A COMPARISON OF THE EFFECT OF COMPUTER AND
MANUAL REMINDERS ON COMPLIANCE WITH
A MENTAL HEALTH CLINICAL
PRACTICE GUIDELINE

by

Dale Sherman Cannon

A thesis submitted to the faculty of
The University of Utah
in partial fulfillment of the requirements for the degree of

Master of Science

Department of Medical Informatics
The University of Utah
August 1999
THE UNIVERSITY OF UTAH GRADUATE SCHOOL

SUPERVISORY COMMITTEE APPROVAL

of a thesis submitted by

Dale Sherman Cannon

This thesis has been read by each member of the following supervisory committee and by majority vote has been found to be satisfactory.

Chair: J. Lincoln

Reed M. Gardner
To the Graduate Council of the University of Utah:

I have read the dissertation of Dale Sherman Cannon in its final form and have found that (1) its format, citations, and bibliographic style are consistent and acceptable; (2) its illustrative materials including figures, tables, and charts are in place; and (3) the final manuscript is satisfactory to the supervisory committee and is ready for submission to The Graduate School.

Michael J. Lincirc
Chair, Supervisory Committee

Approved for the Major Department

Reed M. Gardner
Chair/Dean

Approved for the Graduate Council

David S. Chapman
Dean of The Graduate School
ABSTRACT

The objective of this study was to evaluate the relative effectiveness of a computerized decision support system and a manual reminder system on the implementation of a Major Depressive Disorder clinical practice guideline in an outpatient mental health clinic. The study was a randomized clinical trial in which 76 patients were randomly assigned within clinician to one of the two experimental conditions. The first condition was a checklist inserted in the paper medical record, and the second was a computerized decision support system entitled the CaseWalker. The CaseWalker reminded clinicians when guideline-recommended screening for mood disorder was due, ensured the fidelity of the diagnosis of Major Depressive Disorder with the Diagnostic and Statistical Manual, Fourth Edition (DSM-IV) criteria, and generated a progress note.

The CaseWalker, compared to the paper checklist, resulted in higher screening rates for mood disorder (86.5% vs. 64.1%, \( p = .018 \)) and more consistent application of DSM-IV criteria (100% vs. 5.6%, \( p < .001 \)). Clinician acceptance of the CaseWalker was good, with three of four clinicians strongly preferring it to the paper checklist. Thus, computer reminders were shown to be superior to manual reminders in supporting the implementation of a depression clinical practice guideline in an outpatient mental health clinic. A previous review of the literature on reminder systems concluded there is
insufficient evidence for the superiority of computer reminders over manual reminders. However, the results of the present study provide further empirical evidence that computer reminders are more effective than manual reminders. Further, this study is believed to be the first randomized controlled trial of the efficacy of computerized reminders in a mental health clinic.
This thesis is dedicated to the veterans of United States military service who risked their lives, gave their time, and, in some cases, injured their health to keep this nation free. This thesis in no way compensates these men and women for their sacrifices, but it is hoped that this study will make some contribution, however small, to the improvement of the quality of the healthcare that they so richly deserve.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
<td>iv</td>
</tr>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>ix</td>
</tr>
<tr>
<td>Chapter</td>
<td></td>
</tr>
<tr>
<td>I. INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>Clinical Practice Guidelines</td>
<td>1</td>
</tr>
<tr>
<td>Decision Support for Clinical Practice Guidelines</td>
<td>5</td>
</tr>
<tr>
<td>Veterans Health Administration Clinical Guideline for Major Depressive Disorder</td>
<td>11</td>
</tr>
<tr>
<td>II. METHODS</td>
<td>19</td>
</tr>
<tr>
<td>Research Participants</td>
<td>19</td>
</tr>
<tr>
<td>Study Site</td>
<td>19</td>
</tr>
<tr>
<td>Measures</td>
<td>19</td>
</tr>
<tr>
<td>CaseWalker</td>
<td>21</td>
</tr>
<tr>
<td>Software Developers</td>
<td>31</td>
</tr>
<tr>
<td>Procedure</td>
<td>32</td>
</tr>
<tr>
<td>Data Analyses</td>
<td>34</td>
</tr>
<tr>
<td>III. RESULTS</td>
<td>36</td>
</tr>
<tr>
<td>Screening for Mood Disorder</td>
<td>36</td>
</tr>
<tr>
<td>Documentation of the Diagnosis of MDD</td>
<td>36</td>
</tr>
<tr>
<td>Medical Record Documentation</td>
<td>38</td>
</tr>
<tr>
<td>Six-week Reviews</td>
<td>40</td>
</tr>
<tr>
<td>Interview Findings</td>
<td>41</td>
</tr>
<tr>
<td>IV. DISCUSSION</td>
<td>43</td>
</tr>
<tr>
<td>Relative Effectiveness of the CaseWalker and the Paper Checklist</td>
<td>43</td>
</tr>
<tr>
<td>Critique of the CaseWalker</td>
<td>50</td>
</tr>
<tr>
<td>Future Development using VISTA Clinical Reminders</td>
<td>52</td>
</tr>
<tr>
<td>Clinician Attitudes toward the MDD Guideline</td>
<td>55</td>
</tr>
</tbody>
</table>
Appendices

