Tacit and encoded knowledge in the use of standardised outcome measures in multidisciplinary team decision making: A case study of in-patient neurorehabilitation

Joanne Greenhalgh, Rob Flynn, Andrew F. Long, Sarah Tyson

University of Leeds, Leeds LS2 9JT, UK
University of Salford, Salford, UK

Abstract

This paper explores how multidisciplinary teams (MDTs) balance encoded knowledge, in the form of standardised outcome measurement, with tacit knowledge, in the form of intuitive judgement, clinical experience and expertise, in the process of clinical decision making. The paper is based on findings from a qualitative case study of a multidisciplinary in-patient neurorehabilitation team in one UK NHS trust who routinely collected standardised outcome measures. Data were collected using non-participant observation of 16 MDT meetings and semi-structured interviews with 11 practitioners representing different professional groups. Our analysis suggests that clinicians drew on tacit knowledge to supplement, adjust or dismiss 'the scores' in making judgements about a patients' likely progress in rehabilitation, their change (or lack of) during therapy and their need for support on discharge. In many cases, the scores accorded with clinicians' tacit knowledge of the patient, and were used to reinforce this opinion, rather than determine it. In other cases, the scores, in particular the Barthel Index, provided a partial picture of the patient and in these circumstances, clinicians employed tacit knowledge to fill in the gaps. In some cases, the scores and tacit knowledge diverged and clinicians preferred to rely on their clinical experience and intuition and adjusted or downplayed the accuracy of the scores. We conclude that there are limits to the advantages of quantifying and standardising assessments of health within routine clinical practice and that standardised outcome measures can support, rather than determine clinical judgement. Tacit knowledge is essential to produce and interpret this form of encoded knowledge and to balance its significance against other information about the patient in making decisions about patient care.

Introduction

Outcome measurement in clinical practice

Outcome measurement is one element within the long term trend towards 'scientific-bureaucratic medicine' (Harrison, 2002, 2004) and the increased regulation of medical professionals in both the UK (Flynn, 2002) and the US (Timmermans & Berg, 2003). In practice, outcome measurement involves many different activities in
meeting several different policy agendas. There are many different types of health outcome measures and numerous ways of categorising them exist, based on the content or scope of concepts measured (e.g., impairment, disability, handicap, health-related quality of life), the purpose of the measurement (e.g., to make predictions, to distinguish between groups or individuals or to measure change over time) and how the measurement is performed (e.g., generic versus disease-specific measures, a profile versus an index). Such measures are also designed to be completed by either the patient themselves or by clinicians. However, most health outcome measures have two features in common. Firstly, health outcome measures aim to standardise the way in which dimensions of health states are judged through preset questions or scoring guidelines. Thus, identical criteria are used across different individuals and groups to measure a common dimension of health, to allow comparison between and within individuals or groups. Secondly, they seek to quantify a particular construct or attribute by allocating a set of numbers to descriptions of different health states using an ordinal and sometimes a ratio scale. Thus, they aim to provide an indication as to whether a particular dimension of health has improved or worsened over time, or whether one group or individual has worse health than another.

Outcome measurement has had a dual role in the promulgation of evidence-based practice. Health outcome measures are increasingly used within randomised controlled trials as criteria to judge the effectiveness of different treatments and care packages. Such trials then form the basis of guidelines and protocols which seek to standardise clinical practice and provide a rational foundation for clinical decisions. However, there has also been increasing interest in the collection of outcomes information within decision making in routine clinical practice. Information from an outcome measure is viewed as one of a number of cues to which the process of decision making is then applied. Moreover, this presents a limited view of clinical work; ethnographic studies have shown that clinical decision making is a collective activity that occurs amongst groups of clinicians in a diffuse, iterative way over a protracted length of time, often in several locations.

Outcome measurement has had a dual role in the promulgation of evidence-based practice. Health outcome measures are increasingly used within randomised controlled trials as criteria to judge the effectiveness of different treatments and care packages. Such trials then form the basis of guidelines and protocols which seek to standardise clinical practice and provide a rational foundation for clinical decisions. However, there has also been increasing interest in the collection of outcomes information within decision making in routine clinical practice. Information from an outcome measure is viewed as one of a number of cues to which the process of decision making is then applied. Moreover, this presents a limited view of clinical work; ethnographic studies have shown that clinical decision making is a collective activity that occurs amongst groups of clinicians in a diffuse, iterative way over a protracted length of time, often in several locations.

Outcome measurement has had a dual role in the promulgation of evidence-based practice. Health outcome measures are increasingly used within randomised controlled trials as criteria to judge the effectiveness of different treatments and care packages. Such trials then form the basis of guidelines and protocols which seek to standardise clinical practice and provide a rational foundation for clinical decisions. However, there has also been increasing interest in the collection of outcomes information within decision making in routine clinical practice. Information from an outcome measure is viewed as one of a number of cues to which the process of decision making is then applied. Moreover, this presents a limited view of clinical work; ethnographic studies have shown that clinical decision making is a collective activity that occurs amongst groups of clinicians in a diffuse, iterative way over a protracted length of time, often in several locations.

The nature of clinical decision making

Randomised controlled trials exploring the use of outcome measures in clinical practice and other work exploring the importance clinicians place on outcome measures have, implicitly, drawn on cognitive theories of decision making. Such theories have conceptualised decision making as a time limited activity that occurs within the private world of the clinician's mind at a single site, most often, the clinician–patient interface. Information from an outcome measure is viewed as one of a number of cues to which the process of decision making is then applied. However, this presents a limited view of clinical work; ethnographic studies have shown that clinical decision making is a collective activity that occurs amongst groups of clinicians in a diffuse, iterative way over a protracted length of time, often in several locations.

Outcome measurement has had a dual role in the promulgation of evidence-based practice. Health outcome measures are increasingly used within randomised controlled trials as criteria to judge the effectiveness of different treatments and care packages. Such trials then form the basis of guidelines and protocols which seek to standardise clinical practice and provide a rational foundation for clinical decisions. However, there has also been increasing interest in the collection of outcomes information within decision making in routine clinical practice. Information from an outcome measure is viewed as one of a number of cues to which the process of decision making is then applied. Moreover, this presents a limited view of clinical work; ethnographic studies have shown that clinical decision making is a collective activity that occurs amongst groups of clinicians in a diffuse, iterative way over a protracted length of time, often in several locations.

