Developing an Instrument to Measure Heart Failure Disease Management Program Intensity and Complexity

Barbara Riegel, DNSc, RN; Christopher S. Lee, PhD, RN, CCRN; Julie Sochalski, PhD, RN

Background—Comparing disease management programs and their effects is difficult because of wide variability in program intensity and complexity. The purpose of this effort was to develop an instrument that can be used to describe the intensity and complexity of heart failure (HF) disease management programs.

Methods and Results—Specific composition criteria were taken from the American Heart Association (AHA) taxonomy of disease management and hierarchically scored to allow users to describe the intensity and complexity of the domains and subdomains of HF disease management programs. The HF Disease Management Scoring Instrument (HF-DMSI) incorporates 6 of the 8 domains from the taxonomy: recipient, intervention content, delivery personnel, method of communication, intensity/complexity, and environment. The 3 intervention content subdomains (education/counseling, medication management, and peer support) are described separately. In this first test of the HF-DMSI, overall intensity (measured as duration) and complexity were rated using an ordinal scoring system. Possible scores reflect a clinical rationale and differ by category, with zero given only if the element could potentially be missing (eg, surveillance by remote monitoring). Content validity was evident as the instrument matches the existing AHA taxonomy. After revision and refinement, 2 authors obtained an inter-rater reliability intraclass correlation coefficient score of 0.918 (confidence interval, 0.880 to 0.944, \( P < 0.001 \)) in their rating of 12 studies. The areas with most variability among programs were delivery personnel and method of communication.

Conclusions—The HF-DMSI is useful for describing the intensity and complexity of HF disease management programs. (Circ Cardiovasc Qual Outcomes. 2010;3:324-330.)

Key Words: disease management ■ chronic disease ■ heart failure ■ instrument

Chronic illness afflicts nearly one-half of the US population and is especially prominent among the elderly. More than 83% of persons over age 65 have at least 1 chronic illness; 23% have 5 or more chronic conditions, regularly visit 13 different physicians, and annually fill >50 prescriptions.1,2 Disease management (DM), a systematic process of managing the care of patients with specific chronic diseases or conditions across the spectrum of outpatient, inpatient, and ancillary services, improves outcomes by reducing variations in care and minimizing episodes that require acute care and reducing healthcare costs.

Goals and Vision of the Program

The specific care management processes used in DM programs are called population health improvement processes by the Disease Management Association of America (DMAA).3 These processes are designed to empower individuals and care providers to work together to effectively manage the disease and prevent complications through symptom monitoring and treatment adherence—a goal that is not new or unique to DM. What is new to DM is the addition of an embedded continuous quality improvement process of following clinical guidelines and practice protocols, measuring outcomes, providing feedback to clinicians, and revising protocols as appropriate. These individual continuous quality improvement interventions are combined and standardized within disease management programs.

Local Challenges in Implementation

Descriptions of DM programs are common in the published literature. Making comparisons among DM programs, however, has proven difficult because of wide variability in the program’s recipients, the interventions provided, and the program’s “dose,” that is, the intensity (frequency and duration of interventions) and complexity (number of interventions) of the program. Of note, intensity and complexity are related, as intensity is assumed to be a function of duration. That is, no matter how masterful a specific intervention session is, if it is not reinforced over time, it has little hope of achieving a lasting effect.4 There also is a debate about what elements of DM are responsible for outcomes. That is, is the synergy produced from the DM program key to success or are

From the School of Nursing, University of Pennsylvania, Philadelphia, Pa. Dr Lee is currently affiliated with the School of Nursing, Oregon Health Sciences University, Portland, Ore.

Correspondence to Barbara Riegel, DNSc, School of Nursing, University of Pennsylvania, 418 Curie Blvd, Philadelphia, PA 19104-4217. E-mail briegel@nursing.upenn.edu

© 2010 American Heart Association, Inc.

Circ Cardiovasc Qual Outcomes is available at http://circoutcomes.ahajournals.org DOI: 10.1161/CIRCOUTCOMES.109.877324
specific intervention elements responsible? The purpose of this effort was to develop a tool that can be used to describe the intensity and complexity of disease management programs and thereby allow their comparison.

Design of the Initiative
In 2006, the American Heart Association (AHA) formed a writing group to develop a system of DM program classification, to assist efforts to evaluate the effectiveness of DM programs.5 The AHA taxonomy gave researchers a way to categorize and compare DM programs for chronic diseases and to evaluate their effectiveness. Building on that effort, we developed and tested a tool that can be used to describe and numerically score DM programs. Scoring programs may help discern whether specific elements versus whole programs contribute to producing the desired outcomes. This article describes the manner in which intervention elements were derived from the AHA taxonomy and used to define and build a measurement tool. In addition, we describe the first use of the tool.

