Validation of the UKATT Process Rating Scale (PRS)

Tobey, G., Clyne, W., Finnegan, O., Farrin, A. and Russell, I., in collaboration with the UKATT Research Team

Abstract

Aim: To describe the development and validation of the UKATT Process Rating Scale (PRS), a manual based method for monitoring and rating the delivery of psychosocial treatments of alcohol dependence and misuse.

Methods: Following adaptation and further development of a validated rating scale, the ability of the UKATT PRS to rate the delivery of video recorded treatment in the UK Alcohol Treatment Trial (UKATT) was tested.

Results: Tests of the validity and reliability of the UKATT Process Rating Scale (PRS) show that it is able reliably to detect the two treatments for which it was designed and to discriminate between them.

Conclusions: The UKATT PRS is a valid and reliable method of monitoring and quality of therapeutic style and content in the delivery of two psycho-social treatments of alcohol use and dependence.

Introduction

The purpose of the study was to develop and validate a manual-based, time efficient method of rating treatment fidelity, including frequency and quality of the delivery of a UK version of Motivational Enhancement Therapy (MET) and a new treatment, Social Behaviour and Network Therapy (SBNT) (Copello et al. 2002), compared in effectiveness in the UK Alcohol Treatment Trial (UKATT 2005-2009). Training and supervision of practice were delivered centrally for the purpose of quality control and maintenance of treatment fidelity (Tobey et al., 2008). 742 clients attending for alcohol problems treatment were recruited to the trial and 590 of these attended at least one treatment session. Video recordings of the two treatment trials were used for supervision of therapist adherence and were also independently rated using the UKATT PRS.

Sampling

One video per trial client (where available) was sampled for process rating. The sample of over 400 video tapes was stratified by treatment (MET, SBNT), session number (1-3 for MET, 1-8 for SBNT) and centre. Replacement sampling was used when a video was unrateable, to retain balance between treatments, session numbers and centres. A target of 50 randomly selected videos to be double rated, and of these 25 to be triple rated by two further independent raters was set, ensuring inclusion of ratings throughout the entire treatment phase, balance by treatment, session number and centre.

Analysis

Data were collected and analysed using SPSS version 14.

Validity

To test construct validity, the factor structure of the scale was examined using Principal Component Analysis. Summary scores were calculated for treatment specific items which had factor loadings of more than 0.25 on a single treatment component: METI was the mean of the frequency scores for MET items; METq was the mean of the quality scores for MET items where frequency ratings were greater than 0; SBNTI and SBNTq were calculated similarly.

The ability of the scale to discriminate between the two treatments was investigated by comparing individual item scores and frequency summary scores for each treatment. It was hypothesised that MET items would be rated higher for MET sessions and low for SBNT sessions and vice versa. A t-test was used to compare the mean item scores and the mean frequency summary scores between the two treatments. Concurrent validity was examined by comparing manual derived quality summary scores for the two treatments with global ratings of individual therapist’s skills (low/medium/high) given by the treatment specific supervisors. These global ratings were derived following an instruction to the supervisors to base their response on consistency and quality of delivery across the whole period of treatment.

Results

452 clients (58.4% of 744) had a rateable video: 259 were for MET and 193 were for SBNT. Selection of videos was successful in capturing a spread of equivalent proportions across the sessions.

Construct validity

Principal Components Analysis of treatment specific therapist task and style items showed a dimensional eigenvalue of 5.13 accounting for 43% of variance of the single factor solution, providing an adequate characterisation of the data. Eighteen treatment specific items had a loading greater than 0.25; all of the nine originally hypothesised MET items and seven un-hypothesised SBNT items. MET items all had positive loadings and SBNT items all had negative loadings suggesting a treatment component where the more MET was practised, the less SBNT was practised.

Confirmatory validity

Global ratings of quality of therapists’ delivery treatment were provided in three categories (high, medium and low quality) by the two treatment specific supervisors and compared with quality summary ratings made by the primary rater for the whole sample. The magnitude of ratings between the primary rater and the supervisors showed concurrence in that rater derived scores were highest in the supervisors’ high category and lowest in the supervisors’ low category.

