THE DEVELOPMENT OF A SCALE TO AID CASELOAD WEIGHTING IN A COMMUNITY MENTAL HEALTH TEAM FOR OLDER PEOPLE

Abstract

Background: Community mental health teams (CMHTs) receive high numbers of referrals and staff who work in these teams are under constant pressure to take on more patients. Local CMHT staff expressed the need for a more accurate way to represent their workload than purely reporting the number of clients on their caseload.

Aims: The project aimed to develop a brief case weighting scale (CWS) specifically to rate cases in CMHTs for older people in Leicestershire.

Methods: An eight-item scale was designed to include the main aspects of a client case that contribute to increased work input by the CMHT member. A multiple regression model was used to assign different weightings to each of the items in order to predict time spent on the case over a four week period.

Results: The items on the scale together ‘accounted for’ 58% of the variation in time spent on client cases. The re-test and inter-rater reliability of the scale were shown to be good. Agreement with the HoNOS 65+ was investigated.

Conclusions: The CWS can be used to rate both current client cases, and new referrals following first assessment, and it is hoped that it will provide CMHT staff with a useful and useable indicator of their clinical workload.

Declaration of interest: None.

Keywords: Community mental health, workload, older people
Introduction

Community Psychiatric Nurses (CPNs) often report excessive workload and constant pressure to take on more patients (McFadyen et al., 1998; McNelly et al., 1998). Numerous referrals and caseloads which are excessive are related to high levels of occupational stress for CPNs (Carson et al., 1996; Edwards et al., 2000; McLeod, 1997) and may have a negative impact upon quality of care and upon efficacy of staff involvement (King et al., 2000; 2004; McFadyen et al., 1998; Onyett et al., 1997).

It is therefore important that CPNs and other community team members are able to assess their caseload in an objective way that represents actual workload rather than just the number of cases. Workload increases as the number of cases increases, but it is important to also take into account the case-mix of the caseload. Measures reflecting the individual requirements of each client should be better indicators of workload than case numbers alone (Carson et al., 1996); this would form a valuable tool to assist staff in managing their caseload, enabling them to establish their workload ‘capacity’ in the form of an optimal caseload (McFadyen et al., 1998). It would enable appropriate allocation of new referrals and equity of workload amongst staff (King et al., 2004), would aid the planning of staffing and resources (Bridel, 1993), ensure quality of care, and reduce staff stress and burnout (Meldrum and Yellowlees, 2000).

A comprehensive literature search revealed limited research in this area. Some studies describe relatively simple case ‘weighting’ systems in which cases are allocated to one of a few pre-defined groups according to whether they are assessed to be simple, straightforward, or more complex (Fortune and Ryan, 1996; Haylock and McGovern, 1989; Muir-Cochrane, 2001). Scores for all cases are summed to give a total score for the weight of the caseload. These
systems were reported to be useful in indicating the current weight and mix of a caseload, and establishing a suitable maximum caseload, but were not formally evaluated.

Storfjell et al. (1997) described a successful system that assessed cases on the basis of client characteristics and complexity of care, but it was designed to be used by home health nurses in USA. Cheah et al. (1998) developed a 35-item scale indicating the problems and needs of clients of a community psychiatry team. This showed reasonable correlation with measures of client need and dependency, but poor correlation with staff time spent on the case. Another scale, based on criteria to prioritise clients taken from the DoH handbook of Mental Illness, showed limited inter-rater agreement and poor agreement with measures of client need, quality of life and functioning (Phelan et al., 2001).

Todd and Caffrey (2002) developed a tool to indicate the impact of various factors (client, environment, indirect contact activity, and direct contact) on staff time in a community learning disabilities team. The authors reported that, having decided upon ‘acceptable’ total scores for each professional group, use of the scale aided management of caseloads. However, no evaluation of the tool was undertaken.

