A Quality Assurance Method for Community Occupational Therapy

MaryAnn McColl

Fiscal constraints, service withdrawal, and litigation in the health care sector have created an unquestionable need for accountability in health services. Therefore, this paper presents a method of quality assurance for use in community occupational therapy. Detailed attention is given to methodological issues, particularly those related to selecting an approach. A process audit approach was chosen for this study, and the results of its pilot application to a sample of depressed patients are presented and discussed. In addition, preliminary work on reliability and validity of the method are presented. Test-retest and interrater reliabilities were found to be .734 and .728, respectively, and assertions of normative and criterion validity were supported. The results of the audit are presented to illustrate aspects of the method that need further refinement.

Little more than ten years ago, quality assurance was an ideal entertained only by those practitioners who had the time, resources, and expertise to permit detailed self-examination. Recently, however, public and professional concern about undesirable outcomes, deficiencies in care, and subsequent litigation have underlined the need for quality assurance (1). In addition, recent structural changes and increasingly stringent funding requirements have changed the focus of quality assurance. The focus has changed from enhancing the health of the population to making sure that important services are not cut (2).

During the past several years, occupational therapists have become aware of the need to evaluate and improve services in traditional settings (3-5). However, there is no literature that describes how to apply the principles of quality assurance to the community practice of occupational therapy. Existing models, designed for use in institutional settings, are not appropriate for community application. The objectives and interventions employed by community therapists differ from those of hospital therapists, so that the same evaluation tools cannot be used by both. Therefore, community therapists have little in the way of published standards against which to gauge their practice. In addition, the relative isolation in which community therapists work offers them only infrequent opportunities for intra-professional comparison. Thus, to a large extent, community therapists have neither formal nor informal guidelines for practice or opportunities for professional comparison.

The purpose of this study was to develop a model of quality assurance for community occupational therapy. While the model has the potential to enhance the quality of care given to a specific population by occupational therapists in a community setting, we hope it will also be useful for studying other topics and other community health agencies.

Description of the Model

The model we devised (see Figure 1) is based on the work of a number of authors (3, 5-7). The model is circular, emphasizing the ongoing nature of quality assurance and the need for remedial action and reassessment to complete the cycle. This study is primarily concerned with steps 1 through 7 of the model. Steps 8 and 9 are administrative tasks, which are outside of the domain of the research. This paper describes in-
Figure 1
Generic model for quality assurance

1. Select topic
2. Select method
3. Identify criteria
4. Ratify criteria
5. Collect data
6. Analyze deviations from criteria
7. Identify problems
8. Institute remedial action
9. Evaluate results of remediation

An Account of the Decision-Making Process

Method Selection

On examining the literature on quality assurance in health care, we were confronted with a multiplicity of definitions, activities, and methods. Without a systematic approach, the range of methodological choices is overwhelming and discouraging. The "decision tree" (see Figure 2) attempts to unravel some of the complexity of those decisions and to organize them into a systematic scheme for use by health care practitioners. The decision tree is constructed on five levels (from left to right), each of which is discussed below, along with the decision we reached on each parameter for the present study.

On the first level, Luke and
Modrow (10) propose three overall methods of ensuring quality. The method of choice at this level is usually determined by the economic sector in which a person works. For example, the first two methods of administrative or legislative regulation and competition are most often applied to industry and business, respectively. The third choice, voluntary self-regulation, is generally reserved for "the professions." Traditionally, professional groups have conducted their own quality review and have not been subject to administrative control or public scrutiny.

The second level of the decision tree is related to the overall approach to the process of quality assurance and includes two possible choices: a) the implicit approach, meaning quality of care assessment based on subjective professional judgment, and b) the explicit approach, meaning assessment based on objective criteria (8). Consistent with the principles of empirical inquiry, we chose the explicit approach for this study (11).

The third, fourth, and fifth levels of the decision tree represent a series of dichotomous decisions related to specific areas of methodology: composition of the criteria committee, method of data collection, and type of standards or criteria.

Mono-Versus Multidisciplinary. The option of mono-versus multidisciplinary refers to the professional affiliation of individuals who set the criteria for evaluation. Although some feel that a multidisciplinary committee increases the credibility of the criteria established (4, 8), this study took a monodisciplinary course. There is considerable support in the literature for the use of a monodisciplinary approach in that only an intradisciplinary committee has sufficient understanding of the range of practice to formulate meaningful criteria (10). In addition, only people within the discipline have access to professional sanctions, should these become necessary (10).