Outcome measurement has had a dual role in the promulgation of evidence-based practice. Health outcome measures are increasingly used within randomised controlled trials as criteria to judge the effectiveness of different treatments and care packages. Such trials then form the basis of guidelines and protocols which seek to standardise clinical practice and provide a rational foundation for clinical decisions. However, there has also been increasing interest in the collection of outcomes information within decision making in routine clinical practice. Information from an outcome measure is viewed as one of a number of cues to which the process of decision making is then applied. Moreover, this presents a limited view of clinical work; ethnographic studies have shown that clinical decision making is a collective activity that occurs amongst groups of clinicians in a diffuse, iterative way over a protracted length of time, often in several locations.

Outcome measurement has had a dual role in the promulgation of evidence-based practice. Health outcome measures are increasingly used within randomised controlled trials as criteria to judge the effectiveness of different treatments and care packages. Such trials then form the basis of guidelines and protocols which seek to standardise clinical practice and provide a rational foundation for clinical decisions. However, there has also been increasing interest in the collection of outcomes information within decision making in routine clinical practice. Information from an outcome measure is viewed as one of a number of cues to which the process of decision making is then applied. Moreover, this presents a limited view of clinical work; ethnographic studies have shown that clinical decision making is a collective activity that occurs amongst groups of clinicians in a diffuse, iterative way over a protracted length of time, often in several locations.
standardise how a particular dimension of health should be judged across different populations.

More recent work within medical sociology and organisation studies has explored the social production of knowledge and how clinicians render different forms of knowledge useable within clinical practice (Atkinson, 1995; Blackler, 1995; Casper & Berg, 1995; Newell, Edelman, Scarbrough, Swan, & Bresnen, 2003). Knowledge does not simply reside within the brains of individuals or within organisational processes, but is produced through interactions between social actors and is transformed through its application in local contexts. Thus, new knowledge is created through a dialogue between tacit and explicit knowledge, in which tacit knowledge may be transformed into explicit knowledge and vice versa (Nonaka, 1994). Clinicians integrate these different forms of knowledge in their everyday practice through the development of ‘routines’ (Smith, Goodwin, Mort, & Pope, 2003) or ‘mindlines’ (Gabbay & Le May, 2004) in which tacit knowledge from experience is essential to interpret explicit knowledge and apply it in particular circumstances, thus guarding against ‘cookbook’ implementation of codified knowledge.

Many other writers have also stressed this tension – or disjunction – between professional expertise and intuitive judgement, and formalised or rule-based systems (see May, Rapley, Moreira, Finch, & Heaven, 2006). Nettleton et al. (2008) uncovered doctors’ concerns that current regulatory practices within the NHS were limiting the profession’s opportunity to develop valuable tacit, hands on, experiential knowledge. McDonald, Waring, and Harrison (2006) showed how surgeons’ routines are acquired through frontline experience and that the many contingencies in this work mean that it cannot be readily subsumed within guidelines and protocols. Rafalovich (2005) reported that clinicians perceived diagnostic manuals to be useful as a general guide but were not sufficient to account for the multiplicity of factors they took into account in making a diagnosis of attention deficit disorder. Wood, Prior, and Gray (2003) also observed how clinicians used their clinical experience to override or adjust referral guidelines and computerised decision support systems in their decision making.

With regard specifically to the use of standardised outcome measures, surveys and interviews have found that clinicians prefer to rely on their subjective judgement in assessing one particular dimension of health, namely health related quality of life (Gilbody, House, & Sheldon, 2002; McKevitt & Wolfe, 2002). Tannenbaum (1994) observed that clinicians used outcome research only when they were at the limits of their personal experience. Cowley, Mitcheson, and Houston (2004) also showed how health visitors resisted using standardised needs assessments in the manner required of them by their managers. These findings suggest that, in practice, clinicians prefer to rely on embodied or embedded knowledge and often use this knowledge to override, adjust or resist attempts to dictate their practice through encoded knowledge.

However, to date, research has not directly explored how multidisciplinary teams of clinicians use standardised outcome measures in the process of clinical judgement and decision making. It is not clear how such teams balance this form of encoded knowledge with intuitive reasoning in the care of individual patients. This paper explores these issues by drawing on qualitative data collected from one multidisciplinary in-patient neurorehabilitation unit in which standardised outcome measures were routinely collected. The data were part of an ESRC funded study to examine how (and to what extent) health and social care professionals use outcome measures in routine clinical practice.

**Methods of the research**

**Choice of setting**

In-patient neurorehabilitation was selected for investigation for two reasons. Firstly, the use of standardised outcome measures in this setting could be described as ‘common practice’. A survey of rehabilitation units in the UK found that just over three quarters were routinely collecting a standardised measure within clinical practice (Turner-Stokes & Turner-Stokes, 1997) and guidance exists on the most appropriate measure to use in this setting (Turner-Stokes, 2002). The most common instruments used in neurorehabilitation are the Barthel Index (Mahoney & Barthel, 1965; Wade & Collin, 1988), the Functional Independence Measure (FIM) (Hamilton, Granger, Sherwin, Zielezny, & Tashman, 1987) and the FIM plus the Functional Assessment Measure (FIM + FAM) (Hall, Hamilton, Wayne, & Zalser, 1993). They are all measures of the activities of daily living and provide an indication of the degree of dependency on care. A more detailed description of the Barthel Index is provided in Table 1.

Secondly, patients admitted to neurorehabilitation units have a broad range of conditions (for example traumatic brain injury, stroke, multiple sclerosis) that involve multiple physical and cognitive impairments which require input from many different specialties and therapies. As a result of this complexity of condition, neurorehabilitation teams include practitioners from different professions such as medicine, nursing, physiotherapy, speech therapy, occupational therapy, clinical psychology, and social work. Each professional group brings a different tacit knowledge base and expertise to the team (Anspach, 1987; Colombo, Bendelow, Fulford, & Williams, 2003), which will influence the ways in which they define and measure outcomes. As such, these multidisciplinary teams (MDTs) provide a focal point for debate around outcome measures and thus a means of examining whether, and how, such measures are being used in clinical decisions. The goal of the study was not to examine the ‘effectiveness’ of the MDT’s decision making or identify ways in which clinicians could use outcome measures to ‘improve’ their decisions. Rather, the aim was to explore the contribution that standardised outcome measures made to the process through which the team made judgements about patients.

