Written by Louise Marston and Rumana Omar
Last Updated on 02/08/2013 after the Trial Steering Group

Introduction
The Rehabilitation Effectiveness for Activities for Life (REAL) study is a multicentre national (England only) programme of research into mental health rehabilitation units. The project has four phases. This analysis plan only concerns Phase 4; a cohort study which includes units which were found to be performing well in Phase 1 of this study. This analysis plan does not include information on the economic analysis.

Aims
To investigate service user factors and interactions between service user and service factors that are associated with clinical and service outcomes and costs for service users.

Objectives
- To determine whether service quality is associated with service users' social function
- To investigate whether service quality is associated with successful discharge
- To examine whether service quality is associated with service user length of stay

Study design
This is a prospective cohort study of 50 rehabilitation units. Using early recruitment data as a guide (03/02/2012), it is necessary to obtain data on 315 service users from 35 units for adequate power, although the study has the provision to recruit more than this if recruitment time was maximised. The original protocol was based on recruitment of 350 service users and thus extension of recruitment to 50 units should allow this target to be achieved. The mean number of service users per unit is currently seven. Service user inclusion will be determined by whether the rehabilitation unit they are on at baseline is included in the study. The service users will be followed up at 12 months, and if they have left the unit by then, their allocated care coordinator or key worker will be contacted to provide follow up data. For this study, no data are provided directly by the service users; all data will be provided by the staff on the unit or the given service user's care coordinator or key worker and case notes.

Study period
Baseline data collection began in July 2011. Follow up data will be collected until December 2013.

Inclusion criteria
- The unit scoring above the median on the QuIRC[43] in phase 1 of REAL.
- Service users will be included if they are on one of the included units at baseline and give informed consent to take part in the study. If they lack capacity to consent they will be included in the study. Additionally there will be some service users who will not consent because they are away from the unit when consent is required (although the Research Associates will do as much as possible to contact these people and give them the opportunity to take part in the study).

There are no exclusion criteria

Data collection

Baseline

Service user data collected from the staff

- Demographics (age, gender, ethnic group)
- Diagnosis
- Length of history
- Number of previous admissions (and whether voluntary/ involuntary)
- Length of current admission and whether voluntary/involuntary
- Risk history (there are about 20 questions on this and this is a possible mediator of outcome, especially discharge)
- Social functioning as assessed by the Life Skills Profile[75]
- Substance use, assessed using the Clinician Alcohol and Drug Use Scales[76]
- Challenging behaviours which may make community placement difficult, assessed using the staff rated Special Problems Rating Scale[77]
- Staff attitudes towards service user progress, assessed using a likert scale graded 1 to 5 in response to the statement: "I expect this person to be able to move on to a more independent setting within the next 12 months"
- Time Budget Diary[74]

Unit data

- QuIRC[1]

12 months follow up

The same questionnaires and instruments used at baseline will be used at 12 months.

In addition, data will be gathered on service users' move-on/ successful discharge

- Readiness for or achievement of community discharge
- If discharged where to
- If discharged whether maintained community placement or moved on again or readmitted
- If discharged, length of admission and length of rehab admission
- Overall, whether "positive" outcome i.e. successfully discharged with no placement breakdown or readmission +/- further move to less supported placement

Data entry

Most data will be entered by the Research Associates to a Microsoft Access database. Any possible errors in data entry found by the Statistician will be referred to the Research Associates to check their data collection sheets, correct the database and resend it to the Statistician.

Total scores of standardised measures will be calculated using Stata. A Stata do file will be created to produce these calculations and stored. The senior statistician will check a few of these calculations at random. Explanations for any deviations will be sought from the Research Associates, who will check their data extraction sheets, amend the data and resend to the Statistician as appropriate.

Data that form the 145 items of the QuIRC will be entered directly into the QuIRC website (http://www.quirc.eu/), with paper copies of the responses being made too. The resulting data will be extracted by the Project Manager/ Research Associates and sent to the Statistician for analysis. If there are any queries relating to these data from the Statistician, they will be referred to the Project Manager/ Research Associates, who will check them, correct the data and send back to the Statistician.

Primary outcomes

Life Skills Profile (LSP)[75] is a set of 39 staff rated items which are answered using a four point likert scale with the most socially acceptable/ positive response scoring 4 and the least socially acceptable/ most negative response scoring 1. This measure can be summed to give an overall score ranging between 39 and 156. There are also subscales for this measure; these are composed as follows:

- Self-care is the sum of scores for items 10, 12, 13, 14, 15, 16, 23, 24, 26, 30 (possible range 10 to 40)
- Non-turbulence is the sum of scores for items 5, 6, 25, 27, 28, 29, 32, 34, 35, 36, 37, 38 (possible range 12 to 48)
- Social contact is the sum of scores for items 3, 4, 20, 21, 22, 39 (possible range 6 to 24)
- Communication is the sum of scores for items 1, 2, 7, 8, 9, 11 (possible range 6 to 24)
- Responsibility is the sum of scores for items 17, 18, 19, 31, 33 (possible range 5 to 20)

The subscales will not be analysed in this study.