Internal Versus External. This choice refers to the institutional affiliation of the individuals who set the criteria. In this regard, the present study takes advantage of both internal and external sources of input. By including experienced therapists from within the agency, we were able to use their experience and knowledge of the process of treatment. By obtaining the contribution of a therapist from a major referral source outside of the agency, we enhanced the objectivity and variety of input for criteria development.

Peer Versus Authority. These two alternatives refer to the relationship between the criteria setters and the therapists to be evaluated. A peer committee has the advantage of creating a more supportive and less threatening environment (12) and of providing educational spin-offs to the therapists involved in the evaluation. On the other hand, criteria set by a person of authority are generally felt to be valuable because of that person's greater expertise and objectivity. Again, we chose to take a middle road in this study, using a combination of senior therapists (who had supervisory responsibility) and staff therapists on the criteria committee.

Concurrent Versus Retrospective. This decision refers to the time frame used for the study: concurrent implies the use of patients currently under care and retrospective implies the use of patients already discharged from the service. Although the concurrent time frame offers more up-to-date information, the retrospective time frame was adopted for this study because it makes a more economical use of time and resources.

Audit Versus Observation. The choice of audit versus observation refers to the method of data collection. Audit refers to any data collection process that is based on a review of recorded materials; observation refers to any participative or observational method of direct data gathering. The observational method gives the investigator greater control over the quality and uniformity of data. However, we chose the audit method for this study because a) the chart review process used in the audit method is much less costly and time consuming than is the recruitment, training, and supervision of raters used in the observational method and b) the audit method reduces the element of error introduced by the presence of a participant observer (13).

However, one requirement must be fulfilled to successfully implement the audit method: all charts must be at a certain acceptable level of completeness and uniformity to ensure that the audit is measuring the quality of care rather than the quality of documentation (8). This requirement has been fulfilled in the present study, because a routine monthly documentation audit has been in place at Community Occupational Therapy Associates since March 1980. Therefore any discrepancies found between the criteria and the record in question can be reliably interpreted as oversights in treatment rather than in documentation.

Process Versus Outcome. Se-
lecting the type of criteria that should be used in quality assurance involves perhaps the most complex and controversial decision. The two principal types of criteria used in quality assurance are process and outcome. Process criteria reflect the activities or procedures that should take place as a part of good care; outcome criteria refer to the desired end results of treatment. A long-standing debate exists as to which, in terms of validity, is preferable. The current trend, given the present climate of fiscal constraint in health care and other public sectors, is toward outcome evaluations (evaluations based on achievement of results), because outcomes appear to be inherently valid. Therefore, many health programs take this approach to comply with funding requirements. However, the outcome approach has a number of problems, especially when used with a mental health population.

McAuliffe (14) argues that despite assertions to the contrary (13, 15), outcome measures may be less valid than process measures because they must be considered indirect measures of the quality of care. Since “care” is, in fact, a process, McAuliffe suggests that the quality of care is more directly gauged by focusing on the process of care and that outcome measures may not possess the face validity that (according to their proponents) is supposed to be their principal advantage.

McAuliffe (14) notes two other problems with using outcome measures in quality assessment. He says that outcome measures exclude many of the relevant aspects of care that are not directly reflected in the end result. He suggests that measures of outcome often include variance that is attributable to numerous causes not related to the particular intervention being studied. Because many variables unrelated to medical care account for a major proportion of the final outcome variance, it may not be possible to draw any inferences from many studies of outcome in health care services. Schulberg (16) points out that this is especially true in the mental health field, where outcomes usually are somewhat less tangible and often are attributable to personal and environmental factors that are outside of the realm of therapeutic intervention.

A fourth problem with outcome measures, particularly in the field of mental health, is the inability of therapists to identify outcomes that are universally desirable in patients at discharge. There is no apparent consensus among therapists, or in the literature, as to the objective, measurable status that patients should have achieved when treatment is terminated (17). If outcomes cannot reliably be identified with some degree of professional agreement, then the validity of studies based on those outcomes will surely be called into question, that is, it will be argued that the outcome measured is not the appropriate one to reflect quality of care.

A final problem with using outcome as a measure of quality of care is the general quality of the data produced by many outcome measures. While this is less of a problem with truly objective indicators (e.g., mortality), it must be a concern when somewhat “softer” outcome indicators are used (e.g., health status, functional level, or quality of life) (14). Few measures exist whose empirically documented levels of reliability and validity are high enough to make them truly acceptable (1). Also, the necessity of relying on subjective reports from patients for many of these measures adds a component of error that is difficult to quantify.