Meldrum and Yellowlees (2000) developed a Clinical Load Monitoring (CLM) tool for case managers in community care teams in Australia. The tool rated face-to-face contact, extent of clinical complication and need for community liaison. The authors reported that the CLM tool was very useful, enabled equitable allocation of caseloads, and provided useful data on the size of individual caseloads during supervision of staff, enabling all staff to monitor and manage their caseloads appropriately. Longitudinal CLM data identified chronic ‘high-
demand’ cases to be passed on to a specialist team, enabling more effective use of the case managers’ time.

Several measures have been developed to assess client need in general psychiatry and take into account that factors other than diagnosis. Sociodemographic factors, physical health status, functional ability, social network and caregiver characteristics contribute to need and predict service utilization (Boersma et al., 1997). However assessment of need, either from the patient’s or the professional’s perspective, does not necessarily directly relate to the amount of CMHT input required.

Whilst there is no widely accepted universal measure to indicate workload intensity, many researchers/clinicians have attempted to develop measures that at least meet their needs to manage client cases in their speciality. Many described as helpful lack formal evaluation, and would not be suitable for use with all client groups.

The needs of older people with mental health problems in the community and the demands that are placed upon CMHT staff differ from those in other specialties. Staff in Mental Health Services for Older People in Leicester identified the need for a scale to assist in the management of their caseloads. Ultimately, the ideal scale would predict the amount of input required for each client. Many factors influence this, but a scale that identifies key client characteristics that affect the amount of staff time required could be a useful tool in weighting cases. The aim was to develop a scale based on individual clients’ needs at assessment (or review) that would be a useful indicator of amount of time likely to be required by the case, which can be used to represent, in a meaningful way, the caseload of community team staff.
Methods and Results

Phase 1. Compilation of the scale

Through a variety of exercises CMHT staff identified factors that they felt contribute to demand for staff time. These factors were then collated into a list of 14 items, each of which were described in 4 levels requiring increased input (see Appendix), and ranked 0-3. (An item ‘other factor’ was also included to allow staff to specify other features of a case they felt relevant.)

Phase 2. Testing the scale

There are ten CMHTs for Older People in Leicestershire. For each team a list of current client cases was generated from the Trust information system and a random sample of 20 cases was generated. Cases assessed for the first time within the previous four weeks were excluded (on the basis that some information required for the completion of the form may not be available). The relevant keyworker was asked to complete the caseload weighting scale (CWS) for each selected case, and to estimate the amount of time they had spent working on that case over the preceding four weeks. Staff were asked to make as accurate an estimate of time as possible, by referring to their case notes. They were asked to include in this estimate time spent on face-to-face contact with the client, contact with carers, telephone calls to the client, their carers and other involved agencies/professionals, note-keeping and networking. They were asked to estimate travel time for the case separately.

At all stages of the project, the Consultant Psychiatrists for the CMHTs agreed to the use of client data for the study. Clients who were selected for inclusion in the study were informed by their keyworker that their information was going to be used, that it would be anonymous,
and that they had the option to request that it was not used. Local Research Ethics Committee approval was obtained.

Data analysis

197/200 (98.5%) CWS were returned. Only forms completed by a CPN were included in the analysis (n= 186). Fourteen forms had missing data and 11 had zero value for total time, leaving 161 for inclusion in the analysis.