The AHA Writing Group developed a DM taxonomy guided by the “Principles and Recommendations of the AHA Expert Panel on Disease Management.” The domains in the taxonomy were identified from a thorough review of the DM literature, comparing DM protocols used across heart failure (HF), diabetes, and depression populations. Articles on DM were located by searching the English language literature using the terms heart failure, diabetes, and depression combined with disease management, case management, and care management. Articles published between December 1987 and April 2005 were reviewed if they described a clearly defined protocol that incorporated at least 2 of the domains traditionally associated with DM.5

The Writing Group proposed a conceptual model of DM components that was subsequently refined through an inductive analytic process.6 This analytic process involved carefully reviewing published reports to condense and summarize the details of the DM protocols. A conceptual model of DM was proposed and then refined as program components missing from the originally proposed model were identified.

The first model proposed by the AHA Writing Group consisted of only 4 domains broadly relevant to DM programs: target population, intervention design, method of communication, and intensity. The final AHA taxonomy included 8 domains, each of which includes a number of subdomains: (1) patient population; (2) intervention recipient; (3) intervention content; (4) delivery personnel; (5) method of communication; (6) intensity and complexity; (7) environment; and (8) clinical outcomes (Figure).

In addition to identifying the requisite program components, describing DM program dose requires assessing both the intensity and complexity of programs. Yet, few published studies of DM are described in sufficient detail to make this determination. Most investigators who have studied program dose found that a higher dose of DM appears to improve outcomes7–9 and that low-dose programs are rarely effective. For example, a randomized controlled trial conducted by Weinberger et al10 provided medical inpatients with discharge

---

**Figure.** Disease management taxonomy. Reprinted from Krumholz et al,5 with permission from Lippincott Williams & Wilkins. Copyright 2006, American Heart Association.
planning, a primary care physician hospital visit, and assistance with follow-up appointment scheduling before hospital discharge. A nurse telephoned 87% of patients after discharge and spent, on average, 5.7 minutes per call. Half of patients received an appointment reminder. This program, characterized by few intervention components and a short duration of patient interaction, was associated with increased rather than decreased hospitalizations and hospital days compared with a control group. Similarly, in one of our early trials, a DM program with inconsistent intensity found no effect on HF rehospitalizations. In that study, only some hospitalized patients were visited by team members and telephone calls after discharge were brief and not standardized in format or depth.

A recent study testing basic and intensive DM doses found that neither dose improved outcomes in HF patients. However, in another study, a complex and intense transitional care DM program produced significantly improved outcomes. In this study, advanced practice nurses (APNs) worked collaboratively with patients’ physicians to facilitate the transition of elderly, chronically ill, complex patients from the hospital to home. After intense training by a multidisciplinary team of experts, APNs visited patients within 24 hours of a hospital admission and at least daily during the hospitalization. After discharge, the same APN visited patients at home at least 8 times (once within 24 hours of discharge) to ensure continuity of care and care coordination across settings. Special emphasis was placed on preventing functional decline, streamlining medication regimens, and working with the multidisciplinary team of providers to ease the transition across settings. The program increased time to first hospital readmission or death, decreased the number of readmissions and total costs, and increased quality of life and patient satisfaction.

Other studies suggest that a relationship exists between length and depth of the interaction between patients and providers and clinical and financial outcomes. However, measuring the intensity and complexity of interactions is challenging. Devising a method of assessing differences in DM programs would provide investigators with a standardized method of describing DM programs. With further testing, the proposed method may be useful in evaluating the effect of DM programs on patient outcomes.

Implementation of the Initiative

Few investigators have attempted to devise a method of quantifying the intensity and complexity of HF DM programs. A secondary analysis of data from 1136 patients obtained from 8 clinical sites in the United States found that high-intensity, comprehensive interventions (ie, relatively more contacts and intervention components) improved scores on the Minnesota Living with Heart Failure Questionnaire (LHFQ), a commonly used measure of health-related quality of life. Low-intensity interventions positively influenced LHFQ scores at 3 months, but scores deteriorated by 6 months, with those receiving low-intensity interventions reporting worse physical LHFQ subscale scores than those in the control group. In that study, an algorithm was developed to evaluate treatment intensity.

