Toward a Model for Testing the Relationship Between Quality of Care and Costs

Barbara Dickey1* and Sharon-Lise T. Normand2

1PhD, Associate Professor, Harvard Medical School, Cambridge, MA, USA
2PhD, Associate Professor, Harvard Medical School, Harvard School of Public Health, Boston, MA, USA

Abstract

Background: In mental health services research there is little empirical evidence to support speculation about the relationship between costs and quality.

Aims of the Study: The purpose of this paper is to provide a conceptual model and test its usefulness in determining the cost of care that meets evidence-based standards.

Methods: A case study of individuals treated for an acute episode of schizophrenia is described and a conceptual model for determining the costs of evidence-based care is presented. Statistical tests of difference were used to compare two groups, those with care that met guideline standards and those that did not.

Results: Compared to care that did not meet recommendations, evidence-based care was cheaper. Clinical benefits to patients were the same, but those with poor care (higher than recommended doses of anti-psychotic medication) had higher treatment costs as well as more side-effects.

Discussion: The conceptual model faces many challenges in application, but shows promise as one approach to determining the cost of evidence-based care.

Implications for Health Care Provision and Use: The results support efforts to encourage clinicians to follow practice guidelines.

Implications for Health Policies: The growth of evidence-based medicine must be matched by efforts to assess the costs of adherence to practice guidelines. The pressure of fiscal restraints needs to be balanced with information about what the cost will be to provide recommended treatment.

Implications for Further Research: The greatest effort needs to be in the conceptual development of the model so that we can confidently estimate the costs and effectiveness of evidence-based treatment recommendations.

Received 7 July 2003; accepted 27 January 2004

Background

Two themes appear, usually intertwined, with increasing regularity in published health services research reports: first, the quality of care as measured by evidence-based guidelines and second, efforts to contain the rising costs of health care. The confluence of these two themes appears in the question that is on the minds of many patients and their clinicians: are fiscal pressures to reduce costs compromising the quality of care?

It is commonly believed that we get what we pay for, so, it has been argued, the best care is likely to be more expensive than poor care. A parallel proposition is that more treatment (or more expensive treatment) is equivalent to better care. The root of concern among patients and their clinicians is that practice guidelines have been developed by some managed care companies to justify denial of treatment. However, the introduction of evidence-based medicine in recent years has been an effort to set standards of quality based on research of what treatments work best, not to limit medical spending. Evidence-based standards are not isometric with better quality of care, but it is expected that treatment that adheres to standards will lead to better outcomes. In fact, a recent article in the New York Times (Ableson1) reported evidence that the current Medicare payment system led to perverse incentives: some hospitals employing evidence-based treatments suffered reduced revenues because physicians were able to limit impatient care to only the very sickest patients. This resulted in expenditures above the Medicare episode fixed payment.

In considering how to determine the nature of the relationship between quality and cost, we found that the work of McNeil2 best fit the conceptual approach that we present in this paper. McNeil offered a framework for testing whether poor care or better care cost more, discussing how the twin problems of uncertainty in medical decision-making and cost containment efforts affect service use. Uncertainty has been shown to lead to overuse of services that may raise costs without raising the benefit to the patient. On the other hand, cost containment, especially if it includes financial incentives that reward efficiency, might lead to the under-use of needed treatments.

The small area analyses of John Wennburg’s group3,4 at Dartmouth has provided empirical support for the notion that the best care is also less expensive. Their work illustrates wide variation in the use of certain medical procedures, and cites evidence that those who need more medical attention were treated more cheaply. In a recent New York Times
editorial, Fisher makes this point when discussing the “flawed assumption” that has been at the heart of the Medicare debate in Congress: “more care – and more expensive care – is better care.”

Much of the discussion around quality and cost has been prompted by advocates of the seriously mentally ill who argue that they have been discriminated against as a result of insurance policies that do not provide mental health benefits equivalent to medical benefits, thus compromising care (Frank et al.). Although advocates have argued their case on the basis of unwarranted discrimination, economists and policy makers have resisted parity because without limitations on mental health benefits, it is thought that there will be an overuse of treatments and rising costs, especially for psychotherapy (McGuire, Frank et al.) without an equivalent rise in benefit. The argument about parity has not been played out within the framework of evidence-based guidelines.