Readiness for/ achievement of community discharge These are collected as dichotomous (yes/ no) variables. They will be analysed as separate variables and as a composite variable; the composite being readiness or achievement of community discharge. Although this takes account non-availability of community placements that may prevent discharge, it was decided to separate the two components to see whether there were different factors associated with the two parts.

Secondary outcome
Length of admission will be recorded from the case notes for each service user on the unit. This will either be to the discharge if before 12 month follow up, or to the 12 month follow up if the service user is still on the unit at follow up. Additionally, the length of time on the unit within the admission will be reported. Likewise this will be recorded to discharge from the unit or until 12 month follow up. The more important measure of length of stay will be length of time on the rehabilitation unit in this admission, as that is what the QuIRC is concerned with.

Main independent variables
QuIRC domains

Covariates
These will be selected a priori.
Age
Sex
Length of illness
Mental Illness Needs Index (MINI) score[45]
Baseline measure of the outcome where LSP or length of stay are the outcomes
Risk history (assault on others in the past two years)
Percentage of service users on the unit who are detained (unit level variable).
Special Problems Rating Scale (SPRS) score[77]
Clinician Alcohol and Drug Scale (CADS) score[76]

Statistical analyses
Data will be analysed using Stata version 13.

Descriptive analyses

Service user level data
It is expected that there will be data on more than 300 service users.

The distribution of variables at baseline and 12 months will be explored. Analysis on continuous variables will include measures of central tendency, and variability. For categorical variables initial examination of the data will calculate frequencies and percentages with given characteristics.

Unit level data
Unit level data (questions within the QuIRC) will be described in a similar way to service user data, although caution will be exercised when reporting these as they will be based on 50 rehabilitation units so it is likely that some numbers will be small and it may be necessary to collapse categories for some variables for reporting purposes. If this is necessary, it will be carried out with clinical consultation to ensure clinical meaningfulness is maintained.

Analysis of primary and secondary outcomes
Outcomes at 12 months will be analysed at the service user level accounting for clustering by rehabilitation unit in the analysis with multilevel models. Reporting of analyses will focus on coefficients and odds ratios where appropriate and 95% confidence intervals as the size of effects (and clinical significance) are more important than statistical significance (and p-values).

Each outcome will be considered separately although will include the same covariates, although only one QuIRC domain at a time (so that there will be seven primary models for each outcome; we will not use the overall QuIRC score).

For readiness for/ achievement of community discharge, it may not be possible to include all the covariates listed on the previous page due to a lack of power. We will carry out a precision calculation to determine the maximum number of covariates that can be included in the modelling process using an estimated intraclass correlation coefficient of 0.04 and the assumption that 20%-30% of service users will experience the outcome. If the number of variables that it is possible to reasonably include in the model is lower than the number of covariates listed previously in this document, then variable reduction will either be carried out using univariable associations or propensity scores.

It is likely that the length of stay in months will need to be transformed as it is expected to be right skewed. The most appropriate transformation to normalise the data before analysis is carried out. After modelling, the residuals will be checked. If they are skewed, then an alternative modelling method will be considered.

We will examine the interactions between QuIRC domain and sex, length of illness and social function (using the LSP in the models where the LSP is the outcome only) individually to investigate whether the outcome is influenced by the effect of two variables working together. These models will include the same a priori covariates as for the primary analysis. However, we realise that the study is not powered to provide conclusive results on interactions, so all results will be exploratory and indicative, and secondary to analyses without interactions. Results of these analyses will be presented with coefficients or odds ratios as appropriate and 95% confidence intervals without p-values.

Missing data
It is anticipated that there will be little missing data for the outcomes, however there may be missing data for covariates. If this is more than 10% of the total, then we will

consider employing imputation techniques. As a precursor to this, we will find out the predictors of missingness for that variable and include those as well as other clinically important variables and the outcomes in the imputation.

Before imputation we will calculate the intraclass correlation coefficient (ICC) for the outcome(s) in question. Results from Phase 1 indicate that this may be substantial (~0.04), in which case it will be necessary to account for this in the imputation process; probably using REALCOM-IMPUTE[81] within Stata.

After imputation, similar analyses to the complete case analyses will be carried out.