A plot of ‘observed’ total time (i.e. time spent on the case over the preceding 4 weeks) showed the distribution to be skewed (min 0.25hr, max 8.00hr, mean 2.78hr). Total time was therefore transformed using log10 to produce a distribution nearer to Normal for analysis (hence cases with zero time input were excluded). Each item was placed as a covariate into a multiple linear regression with the log(base 10) of observed total time as the dependent variable. Travel time was included as a variable as this was felt to be an essential factor when considering time spent on a client case. This model produced an R-squared value of 0.55. A backwards selection procedure for general linear modelling was carried out in order to identify those items that were significant. Two-way interaction terms between these variables were selected in the same way. The final model comprised of five of the original 15 items (deliberate self harm, physical problems, adequacy of current care package, quality of informal support & travel time) and two interaction terms. Because a scale comprising of only these five items might have limited face validity, three items (those relating to the complexity of mental health problems, co-operation and compliance with care package, and self neglect) were added to the regression model to improve the credibility of the scale. The resulting final model produced an R-squared value of 0.58. (See Table 1).
The 95% limits of agreement for the ratio of observed total time/predicted total time were calculated according to the method described by Bland and Altman (1986), and were 0.4 to 2.4, with geometric mean 1.0. A plot of the difference between the log(observed total time) and the log(time predicted by the model) shows the agreement between the two (Figure 1).

**Phase 3. Assessment of inter-rater and intra-rater agreement**

The CWS was revised to include only those items in the above regression model. In order to investigate the re-rater agreement of the CWS, the keyworkers for three randomly selected cases from each CMHT were asked to rate the client using the CWS, and to repeat the rating one week later (without reference to the first). In order to investigate the inter-rater agreement of the scale, three further randomly selected clients from each CMHT were rated on the CWS by two different staff members.

It was decided to look at the level of agreement between the CWS and a measure of dependency, on the basis that client dependency and case weight/time would be closely related and likely to show significant agreement. The HoNOS 65+ (Burns et al., 1999) was selected as an appropriate measure for this purpose. It is a brief rating scale developed specifically to measure the particular physical and cognitive problems affecting older people. It is a 12-item measure, with each item rated on a 0-4 scale, giving a score of dependency between 0 and 48.

Each CMHT was asked to complete both the CWS and the HoNOS 65+ scale on a random sample of ten of their cases.

**Data analysis**
Twenty-four cases were rated on two occasions by their keyworker, and 22 by two different members of staff, using the CWS. The CWS was ‘scored’ using the linear regression model derived in Phase 2, giving a score that is a predictor of time spent on the case over 4 weeks. A total of 87 cases were rated using both the CWS and the HoNOS 65+.

**Inter-rater and Intra-rater agreement**

Agreement between the two CWS ratings was analysed according to the method described by Bland and Altman (1986). The geometric mean for the ratio of first rating/second rating by the same CPN was 1.01 (95% limits of agreement -0.43 to 2.37). The geometric mean for the ratio of first rating/second rating of the same client by two different staff was 1.41 (95% limits of agreement 0.39 to 5.07).

**Agreement with the HoNOS 65+**

The CWS score was recalibrated to be on the same scale as the HoNOS 65+ using the ‘rcal’ command from the ‘merror’ add-on to the STATA statistical computer package, and agreement between the two scores was then analysed according to the method described by Bland and Altman (1986). The geometric mean difference was 0.87 (95% limits of agreement between -10.7 and 10.6 points).

**Discussion**

This project aimed to develop a scale that is simple and brief enough for routine use by CMHT staff in the course of their clinical work, and only requires knowledge already available regarding a patient, following an initial assessment. The study has resulted in an 8-item, easy-to-use scale that can be used to represent the time spent on a client case over four
weeks. The scale can be completed quickly and easily by a staff member who is familiar with the client, usually without reference to the case notes.

The items on the CWS together ‘account for’ 58% of the variation in time spent on the client case. The geometric mean value of observed time/predicted time was 1.0 (i.e. 100% agreement) and for 95% of cases the predicted time was somewhere between an under-estimate of 40% of the true (observed) time, and an over-estimate of 250% the true time. Whilst this appears a large range, it is due to ‘outliers’ or extreme cases – the average discrepancy is much more important. A model that accounts for nearly 60% of the variance in time spent on a case is entirely satisfactory, considering that the CWS is estimating a complex phenomenon which potentially is influenced by a wide array of clinical, administrative and personal factors.