In mental health services research there is little empirical evidence to support speculation about the relationship of costs and quality. A recently published study of the quality of care in the Tennessee Medicaid program (Ray et al.) reported that after a specialty carve-out managed care plan was introduced, the use of antipsychotic medication was reduced, especially in high risk patients. Antipsychotic medication is one of the evidence-based treatments for schizophrenia. Few studies have placed a “price” on achieving even minimally adequate mental health treatment, however that is defined. One study by Normand et al. has estimated costs for different depression treatment modalities based on the probability of full or partial symptom reduction. The investigators used a modeling approach based on clinicians’ estimates of how well specific treatments “work” with specific patient types. The study provides outcome-adjusted price per full remission episode.

Aims of the Study

The purpose of this paper is to provide a conceptual model and test its usefulness for determining whether care that meets evidence-based standards results in higher or lower per person expenditures than care which does not. We use the framework of under use or overuse as problems in the delivery of care through a case study to assess the quality of treatment and to guide interpretation of treatment costs. Second, we discuss some challenges that need to be overcome in further testing of the utility of our model.

Methods

Case Study

From a prospective observational study of patients treated for schizophrenia (Dickey et al.), we selected for our case study only those who were treated on an outpatient basis, had no psychiatric inpatient episodes during the 6 month study period, and whose treatment could be classified as meeting guideline recommendations or not. Entry into the original study was triggered by a visit to one of 8 psychiatric emergency screening teams (EST) in Massachusetts that operated on a “walk-in” basis for anyone in crisis, regardless of insurance status.

Conceptual Model

Define Quality in Operational Terms that can be Applied to Available Treatment Data. What do we Mean by “Quality of Care?”

We defined it as a set of evidence-based practice guidelines (Lehman et al.,) for the treatment of schizophrenia. As currently written, the guidelines are quite specific as to who should be prescribed antipsychotic medication (everyone with the disorder) and the dose range for the acute (300-1000CPZ units) and the maintenance (300-600 CPZ units) phase. In addition to specific medication recommendations, the practice guidelines include more general recommendations for psychosocial treatment (various psychotherapies), vocational rehabilitation, case-management, substance abuse treatment if indicated and supportive services. Except for medication, the recommendations are not specific as to amount or duration of the recommended treatment needed.

Although it is common practice to use published practice guidelines to measure quality, the authors of the guidelines we used described the recommendations as “adequate” care, not a mechanism for determining levels of “quality” above the merely adequate. One of the conceptual challenges that we faced was grouping study participants into categories of “better care” and “poorer care.” We chose two indicators of “better care” (medication within the acute dose range recommended and any evidence of substance abuse treatment) and then defined “poorer care” as individuals who got neither treatment. Without more specific recommendations about psychosocial treatments, we chose, using this case study, to limit our definition of adequate treatment to guidelines that were well defined and likely to have the largest impact on overall costs. We did not include individuals with inpatient treatment during the study period because a substantial number of individuals in the “better care” had inpatient treatment did not meet the guidelines with respect to medication dose and substance abuse treatment while at the same time those without adequate care as outpatients sometimes had care that met the guidelines for inpatient treatment.

Define Quality in Operational Terms that can be Applied to Available Treatment Data. What do we Mean by “Quality of Care?”

This is really an extension of the earlier discussion. Incremental levels of “quality” would need to define both the specific combination of treatments received and the timing of those treatments. To date, guidelines have not reached this level of specificity although some medication algorithms provide specific recommendations for treatment

B. Dickey et al.
non-responders. More refined cost estimates would need to characterize care that meets some but not all the guidelines. The treatment of schizophrenia is not limited to medication. A range of psychosocial treatments and case-management are also recommended. To include these, we would have had to identify the most potent “ingredients” in this set of treatment guidelines and then prioritize each guideline by potency and the amount of additional benefit it confers, if used. Today, most guidelines for psychosocial treatment are contingent on specific need and preference, and rarely specify volume or intensity of treatment. For example, for a group of treated individuals, what is the level of need for family therapy, the level of patient preference for family therapy, and most important, what is the level of benefit conferred by this treatment, over and above the benefit conferred by other treatments. If we were able to determine this for all psychosocial treatments, then instead of two quality categories, we could have created an ordinal or continuous variable based on the amount of evidenced-based treatment that a person received. Given the current state of knowledge, it is likely to be some time before we are able to reach this more sophisticated and useful stage of analysis. Our two-category approach, while accurate for purposes of the case example offered, leaves us with a small sample that is under-powered statistically. It also has obvious drawbacks from a policy-development perspective.

