design

- The software has to be clinically useful
- The software should be designed so that audit falls out of the design automatically
- The central and essential data being used for clinical care should feed the audit trail
- Research whether of a cross sectional or longitudinal nature should be facilitated by the design

### Critical Issue 2 – Keeping to aims that are easy to deliver

- Start very small
- Gather data that is easy (alive, dead, male, female, type of disease)
- Rank the variables to be collected in order of importance
- Make sure that they are less than 10 in number, smaller will do
- The rankings should be based on unambiguous definitions e.g. weight is insufficient and weight without external clothing is a better definition
- This section requires a definitions group of experienced experts to be set up chaired by an expert in data project management, not a doctor

### Critical Issue 3 – Coverage

- Buy in from centres should be agreed up front, buy in from the specialist patient organisations should be agreed up front
- Geography should be considered in relation to coverage; how dispersed is the disease?
- Social class deprivation and geography expertise should be considered
- Contacts who are meticulous in their attention to detail need to be sought
- Government should be informed
- The scope of the geography needs to be defined whether at local, national, European or world level

- Ethics needs to be informed

These overarching principles require the assembly of a multidisciplinary team where the clinician or patient group require the advice of a trained project manager (not a programmer!). Such individuals are expensive because they often began as programmers, migrated to systems designer roles and then became managers in their own right. Experience of client management is key and to be affordable, a part time position is often best suited to this role with an increment in time spent as the project develops.

## Other Issues to consider

### Mission Creep

- Once the project starts and timelines are agreed, no further changes are permitted to the software until the deliverables agreed above are met
- Essential changes need to be project planned

### Data Protection

- The laws at national level and European level may be different
- Patient approval and consent should follow an agreed template
- Patient information sheets must accompany the approval in order to obtain informed consent
- The ranking of the data to be collected must be such that the anonymity of the patient is preserved (see types of data below)
- The critical role of a trusted third party in this process cannot be over emphasised who is the point of contact for all parties inside and outside the project and stands alone, unimpeded in data access and is the sole person with access to all the data.

### Data Collection

- The software should be such that its shelf life should be thought of at the beginning, the relationship between paper records and web based or local computer based data collection approaches should be evaluated from the outset

### Data Verification

- Do not try and aim for more than 50 to 70% coverage in the first instance; less will do
- The meticulous nature of the verification requires a timely return of data to the supplier of data within 2 weeks of receipt of the data otherwise the memory fades and errors creep in
- The feedback should be immediately useful reflect audit and enable research capability

## Management and Control

- A steering committee should be elected whose job it is to approve any changes to the software and any additional items to be collected, they should be named on the data protection forms and the consent forms
- The data controller needs to be identified
- The project manager and the project managers brief needs to be specified
- A helpdesk with swift feedback for problems must be established

## Funding

- It is essential to have a minimum 3 years funding at the beginning but preferably a 5 year funding cycle should be established

## Uses

- The function of such software should enable outcomes to be easily measured if outcome parameters are unknown then surrogates for those outcomes must be established at the outset
- The uses must be clearly specified as standards of care if these are available
- If these are not available then they must be established

## Technical Issues on the data to be collected

The principles of design having been explained, there are a few technical issues that may not be very apparent

- Each data field has a number of attributes
  - The first attribute is that the data field is either quantitative,
  - semi quantitative i.e. has a scale associated with it
  - or qualitative
- Each of these data field types also has a second attribute
  - The second attribute is whether the data is public
  - or private to the patient
  - or semi private
- The principle here is that if the data can be used to identify the patient then this is private data, if the data cannot easily be used to identify the patient this is public data, if the data can with a combination of other pieces of information be used to identify the patient, but this is unlikely this is semi private data

In summary, by defining the uses of the software at the beginning, it is possible by thinking about the type of data its attributes and its function to create a rank order of the essential data to be collected in the initial list of ten items.

## Further information can be found in:

[www.cystic-fibrosis.org.uk](http://www.cystic-fibrosis.org.uk)

[www.eurocarecf.eu](http://www.eurocarecf.eu)

**Mehta A** The How (and Why) of Disease Registers *Early Human Development*  
Sept 2010 (Epub) and references therein.