This product is part of the RAND Corporation monograph series. RAND monographs present major research findings that address the challenges facing the public and private sectors. All RAND monographs undergo rigorous peer review to ensure high standards for research quality and objectivity.
Assessing the Performance of Military Treatment Facilities

Nancy Nicosia, Barbara O. Wynn, John A. Romley

Prepared for the Office of the Secretary of Defense
Approved for public release; distribution unlimited

Center for Military Health Policy Research
The research described in this report was sponsored by the Assistant Secretary of Defense for Health Affairs and conducted by the RAND Center for Military Health Policy Research, a joint project of RAND Health, and the Forces and Resources Policy Center of the RAND National Defense Research Institute. The latter is a federally funded research and development center sponsored by the Office of the Secretary of Defense, the Joint Staff, the Unified Combatant Commands, the Department of the Navy, the Marine Corps, the defense agencies, and the defense Intelligence Community.

Library of Congress Cataloging-in-Publication Data
Assessing the performance of military treatment facilities / Nancy Nicosia ... [et. al].
   p. cm.
   Includes bibliographical references.

UH223.A92 2011
355.7’20684—dc23
2011020359

The RAND Corporation is a nonprofit institution that helps improve policy and decisionmaking through research and analysis. RAND’s publications do not necessarily reflect the opinions of its research clients and sponsors.

RAND® is a registered trademark.

© Copyright 2011 RAND Corporation

Permission is given to duplicate this document for personal use only, as long as it is unaltered and complete. Copies may not be duplicated for commercial purposes. Unauthorized posting of RAND documents to a non-RAND website is prohibited. RAND documents are protected under copyright law. For information on reprint and linking permissions, please visit the RAND permissions page (http://www.rand.org/publications/permissions.html).

Published 2011 by the RAND Corporation
1776 Main Street, P.O. Box 2138, Santa Monica, CA 90407-2138
1200 South Hayes Street, Arlington, VA 22202-5050
4570 Fifth Avenue, Suite 600, Pittsburgh, PA 15213-2665
RAND URL: http://www.rand.org

To order RAND documents or to obtain additional information, contact
Distribution Services: Telephone: (310) 451-7002;
Fax: (310) 451-6915; Email: order@rand.org
The U.S. Department of Defense (DoD) has increasingly confronted financial, managerial, and operational challenges in sustaining the TRICARE health benefit, which it provided to 9.2 million beneficiaries in fiscal year (FY) 2006. Medical costs, for example, are projected to increase to 12 percent of DoD’s total budget as of FY 2015, from a level of 8 percent in FY 2007.

In response to such challenges, the 2006 Quadrennial Defense Review motivated a transformation in business practices within the Military Health System (MHS). Performance-based planning and financing would allocate resources based on the value of activities to DoD’s mission, while aligning accountability and authority within the system.

DoD has considered setting targets for health care utilization in its military treatment facilities (MTFs) and rewarding or penalizing MTFs according to their performance. Such an initiative supposes that MTF leaders are able to cost-effectively manage care, much as generalist physicians or managed-care plans are frequently expected to do in the private sector. For example, in areas in which TRICARE costs are high at private hospitals, MTF leaders may be able to encourage beneficiaries to be treated at military hospitals with spare capacity.

The Office of the Assistant Secretary of Defense for Health Affairs (OASD[HA]) has been monitoring utilization and costs “per member per month” (PMPM) among beneficiaries enrolled at each MTF in TRICARE Prime, a managed-care plan similar to a civilian health-maintenance organization. These PMPM metrics include all care received by beneficiaries, whether from the enrollment MTF, from
Assessing the Performance of Military Treatment Facilities

other MTFs, or from civilian health care providers. OASD(HA) has considered assessing each MTF’s performance by comparing current PMPM utilization with past levels.

Assessing changes in performance based on outcomes such as PMPM metrics raises a variety of important questions. What is the relationship between OASD(HA)’s metrics and MTF performance in cost-effectively managing care? What else may influence PMPM outcomes?