**Description of the setting**

To preserve anonymity, we have changed the names of the unit, patients, staff and locations mentioned throughout the paper. Data were collected at the Churchtown Centre, a 19-bed unit providing in-patient neurorehabilitation...
Table 1
Description of the standardised outcome measures used at the Churchtown Centre

<table>
<thead>
<tr>
<th>Name of measure</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Barthel index</td>
<td>Validated measure of activities of daily living. Although several versions exist, the Churchtown Centre used the modified Barthel Index (Shah, Vanclay, &amp; Cooper, 1989). Scoring guidelines indicate that items should be scored based on what the patient does, not what the patient could do. It contains ten items scored on a weighted five point ordinal scale covering personal hygiene, bathing, feeding, getting on and off the toilet, climbing stairs, dressing, bowel function, bladder function, transferring from a bed to a chair, and ambulation. The total can be expressed from 0 (complete dependence) to 100. A number of guidelines for interpretation exist. Shah et al. (1989) suggest that a score of 0–20 indicates total dependence; 21–60 indicates severe dependency, 61–90 indicates moderate dependency, 91–99 indicates light dependency. Granger, Sherwood &amp; Greer (1977) suggest a score of 60 indicates the threshold between more marked dependence and independence. In the MDT, the score for each item was read out.</td>
</tr>
<tr>
<td>Northwick Park dependency score</td>
<td>Validated measure of nursing dependency (Turner-Stokes et al., 1998). It has two parts: (1) Basic Care Needs has 16 items scored on either a three, four or five point scale covering mobility, bed transfers, toileting bladder, urinary incontinence, toileting bowels, faecal incontinence, washing and grooming, bathing/showering, dressing, eating, drinking, enteral feeding, skin pressure relief, safety awareness, communication, behaviour; (2) Special Nursing Needs has 7 items each scored 0 or 5 covering tracheostomy, open wound requiring dressing, requires &gt;2 interventions at night, requires psychological support, in isolation (eg for MRSA), acute medical or surgical intervention, needs one to one ‘specialing’. The total score from each section is summed to give a total from 0 to 100. Those scoring less than 10 are classed as ‘low dependency’, those scoring between 10 and 24 are classed as ‘medium dependency’ and those scoring 25 or more are classed as ‘high dependency’. In the MDT the total score only was read out.</td>
</tr>
<tr>
<td>Leeds assessment of handicap scale</td>
<td>Validated measure of handicap (Geddes, Tannant, &amp; Chamberlain, 2000). It contains four items each scored on a nine point scale covering physical independence, mobility, orientation and social integration. Higher scores indicate a greater degree of handicap. In the MDT, the score for each item was read out.</td>
</tr>
<tr>
<td>Waterlow score</td>
<td>Validated scale of risk of pressure sores. It contains 11 items and each item can have several scores that cover age, gender, appetite, continence, skin visual sins, mobility, build/weight, cardiovascular status, neurological deficit, medical, surgery/trauma. Items are summed and those scoring less than 10 are at low risk while those scoring 20 or more are high risk. In the MDT, the total score was read out.</td>
</tr>
<tr>
<td>‘Home grown’ assessments</td>
<td>Unvalidated, single item assessments covering language reception, language expression, functional communication, memory, concentration, confusion, drive/motivation, snacks and meals, anxiety, depression and behaviour.</td>
</tr>
</tbody>
</table>

for patients aged between 16 and 65. The unit was part of a large NHS teaching hospital, but located away from the main hospital site. The unit was led by two consultants with an interest in standardised outcome measures. The unit had been routinely using the Barthel Index for over 30 years and, as such, outcome measurement in some form was a well established practice within the team. At the time of the field work, a mixture of standardised, validated measures and ‘homegrown’ single item assessments (described in Table 1) were routinely collected on all patients. Access to the Churchtown Centre was initially discussed with the lead consultants and then approved by the whole team. University and NHS ethical approval was subsequently obtained and all members of staff gave written informed consent to take part.

Data collection

Data were collected by one member of the research team (JG) using non-participant observation, semi-structured interviews and collection of documentary data (Stake, 2000). We focused our observation on the regular MDT meetings as this was where team members exchanged information about the standardised outcome measures and patients’ progress was reviewed. MDT meetings lasting one-and-a-half hours were held twice a week, and each was led by a different Consultant. A total of 16 MDT meetings were attended and observed, and team discussions in 14 were fully audio-taped. In the period of fieldwork, 39 patients were considered by the team. Following this, at least one clinician from each of the professional groups, together with the two Consultants, were interviewed (n = 11). Throughout data collection, field notes were taken and distributed amongst the research team and a memo was kept to record analytical insights and the impact of JG’s presence within MDT meetings.

Analysis

Qualitative data analysis was iterative and ongoing throughout the study using the techniques of grounded theory (Charmaz, 2006; Strauss & Corbin, 1998) and was aided by QSR Nvivo. The research team met regularly to discuss and agree on emerging themes.

Based on the field notes and transcripts, we initially developed broad themes around: (1) the process through which the MDT ‘scored’ the outcome measures; (2) MDT member’s subsequent comments and discussions about the resulting scores; and (3) the ways in which the MDT judged patients’ progress and planned discharges. We then developed initial hypotheses about whether and how the scores were used in judging progress and discharge planning. To explore these hypotheses further, we constructed ‘patient journeys’. These consisted of all the data about a particular patient during successive MDT meetings for the duration of the fieldwork. We selected five patient journeys for detailed analysis, based on characteristics that would maximise the potential explanatory power of our analyses. These included patients who were classified as ‘delayed discharges’ and patients who were marked as ‘complex’ or ‘difficult’ because either their cognitive or behavioural problems or their family situation...
(or all three) posed challenges for the timing and destination of discharge. Our analysis of the patient journey entailed searching for instances where the scores were used explicitly either as ‘warrants’ or ‘grounds’ for action but also exploring when and why they may not have been explicitly referred to within MDT meetings.