Define Episode of Care With Clear Beginning and End Points that are Both Clinically and Operationally Interpretable within the Context of Chronic Disease

Beginning and end dates need to be determined and the time frame should represent an episode of care or some other clinically meaningful timeframe. We had a clearly defined starting point for the acute episode, in this case a visit to a psychiatric emergency screening site. We chose 6 months as the length of an acute episode, a choice suggested by the practice guideline for antipsychotic medication duration at the higher (acute phase) level. Although we chose to use 6 months for everyone in this study, acute episodes vary in length and more precise clinical specificity would be desirable to account for variation in length of time the episode lasts.

Identify Data Sources that will Support the Research

Quality of care studies are often conducted using medical records as the source of information about treatment. Administrative data, such as paid benefit claims, have the advantage of large population, but usually lack sufficient data to determine daily medication dose. It is possible to use patient reported treatments within specified time periods, but the accuracy of these data have been questioned by some researchers. Outpatient medical records were the source of data for antipsychotic medication dose and paid claims were used to identify other treatment received and per person treatment expenditures.

Determine the Cost Perspective

Economists argue for taking the social perspective, that is the costs to society should be added to the costs of care. While we agree that this is, in the purest economic sense, the right method if the final step were to be a cost-benefit ratio intended to inform public policy. However, in quality of care studies, there is little debate about whether patients should receive adequate care — the only question is whether we should expect to pay more or less for it. To provide a test for this model, we used only paid treatment claims (i.e., accounting costs), which reflect expenditures by certain agencies in the state and federal government. It is not unreasonable to assume that differences in treatment regimes may lead to costs outside this narrow accounting cost definition we chose. For non-governmental insurance plans, co-payments and other out-of-pocket expenditures are important costs to capture. A more comprehensive approach would include the costs to human service agencies, the patient and their families, and the benefits of employment and improved well-being.

Specify Outcomes that are Linked Directly to Treatment Recommendations and Include Unwelcome Outcomes (e.g., Side Effects) as well as Benefits

If we take seriously Donebedian's conceptual model of quality, then our treatment process measures must be linked to patient outcomes. Although the model we propose could be limited to just treatments and their costs, the ideal measure of quality of care (Donebedian) includes outcomes of treatment. For studies of the quality of care, outcome measures, ideally, should measure treatment response — those symptoms and behaviors that are expected to change as a result of treatment and, importantly, which can be adequately adjusted when influenced by other conditions, including socio-demographic characteristics.

Use Case Mix Adjustments for Outcomes but Not on the Process Measures of Treatment (the Quality of Care Measure)

It is common to use some case-mix adjustment when we undertake observational studies of health outcomes, but in a study of quality, our question is about adherence to practice guidelines. The guidelines are silent as to variation in adherence associated with age, gender, race, living situation, health plan or economic status. In reporting the likelihood of receiving certain recommended treatments we do not need to adjust for these characteristics.

However, baseline clinical information and socio-demographic data are important if we are to have confidence in the outcome results. In the case of this report, it is certainly plausible to ask whether the “poorer” case group are actually much sicker and thus require higher levels of medication and more services. Our baseline clinical measures of symptoms and behaviors, drawn from the emergency room medical record data, showed clinical equivalency. But more importantly, even if the group with poorer care were sicker, strong research evidence supports the recommendation that dose not exceed 1000CPZ units because higher doses do not improve outcomes.
Sample

