Evidence for the Effect of Disease Management: Is $1 Billion a Year a Good Investment?

Soeren Mattke, MD, DSc; Michael Seid, PhD; and Sai Ma, PhD

In the face of double-digit healthcare inflation, evidence of systemwide poor healthcare quality, and an aging population, disease management seems an intuitively appealing way to improve the quality and reduce the cost of care, as well as to enhance health outcomes for the chronically ill. In broad terms, disease management refers to a system of coordinated healthcare interventions and communications to help patients address chronic disease and other health conditions. Commercial health plans and large employers are embracing this strategy, with 96% of the top 150 US payers offering some form of disease management service and with 83% of more than 500 major US employers using programs to help individuals manage their health conditions.

Public purchasers of healthcare services are testing the waters: The Centers for Medicare & Medicaid Services has launched a large Medicare demonstration to evaluate disease management, and several states are offering disease management programs under Medicaid. In addition, prominent politicians have voiced their optimism about the effect of disease management. On February 27, 2006, former US Senate Majority Leader Bill Frist, MD (R-Tenn), in a speech to the Detroit Economic Club, called for slowing the growth in federal Medicare expenditures. He cited chronic disease management as an effective means of controlling costs, estimating that “diabetic disease management in Medicare could conservatively save as much as $30 billion a year.”

Because his enthusiasm is shared widely, the disease management industry has grown rapidly, with estimated annual revenues increasing from about $78 million in 1997 to almost $1.2 billion in 2005 and projected to top $1.8 billion by 2008 (DM Purchasing Consortium, cited by Matheson et al). However, disease management has not been universally embraced. The Congressional Budget Office recently concluded that there is insufficient evidence that disease management reduces healthcare spending. A recent essay in the Times Argus called the economic value of disease management a “fantasy,” sparking a response from the executive director of the Disease Management Association of America that such a statement contradicts the peer-reviewed literature.

In light of the ongoing debate on disease management, we conducted a critical review of the empirical evidence regarding the effect of different types of disease management interventions on quality, cost, and health outcomes for various chronic conditions. Although several reviews and meta-analyses of disease management have been conducted in recent years, they
Evidence for the Effect of Disease Management: Is $1 Billion a Year a Good Investment?

... typically focused on particular diseases or end points. To bring some clarity to the debate on the effect of disease management, we set out to integrate the available body of knowledge and to make it accessible to decision makers by examining various types of disease management interventions for different conditions and the evidence available for its effects on a variety of end points. Before discussing the details of our study, we need first to say more about the meaning of the term disease management.

Part of the controversy surrounding the effect of disease management stems from the fact that a variety of programs and interventions can be labeled disease management. Not all types of disease management programs are receiving the same level of policy interest, nor are all types equally well researched. In addition, evaluations commonly have design flaws, limiting the validity of their conclusions.  

The Disease Management Association of America states that disease management is "a system of coordinated health care interventions and communications for populations with conditions in which patient self-care efforts are significant." The organization notes that a disease management intervention (1) supports the physician or practitioner/patient relationship and plan of care, (2) emphasizes prevention of exacerbations and complications utilizing evidence-based practice guidelines and patient empowerment strategies, and (3) evaluates clinical, humanistic, and economic outcomes on an ongoing basis with the goal of improving overall health.

Although the Disease Management Association of America definition provides a broad characterization of disease management programs, we need to refine this definition to review and categorize different types of disease management programs. We propose characterizing disease management programs along the following 2 dimensions: severity of illness among the target population and intensity of the intervention. The first dimension acknowledges that, although disease management programs have traditionally focused on patients at risk for chronic conditions and lifestyle modification or wellness programs for healthy individuals are increasingly being marketed as a supplement to disease management or as stand-alone offerings. The second dimension, intensity of the intervention, refers to the fact that disease management programs can vary widely from low-intensity interventions (which emphasize mass communication technologies such as mailings and prerecorded telephone messages) to moderate-intensity interventions (which include more direct individual contact such as telephone calls from call centers) to high-intensity intensive case management (which would include face-to-face encounters between patients and disease managers).  

Figure 1 shows how we can use these 2 dimensions to characterize the severity of illness in the target population in relation to the intensity of the disease management intervention. Most programs tend to lie on a diagonal through the grid, so that high-intensity programs are associated with high-risk patients and low-intensity programs with healthy participants or low-risk patients. Population-based programs that target the entire population with a disease trend to segment within that population and to use a mix of low-, medium-, and high-intensity interventions.