Although no clear-cut answer exists to the dilemma over process versus outcome, we considered all of these problems when choosing a process approach for our study.

**Normative Versus Empirical.** The final decision in choosing the method for quality assurance is choosing the level of standard. Empirical standards are based on actual practice and compare the care given in the facility being studied to the usual level of care given in other settings. While empirical standards are attainable, and therefore credible, they may not measure whether or not the care is consistent with the current level of professional knowledge (13). Normative standards, on the other hand, are based on professional values and reflect the optimum level of care compatible with the current state of knowledge. Therefore, normative standards aim to ensure not only a minimum acceptable level of care but also an optimum quality of practice. In this study, we adopted a normative standard.

Thus, we chose a retrospective process audit as the most appropriate method for our study. The committee that formulated the criteria was monodisciplinary; it consisted of therapists internal and external to the agency and included therapists in supervisory and staff positions.

**The Pilot Application**

**Criteria Identification**

The development of criteria for
quality assurance is synonymous with the identification of indicators against which evidence of practice will be compared. The literature unanimously recommends that this task be undertaken by a committee whose members are selected on the basis of credibility, expertise, and variety of input. Through a series of four meetings, 29 process criteria were established to represent the range of desirable interventions with depressed patients. A procedure entitled nominal group process was employed to achieve consensus with a minimum of discussion (5). Based on recommendations in the literature, criteria were then weighted (using the nominal group process) to reflect the relative importance of their contribution to the overall treatment goal, which is, functional and social independence (8, 11). The criteria were then pretested on a sample of ten charts, refined, and incorporated into the scoring sheet (see Figure 3).

Criteria Ratification

Once the committee was satisfied with the scoring sheet, the instrument needed to be ratified by all therapists against whose records it might be applied. Therefore, the scoring sheet was presented to all therapists at Community Occupational Therapy Associates, and their unanimous support was achieved.

Data Collection

Sample Selection. The sample of depressed patients was selected at random from those discharged from the agency over a 12-month period. In the interest of homogeneity, we applied the following cri-

<table>
<thead>
<tr>
<th>CRITERIA</th>
<th>WEIGHT</th>
<th>Adequate</th>
<th>Questionable</th>
<th>Not applicable</th>
<th>FOR OFFICE USE ONLY</th>
</tr>
</thead>
<tbody>
<tr>
<td>Therapeutic relationship 1. Provides positive feedback, reinforces gains</td>
<td>5.00</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>41</td>
</tr>
<tr>
<td>2. Encourages ventilation of feelings</td>
<td>4.75</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>45</td>
</tr>
<tr>
<td>Assessment 3. Functional level</td>
<td>5.00</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>49</td>
</tr>
<tr>
<td>4. Mental status</td>
<td>4.75</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>53</td>
</tr>
<tr>
<td>5. Medical history and physical status</td>
<td>4.25</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>57</td>
</tr>
<tr>
<td>6. Family and social history</td>
<td>4.75</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>61</td>
</tr>
<tr>
<td>7. Functional and work history</td>
<td>4.75</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>65</td>
</tr>
<tr>
<td>8. Ongoing reassessment and documentation on change</td>
<td>4.25</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>69</td>
</tr>
<tr>
<td>Treatment planning 9. Determines achievable goals</td>
<td>4.25</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>73</td>
</tr>
<tr>
<td>10. Seeks patient consensus on goals</td>
<td>4.25</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>77</td>
</tr>
<tr>
<td>Treatment 11. Monitors medical care (medications, appointments)</td>
<td>3.75</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>7</td>
</tr>
<tr>
<td>12. Promotes discussion of social and interpersonal situations</td>
<td>4.50</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>11</td>
</tr>
<tr>
<td>13. Promotes social network</td>
<td>4.25</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>15</td>
</tr>
<tr>
<td>14. Teaches self-care skills</td>
<td>4.50</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>19</td>
</tr>
<tr>
<td>15. Teaches household management and budgeting</td>
<td>4.00</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>23</td>
</tr>
<tr>
<td>16. Teaches organizational skills</td>
<td>3.00</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>27</td>
</tr>
<tr>
<td>17. Teaches leisure time planning</td>
<td>4.00</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>31</td>
</tr>
<tr>
<td>18. Teaches stress management (relaxation, physical exercise, time management, etc.)</td>
<td>4.25</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>35</td>
</tr>
<tr>
<td>19. Provides parent effectiveness training</td>
<td>4.00</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>39</td>
</tr>
<tr>
<td>20. Provides assertiveness training</td>
<td>3.50</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>43</td>
</tr>
<tr>
<td>21. Provides vocational counseling</td>
<td>3.75</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>47</td>
</tr>
<tr>
<td>22. Provides counseling with regard to insight</td>
<td>3.25</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>51</td>
</tr>
<tr>
<td>23. Provides grief counseling</td>
<td>3.25</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>55</td>
</tr>
<tr>
<td>24. Involves family and significant others</td>
<td>4.00</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>59</td>
</tr>
<tr>
<td>Consultation and follow-up 25. Referral made to other agencies as related to treatment goals</td>
<td>3.75</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>63</td>
</tr>
<tr>
<td>26. Contacts maintained with other involved professionals</td>
<td>4.25</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>67</td>
</tr>
<tr>
<td>27. Reports serious concerns to appropriate source</td>
<td>4.75</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>71</td>
</tr>
<tr>
<td>28. Consults with other occupational therapists for supervision when necessary</td>
<td>3.25</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>75</td>
</tr>
<tr>
<td>29. Makes plans for ongoing follow-up on discharge</td>
<td>3.75</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>79</td>
</tr>
</tbody>
</table>
criteria to define the population: a) subjects must have a primary diagnosis of depression, b) subjects must be referred for intervention primarily related to depression, c) subjects must be 18 to 65 years old (this excludes subjects with specific syndromes associated with adolescence or old age), and d) subjects must not have given birth in the six months before referral (this excludes subjects with a postpartum depressive syndrome). These specifications yielded a population of approximately 150 charts, of which a random sample of 35 was chosen for scoring.