In the case study we use as a preliminary test of the model, subjects were English or Spanish speaking adults, age 24-64, disabled Medicaid beneficiaries who received government disability income support, diagnosed with schizophrenia or schizoaffective disorder by a clinician at the psychiatric emergency screening team site and verified through one other source. We excluded those who also had an inpatient episodes, those who were not prescribed an antipsychotic medication or received any other psychiatric treatment during the study period. In the case study we use as a preliminary test of the model, subjects were English or Spanish speaking adults, age 24-64, disabled Medicaid beneficiaries who received government disability income support, diagnosed with schizophrenia or schizoaffective disorder by a clinician at the psychiatric emergency screening team site and verified through one other source. From the original 420 study participants, we excluded 17 study participants who had no record of treatment at all post EST visit (which is the worst but obviously cheapest care), 8 outpatients whose records could not be found, 187 who had inpatient admissions during the study period as well as outpatient treatment and 16 whose medication dose level was below the recommended. These low dose patients were not included because the wording of the guideline suggests that dose lower than 300CPZ units per day is acceptable if it provides clinical benefit. These exclusions were made to ensure that we could interpret our findings. All participants gave consent to participate in the study and were paid a nominal fee for being interviewed.

Evidence-based Guidelines

Quality of care was measured as adherence to evidence-based guidelines, the Schizophrenia PORT treatment recommendations, as documented in their medical records during a 6-month study period. These PORT recommended treatments are considered to be minimally adequate for adults with schizophrenia and were not designed to assess relative increments in quality above this minimum level.

We categorized treatment in such a way as to select those (N=40) as having “better care” whose mean monthly medication dose was within the guideline range (300-1000CPZ units) for the acute phase of the illness over the 6 month study period and who also received treatment for substance abuse if that problem was noted in the record. We classified as “poorer care” (N=51) those who received antipsychotic medication above the recommended dose (>1000CPZ units) and who did not get substance abuse treatment if substance abuse was noted as a problem in the record. Almost everyone in the study received some form of psychosocial treatment at least once and including receipt of this treatment would not have provided a mechanism for differentiating care between groups.

Data

Costs

We chose to use a narrow definition of “costs”: those health benefits paid for by state and federal government through the Medicare and Medicaid programs. Paid benefit claims were used to determine reimbursed psychiatric and substance use disorder treatment, reported here as mean expenditures. These costs are likely to be lower bound estimates and not the true economic costs, which might include such costs as patient and family burden, use of other state agencies, such as the criminal justice system and the value of benefits such as employment or education. We summarize mental health

Table 1. Sociodemographic Characteristics

<table>
<thead>
<tr>
<th>Patient characteristics</th>
<th>Better care (N = 40)</th>
<th>Poorer care (N = 51)</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Females</td>
<td>19 (48%)</td>
<td>23 (45%)</td>
<td>.82</td>
</tr>
<tr>
<td>Males</td>
<td>21 (53%)</td>
<td>28 (55%)</td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;35</td>
<td>12 (30%)</td>
<td>11 (22%)</td>
<td>.84</td>
</tr>
<tr>
<td>35-44</td>
<td>19 (48%)</td>
<td>27 (53%)</td>
<td></td>
</tr>
<tr>
<td>45-54</td>
<td>7 (18%)</td>
<td>10 (20%)</td>
<td></td>
</tr>
<tr>
<td>55-64</td>
<td>2 (5%)</td>
<td>3 (6%)</td>
<td></td>
</tr>
<tr>
<td>Race</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>African-American</td>
<td>10 (25%)</td>
<td>9 (18%)</td>
<td>.16</td>
</tr>
<tr>
<td>White</td>
<td>28 (70%)</td>
<td>33 (65%)</td>
<td></td>
</tr>
<tr>
<td>Other*</td>
<td>2 (5%)</td>
<td>9 (18%)</td>
<td></td>
</tr>
</tbody>
</table>

* Latino, Asian
expenditures in several categories: crisis or emergency treatment, clinic visits, community support services and substance use treatment.

Mean pharmacy and medical expenditures are reported separately and not included in the cost of mental health treatment. We report mean costs in two ways: first, by user when the denominator is the number of people who used the particular service; second, by person when the denominator is the total number of people in the group. The latter permits standardized comparisons within service-type between groups. Because of the small sample size we do not provide tests of difference, but include the standard deviation and the median cost per group.