Criterion validity

Table 1 shows mean frequency scores for the treatment specific items for both treatments. There is a significant difference between frequency ratings in each of the treatment with a higher rating in each case for the treatment for which the item was designed.

<table>
<thead>
<tr>
<th>Item</th>
<th>MET</th>
<th>SBNT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interpersonal focus</td>
<td>1.53 (0.75)</td>
<td>1.40 (0.69)</td>
</tr>
<tr>
<td>Empathy towards patient</td>
<td>2.17 (0.60)</td>
<td>2.00 (0.60)</td>
</tr>
<tr>
<td>Treatment style</td>
<td>2.26 (0.70)</td>
<td>2.10 (0.60)</td>
</tr>
<tr>
<td>Supervision</td>
<td>1.97 (0.85)</td>
<td>1.90 (0.75)</td>
</tr>
</tbody>
</table>

SBNT items were significantly higher (p<0.001) in SBNT (MET item mean = 1.3) than in SBNT (MET item mean = 0.4); 95% CI for the difference = 0.91 to 1.02. Mean scores for frequency of SBNT items were significantly higher (p<0.001) in SBNT (MET item mean = 1.4) than in MET (MET mean = 0.4; 95% CI for the difference = -1.10 to -0.93).

As quality scores were only given if the item was given a frequency rating of 1 or more that is, if the item was rated as having occurred; some items had very low numbers of quality ratings, particularly for the treatment to which those items were not attributed. Items with ten or more quality ratings were included in the analysis. Of the thirteen items with sufficient data, three of seven MET items showed a significantly higher quality score for SBNT than for MET. Six of the seven SBNT items had significantly higher ratings of quality for SBNT than for MET (see Table 1).

SBNT quality summary 2 samples mean±SD

<table>
<thead>
<tr>
<th>Item</th>
<th>MET</th>
<th>SBNT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Supervision</td>
<td>1.97 (0.85)</td>
<td>1.90 (0.75)</td>
</tr>
<tr>
<td>Interpersonal focus</td>
<td>1.53 (0.75)</td>
<td>1.40 (0.69)</td>
</tr>
<tr>
<td>Treatment style</td>
<td>2.17 (0.60)</td>
<td>2.00 (0.60)</td>
</tr>
<tr>
<td>Empathy towards patient</td>
<td>2.17 (0.60)</td>
<td>2.00 (0.60)</td>
</tr>
</tbody>
</table>

SBNT items were significantly higher (p<0.001) in SBNT (MET item mean = 1.3) than in SBNT (MET item mean = 0.4); 95% CI for the difference = 0.91 to 1.02. Mean scores for frequency of SBNT items were significantly higher (p<0.001) in SBNT (MET item mean = 1.4) than in MET (MET mean = 0.4; 95% CI for the difference = -1.10 to -0.93).

Reliability

Item analysis was conducted separately for frequency of MET items and for frequency of SBNT items using Cronbach’s Alpha of .71 for MET items and .76 for SBNT items.

Inter-rater reliability as measured by the intraclass correlation coefficient (ICC) is reported in Table 1. The items generally show high values of ICC, indicating good levels of consistency between raters as the majority of the variation is attributable to the clients, rather than the raters. The SBNT item, active agent for change, shows a low level of consistency as an ICC of 0.28, while the MET items creating conflict, exploration of feelings and empathy show moderate levels of consistency with ICOS of 0.45, 0.51 and 0.60 respectively.

Conclusions

The UKATT-PRS is a valid and reliable method of rating the delivery of two psychosocial treatments for alcohol problems and dependence and identifying which one is being delivered. It is likely to be adaptable to rating the delivery of other psycho-social treatments applying the same principles used in its development. It can therefore form the basis of the measures of performance and treatment fidelity in clinical trials, in treatment audit and in routine supervision of practice.

Acknowledgments

Veronica Morton for assistance with statistical analysis and the UKATT Research Team for their contribution to the study. The research was supported by a grant from the Alcohol Education and Research Council.

References