Figure S.1 suggests some answers. The figure shows OASD(HA)’s metric for inpatient utilization at DeWitt Army Community Hospital during FYs 2004–2005. Actual utilization in any quarter varies around the mean level. Performance may systematically influence mean utilization, yet there also appears to be some randomness.

Figure S.1 suggests some additional questions. If utilization were higher in FY 2006 than the FY 2004–2005 mean, how could OASD(HA) decide whether performance (or some other systematic

**Figure S.1**

*Actual and Mean Inpatient Utilization at DeWitt Army Community Hospital, FYs 2004–2005*
factor) had changed, or whether utilization just happened to be higher by chance? Is the nature of this decision concerning DeWitt, a relatively large MTF, similar to the decision that must be made at MTFs with small numbers of enrollees, where the randomness of utilization could be different? Do catastrophic cases, such as organ transplants, contribute to the random variability of inpatient utilization, making it harder to discern systematic changes?

**Purpose and Approach**

The purpose of this study is to help inform the sponsor’s thinking about the assessment of MTF performance in general and the variability of MTF PMPM utilization and costs in particular. In broad terms, the study included a qualitative review of performance assessment in the nonmilitary health care sector, as well as a quantitative analysis of the variability of the sponsor’s PMPM metrics and the roles played by MTF size and catastrophic cases.

For our qualitative review, we surveyed academic and policy research relating to performance assessment in health care. We visited a large Army hospital that served nearly 53,000 non-active duty Prime enrollees in FYs 2004–2005, where we interviewed MTF line administrators. We also conducted informal telephone interviews of experts in performance assessment at several private health care organizations.

This qualitative information helped guide the quantitative analyses, in which we were able to use two types of information:

- MTF-level data from FYs 2004 through 2006 on MHS-wide PMPM utilization and costs among TRICARE Prime beneficiaries enrolled at 114 “parent” facilities in the United States
- Disaggregate data for FY 2004 on admissions of Prime enrollees to military and civilian hospitals, as well as the personal characteristics of these beneficiaries.

---

1 As discussed below, some MTFs (such as small clinics) are “children” of “parent” facilities.
The analyses distinguished between inpatient, outpatient, and drug utilization. Active-duty personnel were excluded due to deployment-related data concerns.

We first analyzed MTF PMPM utilization and costs at both quarterly and annual frequencies. For each PMPM outcome at each MTF, we determined whether the change between FY 2006 and its mean level in FYs 2004–2005 was significant (as explained below). We then investigated the impact of an MTF’s size on the variability of its PMPM outcomes and the frequency of significant changes. We defined size as the mean number of non-active duty enrollees during FYs 2004–2005; in some analyses, we considered five groups of similarly sized MTFs. We also considered the role of trends across MTFs in PMPM outcomes.

Separately, we analyzed the role of catastrophic cases in MTF performance assessment based on hospital admissions. We defined admissions as catastrophic if their diagnosis groups were typically associated with high levels of resource use. We then explored the role that catastrophic admissions played in PMPM inpatient utilization during FY 2004. We also simulated the impact of excluding these admissions on the identification of significant changes in noncatastrophic inpatient utilization during FY 2006.

**Findings**

Our qualitative review of performance assessment in the nonmilitary health care sector indicates that a variety of factors systematically affect health care outcomes, including PMPM utilization and costs, costs per provider or clinical episode, and so on. The performance of health care managers is such a factor. In our context, MTF leaders cause more or less care to be provided and care to be delivered more or less efficiently. Thus, MTF outcomes may be useful measures of performance assessment.

Health status is another systematic determinant of health care outcomes, since those who are less healthy typically need and use more care than others. Practitioners and researchers frequently attempt to account
for health status by “risk adjusting” outcomes. Indeed, OASD(HA)’s PMPM metrics incorporate enrollee age, gender, and beneficiary status (e.g., retiree or dependent of a retiree). Such risk adjustments, while useful, are necessarily imperfect. When performance measures do not fully account for systematic factors, such as health risk or deployment of medical personnel, there can be substantial bias in assessments of MTF performance. The practical importance of this issue was beyond the scope of this study.