Guideline Adherence

We used the medication dose recorded in the medical record to determine if the daily medication dosage met the guideline for acute episodes (300 – 1000 Chlorpromazine [CPZ] equivalents). This was calculated by converting the antipsychotic medication dose into CPZ units, multiplying the units times the daily frequency and, if more than one was prescribed, summing the units. The CPZ equivalencies used were provided by the Schizophrenia PORT guideline medical record abstraction protocol. Atypical antipsychotic doses were not converted to CPZ equivalents but were determined to be within range as follows: olanzapine, 5-20mg/day; quetiapine, 150-750mg/day; and risperidone, 2-6mg/day. Because not every patient in the study had medication prescribed regularly every month, we had to find methods for estimating whether, on average, the dose was within range. We used a mean monthly standardized dose in this report.

Evidence of substance abuse treatment was gathered from the records and paid claims. Attendance at AA (or similar) meetings was included as treatment data.

Outcomes

We obtained data on health status through patient interviews. The health status measures we used were self-reported health-related quality of life (Ware et al.15), psychiatric and substance use problems (Eisen et al.16) and medication side-effects reported in the medical record.

Data Analytic Procedures

Only the group socio-demographic and outcome data are compared statistically. Descriptive summaries are provided for the cost data. No case-mix adjustments were made because the sample size is small and groups were not significantly different in their socio-demographic characteristics.

Results

We have reported elsewhere (Dickey et al.17) that the patients in the original larger study can be fairly characterized as especially vulnerable high-risk individuals who have serious disabilities and often poor connections with the mental health system. In this subsample, there are no gender, race or age statistically significant differences between the better and poorer care groups (See Table 1).

We found that dually diagnosed adults with the best care had mean per-person unadjusted 6-month psychiatric treatment expenditures of $9,232 (median, $4,690; sd, $7,899); the group with worst care had higher psychiatric unadjusted treatment expenditures: $11,819 (median, $8,779; sd, $6,973). Table 2 summarizes the expenditures by type of treatment received and both the per-user mean cost and the per-person mean cost. Adults with the worst care spent more time in crisis beds or emergency services and used fewer

---

Table 2. Psychiatric and Other Medical Expenditures 6 Months Post Index Visit

<table>
<thead>
<tr>
<th>Mental Health Expenditures¹</th>
<th>Better care N = 40</th>
<th>Poorer care N = 51</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N (%)</td>
<td>Mean per user</td>
</tr>
<tr>
<td>Crisis beds³</td>
<td>14 (35%)</td>
<td>7,005</td>
</tr>
<tr>
<td>Clinic visits</td>
<td>38 (95%)</td>
<td>2,227</td>
</tr>
<tr>
<td>Support services⁴</td>
<td>32 (80%)</td>
<td>5,829</td>
</tr>
<tr>
<td>Total Mental Health</td>
<td>40 (100%)</td>
<td>9,232</td>
</tr>
<tr>
<td>Pharmacy⁵</td>
<td>40 (100%)</td>
<td>2,427</td>
</tr>
<tr>
<td>Medical care⁶</td>
<td>39 (98%)</td>
<td>3,090</td>
</tr>
</tbody>
</table>

¹ For the best group, expenditures include the cost of substance abuse treatment
² mean per person expenditures are calculated using the group N as the denominator, whereas the user mean is calculated using the number of users as the denominator
³ Crisis beds are an alternative to hospitalization
⁴ Support services are include group housing, transportation and other community support services
⁵ Pharmacy expenditures include all medications, not just antipsychotics
⁶ Medical care includes all treatment for disorders other than mental illness or substance use disorder.
clinic visits compared to those with the best care. Those in the worst care category also had much higher pharmacy costs but their medical care expenditures were the same as those in the best care category.