A. GUIDELINE FLOWCHART ................................................................. 57
B. PAPER CHECKLIST ........................................................................ 61
C. CLINICIAN INTERVIEW .................................................................. 66
D. CASEWALKER DATABASE ............................................................ 69
REFERENCES .................................................................................. 72
CURRICULUM VITAE ........................................................................ 78
ACKNOWLEDGMENTS

My deepest love and gratitude are expressed to my wife, Elizabeth (Becky) Cannon, for her positive and unselfish support of me through yet one more graduate program. My pursuit of a degree in Medical Informatics has taken time away from Becky and money out of the family budget at a stage of our lives when most people expect their spouses to be more settled in their careers. Thank you, Becky, for letting me go back to school and for being such an unfailing supporter of my efforts.

I also wish to thank Dr. Thomas B. Horvath, my supervisor at the time I began this degree, for his encouragement. Thank you, Tom, for believing that my obtaining a masters degree in Medical Informatics would be good for Veterans Health Administration Mental Health programs and for permitting me to take time to complete the degree.

I am grateful to the leadership of the Mental Health Care Center at the Salt Lake Veterans Affairs Medical Center, particularly to Craig Hummel, M.D., and Kelly Lundberg, Ph.D., for their administrative approval and support of this research project. I am especially grateful to the staff of the PostTraumatic Stress Disorder Clinical Team (PCT) at the Salt Lake Veterans Affairs Medical Center for their active participation in the study. Appreciation is expressed to Steven M. Allen, Ph.D., Chief of the PCT, for encouraging his staff to participate in the project and for tending to many of the day-to-day administrative tasks required to complete the study. Thanks also are expressed to
Mark Cappel, L.C.S.W., Phillip Christensen, Ph.D., Jan Newren, R.N., and Steve Shore, A.T., for being the research participants. Without their assistance and cooperation, obviously the study would not have been possible.

I wish to acknowledge the brilliant contributions of Jeff Sells, Ph.D. and Robert Feldman to the development of the CaseWalker software. Thanks you, good friends, for your tireless effort to capture in code my changing notions of what the CaseWalker application should do. Thanks, too, for your many independent contributions to the design of the program.

Finally, I am grateful to the faculty of the Department of Medical Informatics for sharing their knowledge with me and for their interest in my education. In this regard, I am especially grateful to my thesis committee for their contributions to the design of this study. I thank Michael Lincoln, M.D., for chairing the thesis committee and for his insightful contributions to the initial design of the CaseWalker program. I am grateful to Peter Haug, M.D., for the decision support course he taught that stimulated my interest in clinical decision support. I thank Reed Gardner, Ph.D., for his willingness to join the committee after a vacancy occurred and for his many helpful comments on the written thesis. I also wish to thank Helmuth Orthner, Ph.D., for being a member of the thesis committee during most of my graduate tenure until he left to assume an academic post elsewhere.
CHAPTER I

INTRODUCTION

Clinical Practice Guidelines

A clinical practice guideline is a set of diagnostic and treatment recommendations for a medical disorder. The Institute of Medicine ([IOM], 1990) defines guidelines as "systematically developed statements to assist practitioner and patient decisions about appropriate health care for specific clinical circumstances." The Veterans Health Administration ([VHA], 1996, p. i) defines clinical practice guidelines as "recommendations for the performance or exclusion of specific procedures or services derived through a rigorous methodological approach that includes the following:

- "Determination of appropriate criteria, such as effectiveness, efficacy, population benefit, or patient satisfaction; and
- "Literature review to determine the strength of the evidence (based in part on study design) in relation to these criteria."