We analysed the interviews alongside the MDT meetings to understand team members’ perceptions of the role the scores played in their judgement of patients’ progress and the extent to which they felt they influenced the team’s judgements and decisions. We explored whether some hypotheses generated from the MDT analysis were supported by the interviews and also looked for examples of team members using outcome measures within the MDT meetings in the ways that they reported using them in their interviews. We were, however, careful not to accord one data source as having greater ‘truth’ than another.

**Findings**

Structure of MDT meetings

Before presenting our analyses it is first important to provide some contextual information about the setting and format of MDT meetings and the nature of the discussions within them. MDT meetings were held in a ‘boardroom’, with a notice board summarising the meeting agenda, the dates and times of family and goal setting meetings for that week and the discharge dates that had been set. Booklets about the scoring guidelines for the outcome measures used by the team and laminated cards detailing onward referral criteria and the process of referral to the early discharge community stroke team were placed on the table. Meetings were usually chaired by a Consultant, and at every meeting there were representatives (both junior and senior) from the medical, nursing, occupational therapy and physiotherapy staff. A family liaison health visitor attended most meetings unless she was engaged in community visits. Medical staff were present throughout the meeting, while different practitioners entered the meeting when ‘their’ patients were being discussed. The speech therapist and clinical psychologist attended when patients receiving their input were discussed, and social workers attended when discharge arrangements were being finalised. Medical, nursing and therapy students also occasionally attended.

At each meeting, six or seven patients were discussed and each was allotted a 15-min timeslot. New patients were usually discussed first. For each patient, the discussion generally (though not always) followed a seven item agenda in which the ‘scores’, as the team called them, were read out first, then new diagnostic information was detailed, the patient’s achievement of therapy goals were reviewed, home visits discussed, onward referrals and discharges were then planned and finally ‘non goal progress’ was assessed.

**Decision making in neurorehabilitation**

In common with Atkinson (1995), Hughes and Griffiths (1997) and White (2002), decision making within the MDT was not a single unitary act but was an ongoing process of negotiation and searching for consensus amongst team members. Specific occasions when ‘decisions’ were made were difficult to pinpoint. Our initial analyses identified that two key, interrelated aspects of the work of the MDT meetings were assessing patient progress and discharge planning. During the process of rehabilitation, the team reviewed the patient’s progress every two weeks. They formed judgements about whether the patient was changing and making progress or whether she/he was no longer changing, but had ‘plateaued’. Discharge planning occurred throughout the patient’s stay in the rehabilitation unit and was subject to revision and negotiation amongst the team and between the team and the patient’s family. It was complex and required judgements about the patient’s ability to function independently, the amount of support that might be needed at home, the ability and willingness of family members to provide this support and the availability of additional services in the community.

We made a number of overall observations about the team’s use of the scores within MDT meetings. The process of scoring followed a particular format and was done separately from the rest of the discussion about a patient. There were few instances when the scores were referred to by team members beyond this process. During scoring, it was rare for team members to articulate the meaning of the scores and their implications for patient care. Rather, much of the use of the scores in judging progress and discharge planning appeared to be implicit. While discussion amongst the team about what the score ‘should be’ frequently occurred, it was not common for the accuracy of the scores to be subsequently called into question by team members. However, there were times when the validity of the scores was disputed. Our analyses expand these observations further and explore a number of possible explanations for these findings. Within each section, we focus our discussion on a particular analytical theme, while recognising that the extracts of data presented often display more than one of these themes.

**The reporting and production of ‘the scores’**

Tacit knowledge played an important role in both producing and reporting on ‘the scores’ within the meeting. The recitation of the scores strongly echoed Strong’s (1979) ‘ceremonial order’ of the clinic, and Atkinson’s (1995) ‘liturgical’ format. It entailed the senior house officer (SHO) or specialist registrar reading out each item on the particular measure, in the same order for each patient, to which the team member would reply with a ‘score’. Team members knew which order the scores were read out in and engaged in repair work to correct this when a new registrar read them out in the ‘wrong’ order. It was also clear that particular scores ‘belonged’ to certain team members, reflecting their professional background; for example, the physiotherapist would score the items on the Barthel that related to mobility (ambulation or wheelchair, transfers and stairs). This division of labour in the scoring was rarely made explicit within team meetings, except when the relevant professional was absent. On these occasions, team members would often comment that the score, for example, was “an OT one” which then allowed other team members
to produce a score in the member’s absence. The ritualistic recitation of the scores necessarily required tacit knowledge about the ordering and the division of labour in the scores, which was shared amongst the team.

Team members also drew on tacit knowledge in order to produce the scores themselves. In many instances, the score was simply read out by the team member. On other occasions, particularly for new patients when the team had not had much opportunity to assess the patient prior to the meeting, the score was produced within the meeting through discussion amongst team members. The extract below, where the team includes a Speech and Language Therapist (SALT), provides an example of such a discussion. The team are trying to determine a score on the orientation item from the handicap measure for a new patient, Mr. Edwards, a 40-year-old man who had severe cognitive difficulties resulting from a head injury.

This extract shows how the process of producing a score involved fitting team members’ clinical knowledge of the patient into the framework set out within the scoring guidelines. In this way, tacit knowledge was transformed into encoded knowledge; the scoring guidelines provided a lens through which clinical knowledge was filtered and made relevant to the task of producing a score. This could be seen as similar to the ways in which devices such as data abstraction forms and protocols channel the ways in which systematic reviewers read and then ‘rewrite’ papers in order to disentangle data from their original sources (Moreira, 2007). The validity of the scoring guidelines themselves were not disputed by team members, and once produced, the scores themselves were rarely called into question but taken as a ‘fixed’ and ‘objective’ form of encoded knowledge. Thus, shared, tacit knowledge underpinned the ritualistic recitation of the scores and the means through which the scores themselves were produced as ‘objective’.