A third defining characteristic of disease management programs concerns the condition being addressed because the clinical course and baseline utilization patterns differ substantially across chronic conditions, creating a different set of challenges and opportunities for the intervention. For example, potentially avoidable high-cost events such as hospital admissions are common in patients with CHF, so a disease management program for this condition might seek to reduce the number of unnecessary high-cost treatments. On the other hand, patients with depression are commonly undertreated, so a disease management program might seek to increase levels of treatment. Therefore, although disease management will always strive to improve care and patient outcomes, whether it can save money by doing so may depend on the targeted disease. Finally, disease management programs vary by other important attributes such as the integration and involvement of the patient’s primary care physician and the presence of financial incentives.

Population-based disease management programs that identify and target all patients with a given condition irrespective of severity are receiving the most interest from public and private purchasers. Most notably, the Medicare Health Support Demonstration is a randomized trial to evaluate the effect of commercial disease management programs on fee-for-service beneficiaries with chronic conditions. It should also be mentioned that the Medicare program has a long tradition of experimenting with case management approaches that target high-risk, high-cost beneficiaries. Most recently, the High Cost Beneficiary Demonstration, which tests a variety of case management models, was launched. Given that most of the interest of public and pri-
private purchasers focuses on large population-based programs that are typically offered by third-party vendors or by health plans, our review pays particular attention to the evidence for population-based programs.

METHODS

Data

We conducted a library search of the PubMed and MEDLINE databases to identify relevant studies regarding the effect of disease management on chronic disease conditions, restricted to studies published from 1990 to 2005, with abstracts and in English, that focused on human subjects using the following search terms: disease management, care management, case management, self-management support, care coordination, management program, and care program, combined with utilization, admission, effectiveness, quality, outcomes, process, results, performance, efficiency, experiment, randomized controlled trial, pseudoexperiment, comparison, evaluation, analysis, impact, demonstration, study, and controlled clinical trial. We screened the full sample to identify reviews and meta-analyses as the focus of our analysis. Identified studies were reference mined, and experts in the field were asked to nominate additional articles. We did not attempt to re-review all individual articles discussed in the meta-analyses because of the large number of studies but rather used the findings and conclusions as presented in the reviews. In addition to the identified reviews and meta-analyses, we tried to identify studies for further review that analyzed the effect of population-based disease management programs, which attempt to identify and target every patient with a given condition irrespective of severity.

Our library search yielded 3831 articles. Based on expert recommendation, we also identified 4, existing reviews from sources that are not included in PubMed or MEDLINE.4,12-14 We screened abstracts of the 3835 articles and then identified 42 articles for full-text review (complete citations and data are extracted in the online Appendix, available at www.ajmc.com). However, 3 of those articles could not be retrieved. While screening the remaining 39 articles, we rejected 10 studies because their focus was irrelevant to our analysis (eg, no appropriate disease management components in their interventions), so we retained 29 studies. Of those, 3 represented evaluations of large-scale, population-based programs, 10 were meta-analyses, and 16 were systematic reviews of small-scale programs, covering 317 unique studies (Figure 2).

Data Analysis

For each article used, we abstracted into a database the conditions covered, the type of population the interventions targeted, and the organization responsible for designing and operating the intervention. For review articles, we also recorded the number of studies included and their design, sample size, and time period covered. Studies were grouped by condition, and 2 of us (S. Mattke and M. Seid) independently extracted the conclusions drawn by the authors (with differences resolved by consensus) on the effect of disease management using the following list of end points:

- Clinical processes of care (such as adherence to guidelines and best practices)
- Health-related behaviors (such as exercise patterns and medication compliance)
- Intermediate outcomes (disease control such as glycosylated hemoglobin [A1C] control in persons with DM)
- Clinical outcomes (such as mortality and functional status)
- Patient experience (such as patient satisfaction and health-related quality of life)
- Healthcare utilization (such as hospital admission rates)
- Financial outcomes (such as overall direct medical costs and net savings attributable to the intervention)

Reviews that covered more than 1 condition were included once under each condition (ie, they contributed multiple observations to our analysis). We recorded additional observations.
on the studies such as the strengths and weaknesses of their analyses. Based on the information extracted, we determined what conclusions could be drawn from the literature about the effect of disease management on the selected end points, considering the depth, consistency, and quality of the identified evidence, and developed condition-specific summaries.

RESULTS

What Do We Know About the Large Population-based Programs?