**Scoring.** The sample was then scored to determine adequacy of intervention. This was accomplished by reading the chart in search of documented evidence of the occurrence of each criterion as a part of treatment. If such evidence was found, the particular criterion was scored "adequate" and was given one point. For criteria 1 through 13, the absence of documentary evidence for a criterion meant that it was scored as zero or "questionable." The committee declared these first 13 criteria to be necessary in all cases, and therefore the option of "not applicable" was not available. However, for criteria 14 through 29, a rating of "not applicable" could be given where no evidence of the intervention was found and no reference to it had been made in the referral or the treatment plan. These "not applicable" criteria were not used to contribute to the total score.

The total score for each chart was calculated by multiplying the criterion scores (0 or 1) by the weight and then summing over all applicable criteria and converting to a percent. This total score represents the percentage of applicable criteria with which the therapist complied. However, the unit of analysis for quality assurance is not the individual therapist, but the agency. Therefore, in addition to scores for individual therapists, cumulative scores were calculated a) for each criterion, b) for each section (e.g., therapeutic relationship, assessment, treatment planning), and c) for the total scores. These cumulative scores are expressed as percentage of therapists sampled who complied with the criterion or criteria.

A normative standard of 85% compliance with the criteria was set for acceptability of the service (3, 5). To account for sample size in the pilot study, we constructed a one-sided 95% confidence interval around the standard (5), thus reducing the threshold score for acceptability to 75%. The standard was applied in two ways: first, individual charts in which there was less than 75% compliance with the criteria were considered unacceptable, and second, individual criteria or sections where fewer than 75% of therapists complied were identified as problem areas. An overall measure of acceptability of the service was obtained by determining if at least 75% of charts exceeded the overall standard for acceptability.

**Reliability.** For the purpose of the pilot study, the chart review was conducted exclusively by the researcher; thus, the consistent application of the criteria was ensured. However, for the future use of the score sheet, test-retest reliability and intrarater reliability were estimated (using a Pearson r correlation) at .734 and .728, respectively. These coefficients reflect the fact that a degree of interpretation was required in applying some of the criteria to written records.

**Validity.** Validity, in a study such as this one, refers to the extent to which the criteria used actually reflect the quality of care delivered. The literature recommends two types of validation for use with process criteria. Donabedian (11) suggests normative validity, which is achieved by using both a consensus-seeking process and professional experts when establishing criteria. These features of normative validity were used in the present study. Senior mental health consultants from within the agency and outside of it were included on the criteria committee; and a consensus-seeking procedure (namely the nominal group process) was used for all aspects of criteria development.

A second validation procedure was undertaken to assert criterion validity. Because no "gold standard" is currently available for the assessment of quality in community mental health care, it was necessary to use professional judgment as the criterion. Ten additional charts were chosen at random from the same population. The charts were then ranked ordered by a) a panel of therapists, who used their professional judgment, and b) the investigator, who used the scoring procedure. The median rank given by the panel of therapists was compared with the rank based on the score sheet. A Wilcoxon signed-rank test showed no significant difference between paired rankings (p < .05), thereby contributing to the assertion of criterion validity for the new instrument.