Articulating the ‘taken for granted’

Much of the ways in which the team then used the scores in judging progress or planning discharge was implicit. A typical structure for a case discussion was for the scores to be recited at the beginning and then not referred to again for the rest of the discussion about the patient. Whilst they were being read out, the scores were implicitly employed by the team member giving the score to flag up a lack of change in the patient. Therapists and nurses would often report the scores in such a way that made it clear that change had not occurred, for example, by simply saying that ‘no change’ had occurred, by stating ‘all the same’ or ‘still zero’ instead of, or as well as, giving the score itself. Thus, the scores were not just a piece of information but a form of judgement about the patient’s progress. As such, there was no distinction between ‘information’ and ‘decisions’ in the ways in which the scores were used in patient management (Atkinson, 1995). Immediately after the scoring had been completed, the total Barthel score would be added up by the SHO or specialist registrar who would often make explicit whether the score had improved or not. However, these comments tended to be an ‘aside’ from the main discussion about the patient. The implications for patient management of the lack of change or improvement or the absolute values of the scores themselves were rarely spelled out. To illustrate this, we present an extract from the second time the team scored a stroke patient, Mr. Gray, two weeks following admission.

<table>
<thead>
<tr>
<th></th>
<th>SHO</th>
<th>Transfers?</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>SHO</td>
<td>Transfers zero,</td>
</tr>
<tr>
<td></td>
<td>Physio</td>
<td>ambulation zero, umm...</td>
</tr>
<tr>
<td>2</td>
<td>Physio</td>
<td>five, epic [looking at scoring guidelines] (...) and</td>
</tr>
<tr>
<td></td>
<td>SALT</td>
<td>stairs not assessed.</td>
</tr>
<tr>
<td>3</td>
<td>SHO</td>
<td>Dressing?</td>
</tr>
<tr>
<td>4</td>
<td>OT</td>
<td>Umm two</td>
</tr>
<tr>
<td>5</td>
<td>SHO</td>
<td>Hygiene?</td>
</tr>
<tr>
<td>6</td>
<td>OT</td>
<td>One</td>
</tr>
<tr>
<td>7</td>
<td>SHO</td>
<td>Bathing?</td>
</tr>
<tr>
<td>8</td>
<td>OT</td>
<td>One</td>
</tr>
<tr>
<td>9</td>
<td>OT</td>
<td>One</td>
</tr>
<tr>
<td>10</td>
<td>(Pause while SHO adds up the Barthel scores)</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>SHO</td>
<td>Fifty-seven</td>
</tr>
<tr>
<td>12</td>
<td>(Pause)</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>SHO</td>
<td>Language?</td>
</tr>
<tr>
<td>14</td>
<td>SALT</td>
<td>Two, two, two.</td>
</tr>
<tr>
<td>15</td>
<td>(further ‘homegrown’ scores are read out)</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>SHO</td>
<td>Northwick Park is twenty-seven.</td>
</tr>
<tr>
<td>17</td>
<td>SHO</td>
<td>Yes</td>
</tr>
<tr>
<td>18</td>
<td>Registrar</td>
<td>Erm ...handicap mobility ... and Orientation?</td>
</tr>
<tr>
<td>19</td>
<td>SALT</td>
<td>Five b.</td>
</tr>
<tr>
<td>20</td>
<td>Registrar</td>
<td>They’re the same, the same <a href="">quietly</a></td>
</tr>
<tr>
<td>21</td>
<td>OT</td>
<td>Shall we go through the goals then?</td>
</tr>
<tr>
<td>22</td>
<td>Registrar</td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>SHO</td>
<td></td>
</tr>
</tbody>
</table>
This extract clearly shows that there was often little discussion about the meaning or implications of individual item scores or for the total Barthel and Northwick Park scores. Once the scores had been recited, the SHO swiftly moved the discussion on to reviewing the patient’s goals.

From their interviews, it was clear that team members felt patients’ scores on admission did have implications for their likely progress in therapy or their needs at discharge. For example, the speech therapist explained that “everybody does, you know, if you hear a Barthel score of 0 or 10 or you know that this person is going to be heavily dependent.” The family liaison health visitor noted “if it’s a low Barthel and a high handicap then that’s indicating that we’re going to have to put some service, well quite a bit of service, in the home care”. Team members were also able to interpret the significance of changes in the scores during the process of rehabilitation. Consultant 1 observed that “if they change by only two or three it means nothing at all, but if a dependency or Barthel goes from nought to thirty you know you’ve got a dramatic difference”. One explanation for why this knowledge appeared to be largely absent from MDT meetings is that, through continued use of the scores, team members had developed a shared understanding of the scores and their implications for patient management, which did not need to be verbalised. As the nurse observed about the scores in her interview “I mean obviously the more you do them the more used you get to them”.

In support of this explanation, we found that this shared, taken for granted knowledge was more likely to be verbalised in the presence of third parties, such as medical, nursing, physiotherapy or speech therapy students, who could not be expected to know about the meaning of the scores. The following extract is taken from a meeting in which medical students were present and concerns the discussion of Mr. Smith, a stroke patient who was about to be discharged. Consultant 1 had asked a medical student to play the role of the SHO and read out the names of the score items.

<table>
<thead>
<tr>
<th></th>
<th>Medical student</th>
<th>Handicap (pause) is MO, motor.</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>Nurse</td>
<td>Motor, yep</td>
</tr>
<tr>
<td>3</td>
<td>Physio</td>
<td>Eight</td>
</tr>
<tr>
<td>4</td>
<td>Consultant</td>
<td>That means he can’t get out of bed.</td>
</tr>
<tr>
<td>5</td>
<td>Physio</td>
<td>No, not quite</td>
</tr>
<tr>
<td>6</td>
<td>Medical student</td>
<td>And OR? Orientation.</td>
</tr>
<tr>
<td>7</td>
<td>OT</td>
<td>Zero.</td>
</tr>
<tr>
<td>8</td>
<td>Consultant</td>
<td>That’s normal.</td>
</tr>
<tr>
<td>9</td>
<td>(Long pause while Medical student adds up scores)</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Medical student</td>
<td>Sixty.</td>
</tr>
<tr>
<td></td>
<td>Consultant</td>
<td>Right, so what can you, a Barthel of sixty out of hundred, what… it’s an unfair question but what do you think? Er, is he likely to need help or not? And how much help?</td>
</tr>
<tr>
<td>12</td>
<td>Medical student</td>
<td>Likely to need some help.</td>
</tr>
<tr>
<td>13</td>
<td>Consultant</td>
<td>Yes he is. And what is his Northwick Park score?</td>
</tr>
<tr>
<td>14</td>
<td>SHO</td>
<td>Fifteen.</td>
</tr>
<tr>
<td>15</td>
<td>Consultant</td>
<td>Fifteen, so he needs a fair amount of help.</td>
</tr>
</tbody>
</table>