Clinical practice guidelines are differentiated from clinical pathways in that guidelines are more general in content than clinical pathways. In VHA, clinical practice guidelines are established centrally for use throughout the agency's 172 medical centers, and clinical pathways are established at each medical center for local implementation (VHA, 1996).

The use of clinical practice guidelines has become widespread over the past decade for many reasons. In 1989, a federal agency, the Agency for Health Care Policy and Research (AHCPR), was created "to enhance the quality, appropriateness, and
effectiveness of health care services” (ACHPR, 1993). One of the original missions of AHCPR was the development and dissemination of clinical practice guidelines, and it has published a number of nationally recognized guidelines (e.g., Rush, Golden, Hall, et al., 1993a,b). Another reason for the increased use of clinical practice guidelines is the belief that guidelines are useful for a wide variety of purposes. For example, the IOM (1990) stated that guidelines could be used to “establish quality review criteria and set quality improvement goals, guide medical insurance benefits decisions and regulate reimbursement, govern licensure and accreditation, develop information and decision-support systems, guide cost-control efforts, determine tort liability, and inform the decisions of consumers and purchasers of health insurance and managed care” (Brown, Shye & McFarland, 1995). Finally, there is an economic motivation driving the use of clinical practice guidelines. This motivation is based on the assumption that practice variation is at least partially the result of the overuse of medical services and the hope that, if clinical practice guidelines reduce practice variation, the expense of providing medical care will decrease as a consequence (Walker, Howard, Lambert, & Suchinsky, 1994).

If guidelines are to accomplish their intended purposes, physicians and other health care providers must follow them. Grilli and Lomas (1994), in a review of the English literature on guideline compliance published between 1981 and 1990, found that the mean compliance rate with guideline recommendations was only 55%.

Before considering ways in which guideline compliance might be increased, it should be noted that the goal is not necessarily 100% compliance. A caveat attached to virtually every guideline is that it is just a guideline and that it should not replace sound
clinical judgment about the individual case. Ellrodt, Conner, Riedinger, and Weingarten (1995) illustrated the point that sometimes it is medically appropriate not to follow a guideline. They studied the reasons physicians did not comply with a guideline that was intended to reduce hospital length of stay for patients who presented with chest pain. The guideline compliance rate was 66%. Physicians failed to follow the guideline without some overriding justification in only 6% of the cases. A retrospective review by other clinicians determined that the physician appropriately extended the hospital stay in 13% of the cases either because the patient’s condition had deteriorated or because the patient had been misclassified by the guideline as a low risk patient. In 5% of the cases, the reason for “noncompliance” (i.e., keeping the patient in the hospital too long) was that the hospital staff were too slow in scheduling or reporting the results of tests the doctor had ordered.

The usual first step in implementing a clinical practice guideline is to educate practitioners about the guideline through some passive pedagogical technique such as lectures or publication of the guideline. Although such techniques may increase awareness of the guideline and may be necessary, they have not proven to be very effective in changing actual practice patterns. For example, Lomas et al. (1989) report the results of a very thorough effort to disseminate a guideline on cesarean delivery in Ontario, Canada. The guideline was mailed to every obstetrician and gynecologist as well as to every hospital in the province. In addition, it was published in medical journals, society bulletins, and the newsletters of consumer groups. A year following this dissemination effort, 94% of the obstetricians and gynecologists in the province said they were aware of the guideline, 83% said they agreed with it, and 33% said they had
changed their practice as a result of the guideline. However, these physicians did poorly on a test of their knowledge of the content of the guideline. Sixty-seven percent of their answers on a test of guideline knowledge were correct, which was only slightly higher than the 50% correct expected by chance. Further, a review of hospital discharge records failed to substantiate the change in practice patterns that was reported by physicians. There was in fact no change in cesarean delivery rates over the year following dissemination of the guideline.