The algorithm was developed from published research that was used to develop a coding scheme that included the components common to successful HF DM programs. The draft algorithm was refined by study coinvestigators and then validated by experts in HF and quality of life. The lead authors coded and analyzed detailed descriptions of the interventions provided by investigators at the 8 sites, allocating 1 to 3 points for each component of an intervention as a measure of its ability to influence quality of life in persons with HF. For example, pharmacological therapy has been shown to influence survival and functional status but perceived quality of life has not always improved with medications alone. Alternatively, intense patient teaching and counseling can greatly improve quality of life in persons with HF, so the number of teaching sessions and the environment in which it was provided were accounted for in scoring. The coding of each intervention then was verified by the coinvestigator at each site. A summed score of intervention intensity was computed, with raw scores ranging from 2 to 9 and higher scores indicating higher intensity.

In a complementary effort, Driscoll et al developed a scoring system to evaluate complexity of HF DM programs in Australia. An intervention score was computed using (1) patient assessment, which was similar to patient population in the AHA taxonomy; (2) patient education and counseling, which addressed both type of education and method of communication; (3) promotion of self-care behavior—content emphasized during education and counseling; and (4) transitional care, which might be captured in the AHA taxonomy under environment. Programs with the highest comprehensiveness score had a 40% reduction in risk of all-cause hospital admission and mortality (hazard ratio, 0.60; 95% CI, 0.40 to 0.90; P=0.01).

Although limited by a focus on only intensity or only complexity, early efforts at coding and scoring HF DM interventions were useful in informing the approach used in this study to score HF DM programs. A convenience sample of DM programs provided in 12 prominent and frequently cited HF DM randomized clinical trials served as the data resource for the development of the HF DM program scoring instrument described here. These 12 trials were selected from studies cited in 2 comprehensive meta-analyses of DM interventions and chosen based on existing collaborative relationships among the nine investigators. The framework and taxonomy devised by Krumholz and the AHA Writing Group 5 was used as the basis for this effort.

Success of the Initiative

The HF Disease Management Scoring Instrument (HF-DMSI) takes the AHA taxonomy 1 step further by proposing a specific way to describe the domains and subdomains of HF DM programs. With further work, the ultimate goal is to develop a method to quantify the intensity and complexity of the DM provided in a particular HF program.

The instrument shown in the Table was developed from the AHA DM taxonomy. The HF-DMSI incorporates 6 of the 8 domains from the taxonomy: recipient, intervention content, delivery personnel, method of communication, intensity/complexity, and environment. Patient population was not
Table. Heart Failure Disease Management Scoring Instrument

<table>
<thead>
<tr>
<th>Intervention Category</th>
<th>Points to Be Assigned</th>
<th>Comment/Rationale</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recipient</td>
<td></td>
<td>Most interventions focus on the patient; yet, coding scheme recognizes that some interventions are aimed at improving provider behavior (eg, system intervention aimed at evidence-based care). Most points are given to interventions that focus on the patient but also include caregivers because an engaged family member act as 2nd set of eyes and memory support, which can deter hospitalization. Thus, 2 points assigned if focus was on patient alone; 3 points if there is some inclusion of the caregiver, 4 points if inclusion of the caregiver was a major component of the intervention.</td>
</tr>
<tr>
<td>Education and counseling aimed at supporting self-care</td>
<td>0 = No mention of education</td>
<td>Interventions are derived from Krumholz et al, with point allocations assigned to reflect current literature that suggests these interventions are not comparable in efficacy. Individualized patient education and counseling is essential because patients must be engaged in the process of self-care and helped to learn how to make decision about managing their HF. However, true self-care is more than treatment adherence.</td>
</tr>
<tr>
<td>Medication management</td>
<td>0 = No mention of medication regimen</td>
<td>Patients on optimal, evidence-based therapies are significantly less likely to have acute exacerbations and hospital admissions.</td>
</tr>
<tr>
<td>Social support Peer support</td>
<td>0 = No mention of a peer support intervention</td>
<td>Peer support interventions not used commonly but when used they appear to improve perceived support rather than self-care. Support has been conceptualized as a moderator of the relationship between intervention and outcome.</td>
</tr>
<tr>
<td>Surveillance by provider: Remote monitoring</td>
<td>0 = No use of remote monitoring or telehealth</td>
<td>Remote monitoring is distinguished from other methods of communication. Video monitoring may become a common method of communication. For now, remote monitoring is conceptualized as method of engaging patients in process of learning self-care by active engagement.</td>
</tr>
<tr>
<td>Delivery personnel Method of communication</td>
<td>0 = Mechanized via internet or telephone</td>
<td>Generalist: Provider specifically noted to not have training in heart failure.</td>
</tr>
<tr>
<td>1 = Single generalist provider (eg, physician, nurse, pharmacist)</td>
<td>1 = Person-to-person by telephone</td>
<td>Multidisciplinary interventions: Multidisciplinary team involved with all or most patients.</td>
</tr>
<tr>
<td>2 = Single HF expert provider (eg, physician, nurse, pharmacist)</td>
<td>3 = Face-to-face, individual, or in a group</td>
<td>Integrated/choreographed multidisciplinary intervention: Provided by multiple disciplines in collaboration; provided in an HF clinic with policies/protocols specified for HF care.</td>
</tr>
<tr>
<td>3 = Multidisciplinary intervention</td>
<td>4 = Combined: Face-to-face at least once alone or in a group with individual telephone calls in between meetings</td>
<td>Optimal mix of program delivery personnel is not known, thus assigned points are hypothesized in this study.</td>
</tr>
<tr>
<td>Points assigned if focus was on patient alone; 3 points if there is some inclusion of the caregiver, 4 points if inclusion of the caregiver was a major component of the intervention.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
specified because all the studies focused on HF DM. Outcome measures were not included because the purpose of this instrument was to measure the content that informs outcomes. That is, with an adequate measure of program content, outcomes can be evaluated in relation to content. In addition, each of the 3 intervention content subdomains (education/counseling, medication management, and peer support) was scored separately. Intervention content was anticipated to differ significantly among HF DM programs, so each area was detailed.