Here we see the consultant teaching the medical students about the meaning of individual items in the handicap measure and the implications of the total Barthel score and Northwick Park Dependency score for the amount of help this patient will need at home. She describes her question as ‘unfair’ suggesting that the medical student cannot be expected to know what the total Barthel score means. This is reminiscent of Atkinson’s (1995) observations that the presence of medical students led senior staff to be more explicit than they otherwise would about the knowledge they drew upon in conducting their work. Wood et al. (2003) also found that, for unproblematic cases, the rules for interpreting referral guidelines were not articulated as the team shared a common understanding of these rules. Thus, the scores did have implications for patient management but, in most cases, these implications did not need to be articulated as the team held a shared, taken for granted understanding of them.

**Encoded versus embedded knowledge**

Although the scores did have implications for patient management, team members did not use them in a deterministic fashion. They had to be integrated with and balanced against team members’ own intuitive assessments of the patient. The structure of MDT meetings, with the ‘scores’ dealt with first followed by a more narrative discussion about the patient, created an apparent division between the ‘scores’ and other clinical and social information about the patient. In their interviews, team members also drew a distinction between the ‘scores’, or encoded knowledge, on the one hand, and their ‘clinical impression’, or tacit knowledge, on the other. For example, the nurse indicated that the scores do “show whether the patient’s improved, you know. We ourselves know whether the patient’s improved from week to week but that’s just going to put down on paper that the patient has improved” suggesting that the scores provided written evidence of what the clinicians themselves already ‘know’. The occupational therapist felt that the Barthel “does help you to, I guess, categorise that person going on past experience of similar patients and I guess that’s why more experienced staff will be more confident in getting a discharge date” indicating that the scores only made sense when interpreted in the context of years of clinical experience of similar patients. Consultant 2 also explained that she felt someone was not progressing “based on a mixture of seeing no change in the scores and clinical experience and expertise” and warned that it “certainly shouldn’t be one without the other”. Thus, encoded knowledge in the form of the scores was not a sufficient basis for clinical decisions; tacit knowledge was also necessary to interpret them.

These comments raise questions about how the team integrated these two types of knowledge in judgements about patients’ needs or their progression in therapy. Our observations of MDT meetings suggested that at times when the scores and other clinical information about the patient were in broad agreement the scores were rarely discussed further, as there was no time, and no need to do so because they simply supported ‘what everybody knew’. In some instances, the scores were strategically
used by the team to augment the credibility of the team’s clinical assessment of the patient’s progress in therapy (or lack of) to family. While discussions about what the score should be was common, we could find no examples where the scores were used to challenge a team member’s clinical opinion of a patient. However, when there appeared to be a discrepancy between the scores and other clinical information, team members challenged the validity of the scores or engaged in repair work to explain why the scores were not accurately reflecting the patient’s progress. Less commonly, the team ‘moulded’ the scores to fit with their clinical impression through invoking an ‘unwritten rule’ of giving the patient a low score if they were uncertain of the score. We present two examples to illuminate these observations further.

Mr Edwards was a 40-year-old man who had sustained a serious head injury resulting in cognitive problems. He had also broken his leg, which prevented him from weight bearing and limited the work that the physiotherapists could do with him during rehabilitation. On admission, Mr Edward’s Barthel score was zero and was still zero once the scoring had been completed at the start of his third case conference, four weeks after admission. However, during this case conference, it transpired that Mr Edwards had met a number of his rehabilitation goals which led the occupational therapist to ‘correct’ his Barthel feeding score and the nurse to adjust his Barthel personal hygiene score to reflect their impressions, joking that they had been “diddling him out of a two!”. The physiotherapist then explained “you see his other scores are down because we can’t get him on his feet. He’d be scoring much more because he wants to put his foot down and walk and he has walked on the ward”. Thus, the scores were not adequately representing the patient’s progress and required both adjusting and supplementing with clinical information to paint a more accurate picture of the patient’s progress.

Mr Smith was a 63-year-old man who had a subarachnoid haemorrhage that had led to a car accident, resulting in severe cognitive difficulties. At his first case conference shortly after admission, the consultant pronounced that this patient was “high dependency” and estimated that his Northwick Park Dependency score would be “about forty-two”. The SHO indicated that “some of his standard scores, most of his Barthel’s like the feeding, and, he’s zero on every one”. Consultant 2 predicted that “unfortunately, I don’t think we’re going to get very far with this man” and that he was “destined for a nursing home”. In this instance, the scores confirmed and were used to reinforce the consultant’s clinical impression of this patient. The nurse suggested that they completed the scores on the patient because “I think the wife will have a lot of questions about just how good he will get and if we are saying we’ve not assessed them it will look a bit shabby”.

During the second meeting about Mr Smith, two weeks after admission, a number of his ‘homegrown’ cognitive scores had decreased. The registrar joked “It doesn’t look good for scoring, he’s down from two to one” on a number of the scores. The consultant felt that they had been “generous” with his scores the previous week and the registrar rationalised that “we may have been guessing last week”. Consequently, Consultant 2 explained “I think that this is indicating getting to know him rather than any deterioration” which served to downplay the accuracy of the scores and revealed that this accuracy relied on their clinical knowledge of the patient.