Limited information exists concerning large population-based disease management programs that target an entire diseased population. Two studies\(^{15,16}\) evaluated the effect of disease management programs run by large integrated delivery systems, and 1 study\(^{17}\) looked at a vendor-based program (Table 1). All 3 studies reported better quality of care. Two studies\(^{16,17}\) found that disease management for DM resulted in net cost savings, whereas the third study,\(^{15}\) which looked at 4 conditions, did not. None of those studies used a randomized design, and only 1 study\(^{17}\) had a rigorous quasiexperimental comparison strategy.

What Other Types of Disease Management Programs Have Been Reviewed?

Most of the evidence addressed small-scale programs, as none of the 26 meta-analyses or systematic reviews analyzed the effect of population-based disease management. The types of interventions covered in the identified publications were heterogeneous, and frequently little detail was provided about the intervention. However, the typical program was small (about 30-500 patients), was operated by the providers at a single site such as a hospital or a group practice, and was targeted to high-risk patients such as those at risk for hospital admission or readmission. The interventions typically combined individualized patient education, care planning, and follow-up delivered by a nurse or case manager by telephone or in person.

Which Conditions Have Been Included in Reviews?

As Table 2 shows, CHF and DM are the conditions for which the most evidence exists, with CHF clearly standing out. A sufficient body of research was also identified for coronary artery disease (CAD), asthma, chronic obstructive pulmonary disease (COPD), and depression, but not for other chronic conditions such as cancer, dementia, Alzheimer's disease, and musculoskeletal disorders. However, although the volume of evidence was similar for those 4 conditions, the results were most consistent for depression and CAD, somewhat less consistent for asthma, and inconclusive for COPD. It should be emphasized that hypertension and hyperlipidemia, as common disorders, are typically not managed as stand-alone conditions but as part of programs dealing with CAD, CHF, or DM.

What Are the Findings of the Reviews?

Across all conditions except asthma and COPD, there is consistent evidence that disease management can improve processes of care (eg, increased A1C screening for persons with DM). The results of the studies suggest that improved clinical care seems to lead to better intermediate outcomes and improved disease control (such as lower A1C levels in persons with DM), which was demonstrated for CHF, CAD, DM, and depression. Although the positive effect of disease control on long-term outcomes is well documented in the literature, we found no consistent evidence linking disease management to improved long-term outcomes. In most original studies, this result may be a consequence of the limited follow-up, which was commonly limited to 1 year, because changes in the long-term outcomes of patients may take

---

**Table 1**

|-----------|--------------|-----------|----------------|-------------|-----------|-----------------|

**Table 2**

|-----------|--------------|-----------|----------------|-------------|-----------|-----------------|

---

**Figure 2. Selection Process for Identifying and Including Studies**

PubMed (n = 3831) Identified by Experts (n = 4)

3835 Abstracts Screened

42 Articles Passed On to Systematic Review

29 Articles Contributed to Systematic Review:

- CHF: 18/118
- CAD: 7/78
- Diabetes: 9/64
- Asthma: 4/37
- COPD: 5/25
- Depression: 4/24

3 Evaluations of Population-based Programs

---
longer to materialize. If the documented effect of disease management proved sustainable, long-term outcomes should improve as well.

Overall, the evidence on the role of disease management in reducing utilization of health services was inconclusive, with the following 2 exceptions: disease management was found (1) to reduce hospitalization rates among patients with CHF and (2) to result in higher utilization of outpatient care and prescription drugs among patients with depression. When the costs of the intervention were appropriately accounted for and subtracted from any savings, there was no conclusive evidence that disease management leads to a net reduction of direct medical costs. However, the strength of this conclusion is limited because many studies do not address the issue of cost but focus on quality of care and outcomes; and many studies have methodological flaws such as the incomplete accounting of costs and the absence of a good comparison strategy. The findings are summarized in Table 3.

**Discussion**

Disease management enjoys much support among some in the healthcare community and among many policy makers and politicians. Recent legislation has put the Medicare program on a path toward adopting disease management as a standard benefit if the results are favorable in the ongoing Medicare Health Support Demonstration, which uses a randomized design to evaluate the effect of disease management for DM and CHF among Medicare beneficiaries. The potential for disease management to help solve the problems of the US healthcare system is likely to be subject to debate during the run-up to the 2008 presidential election. Financial markets share an optimistic view of disease management, as market leader Healthways Inc, for example, trades at a price-to-earnings ratio of approximately 43, compared with approximately 22 for the New York Stock Exchange Health Care Stock index and approximately 18 for the broad-based Standard & Poor's 500 index.