Utilization and costs also vary randomly. Whatever their health status, people use less care than usual in some periods and more in others. As a result, an observer cannot be certain about the true cause of a change in outcomes. On some occasions, an observer will mistakenly conclude that a change is systematic when in fact it is random (“false positives”). In other cases, the observer will conclude that a systematic change is random (“false negatives”). In reality, big changes are sometimes random noise, while small changes are sometimes meaningful.

An observer’s confidence that a change is truly systematic can be enhanced by requiring that an outcome increase (or decrease) by a large magnitude. When this threshold is exceeded, an observed change is “statistically significant.” A higher threshold for statistical significance results in fewer false positives, but more false negatives.

Given a confidence level, a lower rate of false negatives is desirable, because an observer has greater power to discern systematic changes. The false-negative rate is higher, however, when the randomness of an outcome is greater. PMPM utilization and costs may be more random at smaller MTFs, as there is less opportunity for enrollees’ random health care needs to balance out when there are fewer enrollees. Catastrophic cases may also contribute substantially to the randomness of PMPM outcomes.

Table S.1 highlights some important findings concerning the frequency of statistical changes during FY 2006 when MTF outcomes are analyzed at a quarterly frequency. We found similar patterns (though generally higher frequencies) in the annual analysis. For outpatient utilization, drug utilization, and total cost, the frequency of significant changes was lower for the smallest MTFs than for the largest ones. For total cost, for example, the frequencies were 20.7 percent and 42.0 per-
Changes in costs would be statistically significant in 5 percent of cases (given the 95 percent confidence level) even if there were no changes in the systematic determinants of outcomes. As a result, the share of significant changes in cost that are false positives could be as high as one in four (5%/20.7% = 24.2%) for the smallest MTFs, versus less than one in eight (5%/42.0% = 11.9%) for the largest ones. Unfortunately, the associated false-negative rates are unknown because the actual changes in performance and other systematic factors are unknown (though it would be possible to simulate these rates under various assumptions).

For inpatient utilization, the frequency of significant changes is actually lower at the largest MTFs. One possible explanation for this is that inpatient utilization became less variable at these MTFs. Among all MTF outcomes, the frequency of significant changes is lowest for inpatient utilization. While these outcomes were especially variable, the other outcomes tended to grow faster throughout the MHS in FY 2006, potentially making changes easier to discern. It is possible that such trends are partly attributable to changing performance across MTFs.

We also found that catastrophic cases, such as organ transplants and low-birthweight deliveries, play an outsized role in inpatient utilization. Diagnoses that ranked high in resource use accounted for a much larger share of utilization than of admissions. There is some reason to believe that excluding such cases would substantially increase the fre-

---

**Table S.1**

<table>
<thead>
<tr>
<th>MTF Outcome</th>
<th>Smallest MTFs</th>
<th>Largest MTFs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inpatient utilization</td>
<td>10.9%</td>
<td>3.4%</td>
</tr>
<tr>
<td>Outpatient utilization</td>
<td>19.6%</td>
<td>30.7%</td>
</tr>
<tr>
<td>Drug utilization</td>
<td>12.0%</td>
<td>15.9%</td>
</tr>
<tr>
<td>Total cost</td>
<td>20.7%</td>
<td>42.0%</td>
</tr>
</tbody>
</table>

NOTES: The smallest MTFs averaged no more than 7,187 non-active duty enrollees during FYs 2004–2005; the largest MTFs averaged at least 27,911. The confidence level is 95 percent.
quency of statistically significant changes in noncatastrophic inpatient utilization. It is possible, however, that MTF performance in managing catastrophic care is critical but hard to assess.

Altogether, our findings suggest that performance assessment of MTFs could be useful, though its effectiveness would generally be greater for larger facilities. Excluding catastrophic cases is practical and could be useful. In theory, systematic factors unrelated to performance could undermine the value of MTF outcomes as performance measures, and the practical importance of this issue may merit investigation. Finally, it is possible that alternatives, such as more targeted but complex assessments—for example, of cost per clinical episode—could help to diagnose MTF performance problems more reliably and to treat them more effectively.