The inter-rater and re-rater reliability, and the agreement between the HoNOS 65+ and CWS, were analysed according to the method described by Bland and Altman (1986) rather than often-used but inappropriate correlation coefficients. The geometric means of ‘rating 1/rating 2’ for both the intra-rater agreement and the inter-rater agreement were good. The lower level of agreement between two different CMHT staff was unsurprising. The level of involvement of the keyworker is likely to be different from that of another member of the CMHT. Furthermore the CWS was developed solely on data provided on CPN caseload, and is likely to be a less good predictor of other CMHT members’ time.

The HoNOS 65+ is a scale that incorporates a wide range of social, physical and psychological functions in one score. Analysis of the agreement between the two scales showed that, having converted the CWS score to the equivalent scale of the HoNOS 65+, the
two scales were within 10 points. Complete agreement between a scale designed to reflect
client dependency and a scale designed to reflect time-input for the CPN would not
necessarily be expected, but some degree of concordance was anticipated.

The CWS was developed using ‘time spent on the case over the previous four weeks’ as the
‘gold standard’ indicator of input required for a client. Therefore rating a current or newly
assessed case using the CWS should predict the time-input that will be required for that client
over the next four weeks. CMHT staff can rate their whole caseload giving a total workload
score. This total score is a means of quantifying individual workload in a more representative
way than number of cases, and could be used to aid caseload management, supervision, and
allocation of new referrals. However, for an individual staff member, the total score will vary
depending upon their level of experience, time available for clinical work and perhaps
personal characteristics such as coping mechanisms (and probably other factors, such as team
structure). Comparisons between individual staff (or CMHTs) using total CWS score may be
useful but are not necessarily straightforward as there are many factors not included in scale
that effect the total workload that an individual can manage effectively. A measure of a
CMHT total caseload may also provide service managers with information to aid planning
and management of resources within the service.

The CWS is an indicator/predictor of time, and its development was based on retrospective
estimates of time provided by staff from their case notes. The extent to which the scoring
model predicts the amount of time required for the client’s care is dependent upon the
accuracy of those estimates. Staff were encouraged to refer to their notes, and every effort
was made to make those estimates as accurate as possible. Whilst it would have been ideal to
assess the accuracy of the estimates it would have been difficult and impractical to do so.
The collection of more data using the CWS, along with prospective records of the amount of time spent on a client case over 4 weeks, would allow the scoring model to be made more accurate.

It is most important that the clinical utility of the CWS is established. It is only through extensive use of the CWS, by clinicians in the course of their every day clinical work, that it can be established whether this scale is indeed a useful and useable indicator of workload. This essential next step in the project is currently on-going. A simple spreadsheet, developed to score the scale, enables quick rating of a large number of cases, so that staff can rate their whole caseload easily, without needing to know the complicated weighting formula. Wider usage of the CWS may enable refinement of the scoring formula, and possibly validation of the scale for use by other professional groups in the CMHTs. However adapting the CWS for ‘non-CPN’ CMHT staff may require new scoring/weighting models, or even inclusion of different items in the scale, to reflect the different roles and responsibilities that are part of their work with clients.

CMHT cases will vary over time in terms of the amount of assistance they require from their keyworker. Commonly there will be times when ‘high input’ is required, and the client and their families need more support, typically when a client is first referred, and then at times of change or ‘crisis’. Use of the CWS scale over time will also provide data regarding the sensitivity of the scale to change in a client’s case.

The incentive to commence this project came from a clearly identified need by a local CMHT. Whilst there are many scales that assess severity of mental illness, or client need, these do not necessarily reflect how time-consuming working with a client will be for a CPN.
When this study began no suitable measure existed. Although new research has been published more recently it would have to be established whether the scales described would be appropriate for use with older people who have mental health problems, as there are issues particular to dementia and to older people that may not be covered by a scale designed for use with another client group.