---

**Table 1. Findings from Studies of Large-scale, Population-based Disease Management Programs**

<table>
<thead>
<tr>
<th>Author</th>
<th>Setting</th>
<th>Managed Condition(s)</th>
<th>Comparison Strategy</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sidorov (2002)</td>
<td>Program developed and operated by integrated delivery system</td>
<td>Diabetes</td>
<td>Program participants versus nonparticipants, controlled for age, sex, insurance type, duration of plan enrollment, presence of pharmacy benefit</td>
<td>Improved quality of care and disease control, lower cost and utilization, net cost savings</td>
</tr>
<tr>
<td>Fireman (2004)</td>
<td>Program developed and operated by integrated delivery system</td>
<td>CAD, CHF, diabetes, asthma</td>
<td>Patients with the condition against nondiseased group matched by age and sex</td>
<td>Improved quality of care and disease control, cost increased less in intervention group than in reference group, no net cost savings</td>
</tr>
<tr>
<td>Villagra (2004)</td>
<td>Program developed and operated by disease management vendor for a health plan client</td>
<td>Diabetes</td>
<td>Natural experiment created by phased roll-out, plus pre-post comparison, adjustment for risk, and demographic differences</td>
<td>Improved quality of care, lower cost and utilization in both comparisons, net cost savings</td>
</tr>
</tbody>
</table>

CAD indicates coronary artery disease; CHF, congestive heart failure.

**Table 2. Summary of Meta-analysis/Reviews Included**

<table>
<thead>
<tr>
<th>Disease</th>
<th>Meta-analysis</th>
<th>Reviews</th>
<th>Individual Studies</th>
<th>Evaluation of Large-scale Interventions</th>
</tr>
</thead>
<tbody>
<tr>
<td>CHF</td>
<td>7</td>
<td>11</td>
<td>118</td>
<td>1</td>
</tr>
<tr>
<td>CAD</td>
<td>3</td>
<td>4</td>
<td>78</td>
<td>1</td>
</tr>
<tr>
<td>Diabetes</td>
<td>4</td>
<td>5</td>
<td>64</td>
<td>3</td>
</tr>
<tr>
<td>Asthma</td>
<td>2</td>
<td>2</td>
<td>37</td>
<td>1</td>
</tr>
<tr>
<td>COPD</td>
<td>2</td>
<td>3</td>
<td>25</td>
<td>0</td>
</tr>
<tr>
<td>Depression</td>
<td>2</td>
<td>2</td>
<td>24</td>
<td>0</td>
</tr>
</tbody>
</table>

CHF indicates congestive heart failure; CAD, coronary artery disease; COPD, chronic obstructive pulmonary disease.
Evidence for the Effect of Disease Management: Is $1 Billion a Year a Good Investment?

However, the results of our review suggest that, to date, support for population-based disease management is more an article of faith than a reasoned conclusion grounded on well-researched facts. Evidence specifically addressing large population-based disease management programs is limited to only 3\textsuperscript{15-17} evaluations, with 2\textsuperscript{16,17} of them being diabetes programs and only 1\textsuperscript{15} using a rigorous quasiexperimental design. Results from those studies are inconclusive with respect to the cost savings. Most of the evidence on disease management programs to date is derived from small high-intensity programs focusing on high-risk patients that are typically run as part of a demonstration project by the providers at a single site. This evidence suggests that those programs typically lead to better processes of care, but the evidence for improved long-term health outcomes and cost savings is inconclusive.

Research so far has concentrated on 6 common chronic conditions (CAD, CHF, DM, asthma, COPD, and depression). Little is known about the effect of disease management programs on other disorders such as cancer, dementia, and rare and expensive conditions such as hemophilia, which are increasingly being marketed by vendors. That there are only 3 studies\textsuperscript{15-17} of large-scale, population-based programs is somewhat surprising given that disease management vendors typically give performance guarantees and lose their fees if those targets are not met, creating means, motives, and opportunities for evaluative research. However, the vendor-run assessments typically do not meet the requirements of peer-reviewed research in terms of the comparison strategy, and adequate control for selection bias and regression to the mean. This does not imply that large-scale programs do not work: If the extensive use of mass communication and information technology allowed large-scale disease management programs to deliver an intervention of effectiveness similar to that of high-intensity programs with a lower cost per case, they might be able to improve quality and to reduce cost of care. This hypothesis needs to be further evaluated empirically.

**Limitations**

Our study largely relies on a review of reviews, and we based our conclusions on those of the original authors without independently verifying them. The reviews had some degree of overlap in the analyzed studies, which implies that some original studies entered our analysis more than once. It is difficult to determine how much bias this limitation may have introduced without reabstracting the original studies; however, the degree of overlap is not large, and some reviews and meta-analyses focused on different end points within those studies.