**Analysis of Deviations**

The results of the pilot application are of general interest only
because they have methodological implications for future study. Usually the results of quality assurance studies are intended to be used only internally by the agency to identify problems and direct remedial action. Therefore, while the detailed results are not presented, they are discussed to the extent that they recommend improvements for future study.

Criteria refinement. The first recommendation is related to several criteria (e.g., criteria 1, 2, and 10) that seldom appeared as such on the chart. To apply these criteria, it was necessary to undertake an extensive interpretation of the documented materials. In some cases, circumstantial evidence led the rater to give a positive score on these criteria; however, in most cases, these criteria were scored negatively because no documentary evidence was available. As a result, these criteria probably received a score that was much lower than the level of performance. Clinicians studying these scores will recognize how unlikely it is that an intervention such as provision of feedback would be recorded on the chart. In a therapeutic relationship, interactions such as providing verbal or nonverbal reinforcement are seldom recorded, and therefore they are difficult to score by this method.

For future applications of the model, we suggest three strategies for overcoming this problem of subjectivity in the criteria. First, the criteria in question could be refined to more objectively represent the desired qualities. Examples of objective behaviors indicative of compliance with the criteria could be included in the criteria definitions. In this way, raters would be alerted to specific indications of developments in the therapeutic relationship, and they would not have to search for somewhat nebulous evidence. Therapists have suggested the following indicators: the fact that the therapist was allowed back into the home on subsequent visits, the fact that the patient kept appointments, and the fact that personal or sensitive material was discussed. However these are indirect indicators that could represent many other issues besides the development of a therapeutic relationship.

As a second alternative, the monthly documentation audit (mentioned earlier) might be used to overcome the problem. If renewed emphasis were placed on the recording of abstract, as well as concrete, aspects of care, then perhaps therapists would be more likely to document abstract care components (e.g., the therapeutic relationship and patient consensus). Although the documentation audit ensures a high level of uniformity and completeness of records, it does not specifically address the abstract component of care. Emphasis on recording interpersonal developments between patient and therapist would help to ease this problem for the future use of the model.

Finally, the problem connected with recording this kind of information may lie not in the criteria but in the method itself. Perhaps the process chart audit is not the method of choice for assessing the "therapeutic relationship" aspect of total care. While we continue to believe that the process audit is the best method for the study as a whole, another method might be more appropriate for assessing the therapeutic relationship. As a third possible strategy for overcoming the problem of criterion subjectivity, we recommend one of the observational or direct data collection methods (e.g., participant observation, interview, or questionnaire), which might more satisfactorily uncover this type of information.

Compliance rate. The use of the 85% compliance rate, adjusted for sample size by application of the 95% confidence interval, was found to be a useful and sensitive standard. Although we have no empirical evidence to support the ability of the current standard to detect adequate versus inadequate care, we feel that the standard was effective. The absence of any empirically valid criterion for the assessment of community occupational therapy makes further validation of the standard impossible at this time.

Committee structure and function. The committee structure and format (i.e., the nominal group process) were found to be effective and expedient. The committee offered an extremely satisfactory level of expertise, and the nominal group process eliminated much of the lengthy discussion required for more traditional forms of consensus. Therefore, we recommend that a similar procedure be used in future quality assurance efforts.

Conclusion

The development of quality assurance based on this model could ultimately take two routes. First, future endeavors could be aimed at the evolution of a collection of criterion lists, each with a specific population or topic in mind. Alternatively, future research might be aimed at the development of a generalized or "generic" list of criteria that could have application to the
study of quality assurance in all fields of community occupational therapy. Based on the experience of other professions, we recommend the latter course. The proliferation of multiple criterion lists for all imaginable topics threatens to impose a "recipe" mentality on the practice of community occupational therapy. This trend would diminish the vital element of case-specific judgment, and in so doing, could undermine two very important concepts in occupational therapy: the holistic approach to individual, patient-centered care and the use of judgment based on professional knowledge rather than on the technocratic application of treatment modalities. We caution future researchers against a "technologist" mentality in applying quality assurance. We encourage the development of quality assurance in a direction that preserves the creative, holistic, and patient-oriented perspective that is the basis of occupational therapy.

ACKNOWLEDGMENTS

The authors thank the Research and Program Development Fund at Community Occupational Therapy Associates for financial support. Dr. J. I. Williams and the Health Care Research Unit, University of Toronto for technical assistance, and the criteria development committee for providing clinical expertise.

REFERENCES