Both intensity (measured as duration) and complexity were rated using an ordinal scoring system devised for each of the 10 items in the HF-DMSI instrument. In the first category, recipient, 4 potential situations were listed in a hierarchy reflecting the support available to facilitate disease management: 1 = provider alone, 2 = patient alone, 3 = patient with some inclusion of caregiver, 4 = patient with a caregiver who is central to the intervention. The decision about scoring was made based on evidence that patient centered care is superior to standard, provider-focused efforts. Support from a caregiver also facilitates the success of DM efforts.

For the category of education and counseling aimed at supporting self-care, scores of 1 to 4 can be assigned based on the rationale that education that emphasizes the importance of treatment adherence (score = 1), one component of self-care, is less effective than education that helps people devise strategies to adhere (score = 2). However, treatment adherence alone is ineffective without the ability to make decisions about symptoms when they occur (score = 3). Better yet, the most effective self-care strategy is routine surveillance, management, and evaluation of symptoms in addition to treatment adherence (score 4).

For the category medication management, responses are scored from 1 to 3 with 1 given if there is some mention of medications (eg, importance of medication compliance), but medication management is not an active part of the intervention, and no attempt is documented to intervene with a provider to assure that patients receive an evidence-based medication regimen. The rationale here is that evidence-based medications have been shown to improve HF outcomes. A score of 2 is given if an evidence-based medication regimen is advocated, but there is no mention of follow-up with the patient or the provider to monitor whether or not the suggestion was implemented. A score of 3 is given if the medication regimen is monitored, if serious attempts are made to get the patient on evidence-based medications, and follow-up monitoring of the medication regimen is done with the patient or the provider.

Other scores assigned to reflect intervention differences (eg, peer support, surveillance by remote monitoring) generally used a 1 to 3 scoring system, with zero given only if the element could potentially be missing. That is, the category of recipient has no true zero because there will always be a recipient of DM. However, HF DM programs can be designed to cover some interventions but not others. For example, a program focusing on education and medical management might not have a peer support component. Thus, when it was possible to design a program without a specific domain, a true zero was available as an option.

Finally, a score of 1 to 4 was used to represent environment (1 = hospital: inpatient only; 2 = clinic/outpatient setting; 3 = home-based; 4 = combination of settings), based on the clinical rationale that hospitalized patients may not be as receptive to intervention when they are ill and hospitalized, whereas those who are contacted in the home environments may be more receptive. Those who receive some teaching in one setting (eg, hospital) with reinforcement in another setting later (eg, clinic or home) probably will have the best outcomes.