### Table 3. Summary of Evidence for Disease Management Program Outcomes by Condition

<table>
<thead>
<tr>
<th>Disease</th>
<th>Clinical Processes</th>
<th>Health-related</th>
<th>Disease Control</th>
<th>Clinical Outcomes</th>
<th>Healthcare Utilization</th>
<th>Financial Outcomes</th>
<th>Patient Experience</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Adherence to Evidence-based Guidelines</td>
<td>Changes in Behaviors</td>
<td>Changes in Intermediate Measures</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CHF</td>
<td>Improved</td>
<td>Inconclusive evidence</td>
<td>Improved</td>
<td>Inconclusive evidence</td>
<td>Reduced hospital admission rates</td>
<td>Inconclusive evidence</td>
<td>Improved</td>
</tr>
<tr>
<td>CAD</td>
<td>Improved</td>
<td>Evidence for no effect</td>
<td>Improved</td>
<td>Evidence for no effect</td>
<td>Inconclusive evidence</td>
<td>Inconclusive evidence</td>
<td>Insufficient evidence</td>
</tr>
<tr>
<td>Diabetes</td>
<td>Improved</td>
<td>Evidence for no effect</td>
<td>Improved</td>
<td>Inconclusive evidence</td>
<td>Inconclusive evidence</td>
<td>Inconclusive evidence</td>
<td>Insufficient evidence</td>
</tr>
<tr>
<td>Asthma</td>
<td>Inconclusive evidence</td>
<td>Inconclusive evidence</td>
<td>Inconclusive evidence</td>
<td>Evidence for no effect</td>
<td>Evidence for no effect</td>
<td>Insufficient evidence</td>
<td>Insufficient evidence</td>
</tr>
<tr>
<td>COPD</td>
<td>Insufficient evidence</td>
<td>Insufficient evidence</td>
<td>Inconclusive evidence</td>
<td>Insufficient evidence</td>
<td>Insufficient evidence</td>
<td>Insufficient evidence</td>
<td>Insufficient evidence</td>
</tr>
<tr>
<td>Depression</td>
<td>Improved</td>
<td>N/A</td>
<td>Improved</td>
<td>Inconclusive evidence</td>
<td>Increased utilization</td>
<td>Increased cost</td>
<td>Improved</td>
</tr>
</tbody>
</table>

Codes: N/A: not applicable, as no relevant health-related behaviors for depression exist. Disease-end point combinations in which disease management seems to achieve the intended result are shaded. Source: RAND analysis using identified articles. CHF indicates congestive heart failure; CAD, coronary artery disease; COPD, chronic obstructive pulmonary disease.
resulting in slightly less duplication. Because the work was conducted in 2006, only studies up to 2005 were considered.

Implications

Our findings suggest that a dynamic and innovative industry has outpaced its underlying science and that substantial efforts will be necessary to allow the evidence to catch up with product development. The Medicare Health Support Demonstration will greatly enhance our knowledge when final results become available in approximately 2010. Results reported in the first interim report showed that not one of 7 vendors could meet its performance guarantee of 5% net savings, and 2 vendors withdrew from the demonstration. As with any experiment, one should also keep in mind that the generalizability of this test is limited. Only Medicare fee-for-service beneficiaries are eligible, so the results may not apply to patients with insurance arrangements that control utilization more tightly or to a younger, commercially insured population with fewer comorbidities. This visible demonstration involves a few leading disease management operators and identifies only patients with DM and CHF. As a result, our ability to extrapolate to the routine operations of other programs will be limited.

Therefore, there is a great need for purchasers and vendors to seek out and pursue opportunities for quasiexperimental research that compares the outcomes of various disease management interventions with those found in an equivalent control group by using robust research designs such as matched-pairs analysis with propensity scoring, instrumental variables, regression discontinuity, or time-series analysis. Until a stronger base of evidence has been developed, public and private purchasers of disease management services should be skeptical about vendor claims and should demand supporting evidence based on transparent and scientifically sound methods.

Acknowledgments

Author Affiliations: From RAND Health, Arlington, Va (S. Martke, S. Ma); and Cincinnati Children’s Hospital and Medical Center (M. Seid).

Funding Source: This manuscript was supported by RAND Health’s Comprehensive Assessment of Reform Options (COMPARE) Initiative which receives funding from a consortium of sources including RAND’s corporate endowment, contributions from individual donors, corporations, foundations, and other organizations.

REFERENCES