Two weeks later, towards the beginning of the case conference, the speech therapist summed up the team’s impression that they were “just not getting anywhere with him really, he’s just not changing” and that further rehabilitation may not be beneficial. However, the patient’s wife felt that her husband was not changing because the team was not giving him enough therapy. The specialist registrar suggested that they scored the patient “just to demonstrate to her if she’s going to query in terms of progress”. The occupational therapist scored the patient a five for the dressing item on the Barthel, suggesting an improvement from a score of two the previous week. However, she qualified this by saying “he’s so variable” and the family liaison health visitor suggested they take the “bottom line”, “otherwise”, the specialist registrar observed, “the score would be improving quite a lot” so the patient’s score remained at two. A similar ‘moulding’ of the patient’s personal hygiene score was also undertaken, again rationalised by the registrar “if you’ve any qualms, it must be zero”.

These examples show that, when the scores and the clinical picture of the patient agreed, the scores were used to back up clinical opinion. However, when the scores and the clinical picture of the patient were at odds, clinicians gave primacy to their clinical impression in judging progress and sometimes adjusted the scores to fit. This was partly because they did not trust the scores. Consultant 2 acknowledged that “scores are not infallible, we know that they have limitations in terms of sensitivity and specificity and sensitive to change in particular, so you can’t rely on a single measure” to make decisions. She then explained that if they judged someone had reached the “level of activity limitation that we were anticipating” then they would not “wait for several weeks to see if the scores the same, we’ll use the combined information that we have in order to reach the decision as a team”. The specialist registrar observed that “sometimes, you know, the scores aren’t changing for that person and at the end of the day, you go on your clinical impression, you know, feeling, if you want to call it that, rather than the score”. The SHO also felt that the scores “reflect the decisions rather than the decisions reflect the scores” and that the score was “just a guide, it’s not used exclusively to make a decision on people’s discharge position or on ongoing treatment or anything like that”.

However, although clinicians preferred to rely on their clinical impression of the patient in assessing their needs and judging their progress, there were times when they strategically used the scores to augment the credibility of this impression to ‘external’ parties such as the patient’s family. In the case of Mr Smith, described above, it was clear that the scores were being used to convince the patient’s wife of the team’s perception that the patient wasn’t changing during therapy. It was perhaps because of pressure to justify their decisions to family members that clinicians were more likely to adjust the scores to reflect their tacit knowledge.
In her interview, Consultant 2 explained that they would use the scores “as evidence” in discussions with the family when the patient appeared “not to change despite continuing rehabilitation”. In these instances, the specialist registrar felt that the scores were “maybe more objective” than “a description of what they are doing earlier and later” in therapy. Tannenbaum (1994) also found that clinicians used outcome research in communicating with family members, not because it influenced their decisions, but because it rendered their clinical decisions more objective.

The limits of encoded knowledge

Even in situations where the scores and other clinical information did not directly contradict each other, it was rare to find instances in which the scores were explicitly used to justify a course of action in relation to the timing or destination of discharge. While the scores did provide an indication of the amount of help patients might need if they were to return home, they were not the determining factor in deciding the time or place of discharge. To have any role in this process, the scores had to be placed in the context of the patient’s social situation and available services. For example, Mr Brooks was a 60-year-old man who had had a stroke. His discharge planning was complicated by the fact that his home was difficult to modify and his wife was also in poor health and was thus unable to support him. During his initial case conference Consultant 1 interpreted his scores for the benefit of attending medical students as follows:

“Alright so that’s a low Barthel and a high handicap. What’s that, sixteen, twenty-six and that for a lone male that would indicate that they would probably need some sort of supported living. But he’s not alone but on the other hand I should think he ought to be, if he was well he’d probably be giving his partner support”.

Here we can see that the consultant interprets the meaning of the scores in relation to the patient’s social situation. The scores themselves often did not capture the real practical constraints affecting discharge and had to be supplemented by clinical and social knowledge about the patient.

A different example was Mr Green who had sustained a head injury resulting in serious cognitive problems. During his stay in rehabilitation his Barthel scores continued to improve but his cognitive and consequent behavioural problems showed little improvement, which were not reflected in his Barthel scores. This led Consultant 2 to describe him as potentially a “mobile wanderer”. The team had referred him to a specialist centre for people with brain injuries but there was a long waiting list and it was by no means certain that the centre would accept him. In the interim, the occupational therapist reported that “the family kind of clearly gave the message that they would be unable to support him” while Consultant 2 judged that the patient would “be rather mobile for Ryeburn House [nursing home] isn’t he, cause most of the people in Ryeburn are really wheelchair bound”. Thus the scores simply did not reflect the complex cognitive problems experienced by the patient, nor his family situation or the availability of existing services, which complicated the process of discharge planning for this patient.

In their interviews, clinicians also highlighted these problems with scores, the Barthel Index in particular, which limited their role in discharge planning. The physiotherapist noted that if their patients had cognitive or behavioural problems “they’ll still have a high dependency which may not be picked up for example on a Barthel, so they may score quite highly on a Barthel but actually be quite impaired”. Consultant 2 also expressed concerns that the “Barthel score is never intended to pick up those sorts of problems but I sometimes feel that outsiders may come in and say ‘oh that person looks independent they must be fine they can go home on their own’ whereas the reality is somewhat different if they had significant cognitive dysfunction”. The occupational therapist also commented that “we’ve also had people who have scored a hundred on the Barthel that still need in-patient rehab because of their cognitive levels and their behaviour. So I don’t see it as a hugely indicative measure of discharge outcome”.

They also commented that the scores did not reflect the patient’s social situation and the availability and ability of the family to act as carers. A social worker explained that discharge was very dependent on the patient’s “social situation, because if they are saying, a fit partner there ready and waiting to take somebody home, they will go home, I think, much earlier” but she didn’t know “how much the measures actually help with that”. The nurse also observed that “there are other things why they can’t go home… That doesn’t cover in the scoring … if they haven’t got anyone to help them at home. There are a lot of other things, it just doesn’t tell. If they get to a hundred then they can go home, but obviously it’s not like that”. The family liaison health visitor cautioned that “before we say ‘oh there’s a husband at home, Barthel score says they’ll cope with the help of a husband or wife at home’” the team needed to look at the family members “abilities to cope as carers”.