We found that the scores on the two outcome measures we used, self-reported problems and mental health-related quality of life, were not statistically different, but the level of side-effects was higher for those on higher doses of medication. (Table 3)

### Discussion

The goal of this study was to present a conceptual model for determining the cost of adhering to treatment guidelines, which we defined as quality of care. We tested the model on a small case study and found that, in this preliminary test, per person psychiatric costs were higher among those with the care that did not meet the guidelines. Furthermore, the group whose treatment did not meet the guidelines had a pattern of care consistent with the overuse and underuse theory of poor care: fewer routine clinic visits and more emergency and crisis treatment. The data that we used did not provide us with information about what led to fewer outpatient visits and more emergency and crisis services, but it seems fair to consider the likelihood that untreated substance use disorder contributed, perhaps substantially, to higher costs. The differences we found cannot be attributed to managed care or benefit limitations – the Medicaid beneficiaries in this sample were not limited in the treatment they were entitled to receive. Engaging substance abusers in treatment is often difficult, but these data and other findings we have reported (Dickey et al.) lead us to conclude that there are considerable cost offsets to be gained by making every effort to provide such treatment.

The small sample prevents us from exploring whether dose is related to patient characteristics or provider practice. One possible explanation is that erratic treatment-seeking patterns by patients lead some physicians to prescribe higher doses in an effort to bring psychotic symptoms under control, but the evidence here is that it does not result in better outcomes and instead only leads to more side-effects.

The cost findings are not the center of the paper, however, although they are intriguing. The purpose of the paper was to test the conceptual model for answering questions about the cost of care that meets recommended standards. Our case example illustrates the many challenges that face investigators seeking to advance this type of work. One limitation of any test of guideline adherence is that medical records only record what treatments were prescribed and the response to that treatment. In this case, we do not know if the patients took some or all of the medication nor do we know if the substance abuse treatment was successful. Patient adherence to treatment regimens opens up a field of research not included in our model, but one that bears investigation. A second limitation, which all work like this faces, is the limitations of medical records. They do not always include all the information that is important, handwriting is poor, dates are sometimes missing and some charts could not be located even after several trips to the record room.

### Implications for Health Care Provision and Use

Our data do not extend to helping us fully understand what treatments are most effective for these patients, but it supports efforts to encourage clinicians to follow practice guidelines, which in this case, appear to lead to relatively lower healthcare expenditures with equivalent benefits and fewer side-effects.

### Implications for Health Policies

It is too early to tell whether evidence-based medicine will contribute to cost-containment, but policy makers should be demanding studies of the type described in this report to shed light on whether cost and quality can be synchronous instead of a series of trade-offs. The growth of evidence-based medicine must be matched by efforts to assess the costs of adherence to practice guidelines. The pressure of fiscal restraints needs to be balanced with information about what the cost will be to provide recommended treatment.

Cost-effectiveness and cost-benefit studies come to mind as a research models for answering this question but these...
studies typically compare two policy choices and identify one with a relatively better cost/effectiveness or cost/benefit ratio. However, when the discussion is about whether better quality of care will be more or less costly than poor care, the discussants are considering quality as an absolute, not as a choice between high and low quality.

**Implications for Further Research**

The model we outline here is one step in the development of cost estimates, but much work remains to be done. Despite the challenges of this work, we believe that model development is essential for work to move forward on cost estimates of evidence-based treatment. Perhaps the greatest effort needs to be in refining the model we propose so that we can confidently group more patients into discrete levels of ‘‘quality.’’

We found that assessing mean dose level over time proved to be one of the most difficult challenges. During the study period almost every patient had dose levels that changed from time to time, that the number (and class) of antipsychotic medications prescribed changed and that the use of depot and oral medications separately and in combination was common.

Once these methodological problems are resolved it will be easier for other researchers to carry out studies with larger sample sizes. This will both provide more statistical power and also make the cost estimates more stable. Unfortunately, larger samples make the collection of data to assess adherence to guidelines more difficult – or at least more expensive if those data come from medical records. This leads to the suggestion that administrative data might yield sufficient information under the right circumstances, especially if health systems increase the use of electronic medical records that can be abstracted for research purposes. Access to administrative data is not always easy, but collaboration with health care providers in through continuous quality improvement projects may provide an introduction to IT mechanisms that could support research.

Finally, for the field to advance, measures of interpersonal care, as described by Donebedian\(^1\) must be joined with measures of technical adequacy. We do not know, but can speculate that stronger ties between patient and doctor lead to greater trust and confidence on the part of patients who collaborate more closely with their doctor and, one hopes, has better outcomes as a result.

**References**