Once developed, the HF-DMSI was sent to each of the 9 principal investigators on the trials—each of whom was a nationally and internationally recognized expert in HF DM—for review and refinement. Content validity was evident in that the instrument matches the existing AHA taxonomy. Minor comments from the consortium members were incorporated into the instrument. Each of the principal investiga-
tors provided a full description of the DM intervention undertaken in their trials. Then, 2 authors (B.R. and C.S.L.) pilot-tested the HF-DMSI by rating each of the 12 trial using the materials describing the interventions provided the principal investigators. In some cases, in addition to the primary report of the clinical trial, there were additional articles and book chapters describing the intervention. When additional sources were available, they were used to assure that the highest score possible was attributed to each individual intervention. The 2 authors read the same reports and, blinded to each other’s scores, individually rated each of the HF DM interventions.

In the first test of the HF-DMSI, inter-rater reliability was only 58.5%, illustrating the need for clearer rationales for each category to facilitate consistent scoring. After that first test of the instrument, it was refined 4 more times. In each revision, the comment or rationale section was refined, expanded, and clarified. The scoring system was revised once. In that revision, the scoring options for medication management were reduced from 4 to 3 possible points after removing content on follow-up, which is covered in duration and complexity. The other scoring system revision was for complexity. The possible score for complexity was reduced from 4 to 3 with the integration of high and extremely options, which were hard to differentiate.

After the fourth revision of the 10-item HF-DMSI, inter-rater reliability was assessed using 2-way mixed-effect modeling to derive intraclass correlation coefficients. Single-measure intraclass correlation coefficient was 0.918 (95% confidence interval, 0.880 to 0.944; P<0.001), indicating adequate inter-rater reliability. To assess intra-rater (test, retest) reliability, the 2 raters individually rescored the HF DM programs 3 years after their original ratings using the same source documents. Spearman-Brown coefficients were 0.945 and 0.971 (both P<0.001) for the 2 raters over this extended period of time, indicating high intra-rater reliability. The high levels of intra-rater and inter-rater reliability over time are attributed to the detailed rationales for each section, although they could be due to the iterative process used in scoring the same studies repeatedly.

At this time, scores on the HF-DMSI should not be added because the manner in which complexity and intensity scores should be integrated is unclear. With use as a descriptive measure, a summary score may be able to be devised in the future and tested for validity.

Summary of the Experience, Future Directions, and Challenges

This article describes the development of a tool to describe the intensity and complexity of HF DM programs. An earlier version of this instrument was used in a subsequent study assessing the contribution of these HF DM programs to the reduction of hospital use and found to both differentiate among these programs and discern which program components were critical to differences in outcomes.27

Further, dose analyses are commonly compromised by failure to control for the fact that treatment dose may vary because patients choose different amounts of services or because different amounts are made available to patients based on individual patient characteristics.28 A program with limited intensity may be sufficient for patients with adequate health literacy and sufficient social support, for example. A limitation of the HF-DMSI is that it does not capture patient characteristics and the match between those characteristics and HF DM programmatic features. Efforts to account for such factors would be useful in research using the HF-DMSI.

The study described in this article was limited by the data available, as are all secondary analyses. Another limitation is that we only addressed HF DM programs and not the variety of programs used for other chronic conditions. Therefore, the patient population did not vary and could not be used in the analysis.

The HF-DMSI provides a consistent manner of describing the intensity and comprehensiveness of existing HF DM programs. The effort of Krumholz and the AHA Writing Group6 greatly facilitated the effort to devise a system of evaluating HF DM programs that functioned well in this study. Others are encouraged to use and refine the HF-DMSI in future evaluations of HF DM programs. Elements of the instrument may be useful to investigators desiring to describe other types of DM programs as well.

In conclusion, evaluations of intensity and complexity can inform and clarify whether inconsistent outcomes from HF DM programs are due to the range of interventions undertaken or their insufficient use. Evaluating what is provided to patients (complexity) and how long the intervention lasts (intensity) may provide the next level of refinement needed to move this field forward.

Sources of Funding

Funding was provided by the Commonwealth Fund (Sochalski).

Disclosures

None.

References

Developing an Instrument to Measure Heart Failure Disease Management Program
Intensity and Complexity
Barbara Riegel, Christopher S. Lee and Julie Sochalski

doi: 10.1161/CIRQUO.109.877324

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://circoutcomes.ahajournals.org/content/3/3/324

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in Circulation: Cardiovascular Quality and Outcomes can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the Permissions and Rights Question and Answer document.

Reprints: Information about reprints can be found online at:
http://www.lww.com/reprints

Subscriptions: Information about subscribing to Circulation: Cardiovascular Quality and Outcomes is online at:
http://circoutcomes.ahajournals.org/subscriptions/