Discussion and conclusions

This paper aimed to understand how multidisciplinary teams use standardised outcome measures in the context of broader literature about the nature of clinical decision making and medical knowledge. It is acknowledged that, given the case study approach and qualitative methods used, any claims to generalisability must be limited. Nevertheless, the observational techniques used were invaluable in providing rich descriptive and narrative data about otherwise opaque procedures. The site we observed was purposively sampled, and (so far as the authors are aware) was broadly typical of similar units elsewhere in the English NHS. Our analyses thus provide some insight into the ways in which multidisciplinary teams balanced tacit knowledge, in the form of clinical experience, expertise and their clinical and social knowledge of the patient, with encoded knowledge, in the form of standardised outcome measures in the process of clinical judgement and decision making. It remains to be determined whether the role of tacit knowledge in using standardised outcome measures in neurorehabilitation is mirrored in other specialties, clinical conditions or with other outcome measures.

Our findings suggest that, at least in the case of standardised outcome measures, there is a complex relationship...
between tacit and encoded or explicit knowledge when applied in practice. Tacit knowledge underpinned the production of the scores; the scoring guidelines were the means through which team member’s judgements of the patient’s abilities were transformed into a number, a particular form of explicit knowledge. Once produced, their means of production were rarely questioned by the team, thus sustaining the status of the scores as ‘fixed’ and ‘objective’ and distinct from clinical experience and judgement. The structure of the meeting which separated the process of scoring from a more narrative discussion of the patient also further reinforced their objective status. Thus, tacit knowledge was integral in producing and maintaining the scores as an explicit form of knowledge.

The team we observed had used the standardised measures for many years and had developed a shared, ‘taken for granted’ understanding of the meaning and the implications for patient management of the absolute value of the scores and changes in the scores over time. Thus, a form of ‘embedded’ knowledge had developed amongst the team, based on experience of using the scores with similar patients, which rendered this form of explicit knowledge useable in practice. This is similar to Atkinson’s (1995) observations that medical students need to be taught how to ‘read’ the body and that medical ‘facts’ require active interpretation. While scoring guidelines were important in assisting with this interpretation, it was through regularly applying these guidelines to patients that enabled the development of this embedded knowledge. This reinforces Polanyi’s (1966) assertion that tacit knowledge is acquired through practical experience and is transferred through close interaction and the development of a shared understanding amongst team members.

In common with studies that have examined clinician’s use of other forms of encoded knowledge in decision making (Rafalovich, 2005; Tannenbaum, 1994; Wood et al., 2003) the standardised outcome measures played a supporting, rather than a determining role in patient management. Clinicians drew on their own expertise and impression of the patient to supplement, adjust or dismiss the scores in making judgements about a patients’ likely progress in therapy, their change (or lack of), during therapy and their need for help on discharge. In many cases, the scores accorded with the clinicians’ judgement of the patient, and were used to reinforce this opinion, rather than determine it. The Barthel Index was not designed to measure the patients cognitive functioning, their social situation, nor take account of the availability of services in the community. Thus, in many cases, the scores provided only a partial picture of the patient. In these circumstances, clinicians employed tacit knowledge to supplement the scores in making decisions about the timing and destination of discharge. In some cases, when progress in therapy was being assessed, the scores and tacit knowledge diverged. In these cases, clinicians preferred to rely on their tacit knowledge and used this to adjust or downplay the accuracy of the scores.

Paradoxically, having recognised the imprecision of some scores, the clinicians often invoked their symbolic value when justifying their decisions to the patient’s family. The credibility of the scores, therefore, was not static but was shaped by clinicians in strategic ways to suit their needs in different contexts and to different audiences. Thus, in the case of standardised outcome measures, the meaning and application of encoded knowledge was open to judgement based on tacit understandings and contextual constraints. This corresponds with Casper and Berg’s (1995) observation that medical data and diagnoses are actively constructed and re-constructed, ‘moulded’ to make cases manageable (see also Berg, 1997).

In summary, our findings suggest that the production and practical application of standardised outcome measurement in practice involved a dynamic interplay between tacit and explicit knowledge, which calls into question a simple dichotomy between the two forms of knowledge (Blackler 1995; Smith et al., 2003). While our findings agree with observations that encoded knowledge fails to reflect the tacit skills and judgement of individuals (Polanyi, 1966; Lam, 2000), they also suggest that knowledge may be transformed from one form to another when used in practice (Nonaka, 1994). For example, tacit knowledge was used to produce the scores, a form of encoded knowledge, and was then essential in enabling team members to recognise the meaning and implications of the scores for patient management. Our findings also show that tacit knowledge itself is heterogeneous and encompassed not only clinical experience acquired through observations of many patients alongside clinical impressions, gut feelings and knowledge of particular patients, but also local, shared understandings of ‘the way we do things around here’, for example, in the specific ways in which the scores themselves were produced, reported and interpreted. Taken together, these findings suggest that, in order to understand how clinicians integrate tacit and encoded knowledge in judgements and decision making, exploring the relationships between different forms of knowledge and how they are produced is as important as delineating their differences (Blackler, 1995).

Our findings also have a number of implications for the implementation and the use of standardised outcome measures in routine practice, and more broadly, for attempts to ensure greater standardisation and uniformity in medical practice. An often cited barrier to using outcome measures in routine practice is that clinicians do not know what the scores mean (Gilbody, House, et al., 2002; Meadows, Rogers, & Greene, 1998). However, we would suggest that it is only through continued use of the scores in clinical practice that clinicians can develop an understanding of the meaning of the scores and their implications for patient management. It has also been argued that the routine use of standardised outcome measures in clinical practice offers a more systematic method of assessing various dimensions of a patient’s health (Schor et al., 1995), with the assumption that they may improve the accuracy clinical judgement. However, our findings indicate that there are limits to the advantages of quantifying and standardising assessments of health within clinical practice and that standardised outcome measures can only ever be a support, rather than a determinant of clinical judgement. Clinical experience and intuition and knowledge of local, shared routines are essential both to produce and interpret.
this form of encoded knowledge and to balance its significance against other information about the patient in making decisions about patient care.

References