In light of this, and because the intention was to develop a scale for our CMHTs, the decision was made to develop the scale “in house”. The items that comprise the CWS were those aspects of client cases identified by local CMHT staff as contributing to increased demand for staff time. Whilst the items ultimately selected to comprise the scale are by no means particular to the local service, it is possible that aspects of the way that local CMHTs work with their clients and the way services are structured locally may have impacted on the selection of the items in the scale and the resultant scoring algorithm. Although originally designed for use by local teams, the potential utility of the scale in other similar services is apparent. As the scale was developed in a large mental health trust, representing a diverse range of communities and demography, it might reasonably be expected that the findings will have some generalisability. Evaluation of the scale in a different trust would be an obvious next step in its development in order to assess its generalisability for other services.
APPENDIX - 15 original items of the CWS (those marked * were retained in the final version of the scale)

RISK / VULNERABILITY

- Deliberate self harm*
- Self Neglect*
- Vulnerability to Abuse
- Safety

NATURE OF THE ILLNESS

- Clarity of diagnosis
- Complexity of mental health problems*
- Physical problems*

CO-OPERATION AND COMPLIANCE

- Co-operation and Compliance with the care package*
- Co-operation and Compliance with your input

CLIENT SUPPORT NETWORK (statutory agencies/formal care)

- Contact with primary care
- Existing care package
- Adequacy of current care package*

INVolVEMENT OF CARERS (family/friends/neighbours)

- Quantity of ‘Informal’ support - family, friends and neighbours
- ‘Quality’ of informal support*

OTHER FACTORS

- To be specified by respondent

TRAVEL TIME (over last 4 weeks)*
Example of an item:

Respondents were instructed “*Where a series of statements are given for each value, select the number which most accurately reflects the client’s case. A client need not score all of the elements within a score category but the majority of them should apply.*”

**Deliberate self harm**

0  No intention to self harm. No history of self harm.
1  Intermittent thoughts of chronic wish to die but no plan.
2  Previous episodes of DSH. Constant suicidal thoughts but not actively planning.
3  History of several DSH episodes of high lethality; actively planning.
References


<table>
<thead>
<tr>
<th>Item</th>
<th>Beta</th>
<th>Sig.</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-0.175</td>
<td>0.008</td>
<td>-0.303, -0.047</td>
</tr>
<tr>
<td>Deliberate self harm</td>
<td>0.336</td>
<td>0.000</td>
<td>0.208, 0.464</td>
</tr>
<tr>
<td>Self neglect</td>
<td>0.030</td>
<td>0.191</td>
<td>-0.015, 0.075</td>
</tr>
<tr>
<td>Complexity of mental health problems</td>
<td>-0.015</td>
<td>0.446</td>
<td>-0.054, 0.024</td>
</tr>
<tr>
<td>Physical problems</td>
<td>0.016</td>
<td>0.281</td>
<td>-0.013, 0.045</td>
</tr>
<tr>
<td>Co-operation and compliance with care package</td>
<td>0.015</td>
<td>0.541</td>
<td>-0.032, 0.061</td>
</tr>
<tr>
<td>Adequacy of current care package</td>
<td>0.045</td>
<td>0.021</td>
<td>0.007, 0.084</td>
</tr>
<tr>
<td>Quality of informal support</td>
<td>-0.045</td>
<td>0.346</td>
<td>-0.138, 0.049</td>
</tr>
<tr>
<td>Travel time</td>
<td>0.206</td>
<td>0.000</td>
<td>0.137, 0.274</td>
</tr>
<tr>
<td>DSH*Travel time</td>
<td>-0.133</td>
<td>0.000</td>
<td>-0.196, 0.070</td>
</tr>
<tr>
<td>Quality of Informal Support*Travel time</td>
<td>0.063</td>
<td>0.025</td>
<td>0.008, 0.117</td>
</tr>
</tbody>
</table>
Figure 1. Plot of the difference between log(observed time) and log(predicted time) against average log(observed time) and log(predicted time)