One set of variables that affect guideline compliance is the nature of the guideline itself. Grilli and Lomas (1994) found two characteristics of a guideline affected compliance rate, viz., “complexity” and “trialability.” “Complexity” was defined as “the degree to which the procedure is perceived to be difficult to understand, or requires the availability of specific resource changes for implementation.” “Trialability” was defined as “the extent to which the procedure can be experimented with on a limited basis before making a final decision to adopt.” Grilli and Lomas report that complexity diminishes guideline compliance and trialability increases guideline compliance.

Several investigators have pointed out the importance of involving clinicians in the development and dissemination of guidelines (e.g., Horne, 1996; Karuza, et al., 1995; Tierney, Overhage, & McDonald, 1996). Karuza et al. (1995) randomly assigned group practices to either an experimental or control condition to assess the effect of small-group consensus building on the implementation of an influenza vaccination guideline. Physicians in the group practices assigned to the experimental arm of the study participated in a 1-hour discussion of the guideline with the other physicians in their group practice. An external facilitator, who followed a semistructured format for the
meeting, concluded the session by asking physicians to make a public oral commitment to increase the number of older patients who received the vaccination. Those physicians in the control condition participated in a 1-hour discussion of an unrelated health topic. Physicians in the experimental arm of the study had a 34% higher rate of immunization among elderly patients.

*Decision Support for Clinical Practice Guidelines*

“Decision Support Systems, Clinical” is defined by the National Library of Medicine MeSH Browser as “computer-based information systems used to integrate clinical and patient information and provide support for decision-making in patient care” (Brylawski, 1999). Such systems have been used for many purposes, including aiding diagnoses (e.g., Berner, et al., 1994; Li, Haug & Warner, 1994; Miller, 1994) and alerting systems triggered by changes in the patient’s condition or by questionable medication orders (e.g., Haug, et al., 1994). There are a variety of models used in decision support systems, including rule-based, mathematical (both deterministic and statistical), cognitive, case-based reasoning, and scoring models (P. J. Haug, personal communication, March, 1997). A comprehensive review of clinical decision support systems is beyond the scope of this thesis, which is limited to a consideration of rule-based systems to support clinical practice guidelines. Such systems include reminders to perform guideline-recommended actions, support for following rule-based algorithms, and documentation support. Reminder systems, in turn, will be expanded in this thesis to include manual, paper-based systems.

McDonald (1976) suggested over 20 years ago that many medical errors are due to the “non-perfectability (sic) of man” rather than to lack of knowledge or wrong intent.
He argued that physicians are faced with more information than is humanly possible to process, and so he suggested that computer-generated reminders would reduce medical errors. His pioneering work (McDonald, 1976; McDonald, et al., 1984; McDonald, Wilson, & McCabe, 1980; Wilson, McDonald, & McCabe, 1982;) and that of Barnett and associates (Barnett, 1984; Barnett, Winickoff, Dorsey, Morgan & Lurie, 1978; Barnett, Winickoff, Morgan, & Zielstorff, 1983) were early demonstrations that computer reminder systems can reduce medical errors.

Subsequently, literally hundreds of papers have been published that describe computerized decision support systems. Unfortunately, much of this literature is flawed from a scientific standpoint (see Haynes & Walker, 1987; Langton, Johnston, Haynes & Mathieu, 1992; and Lundsgaarde, 1987 for methodological critiques). Fortunately, several excellent systematic review articles and two meta-analyses have been published recently that summarize the best of this vast literature, so the present review will be limited to a consideration of these published review articles and the meta-analyses.

Balas et al. (1996) reviewed all published randomized clinical trials of computerized information systems. Reminders targeted towards both providers and patients were found to be effective, but in 64% of all the studies the provider rather than the patient was the target of the information system. The majority of information systems (82%) were tested in outpatient settings, and 66% were tested in primary care clinics. The dependent measure was the process of care, as opposed to the outcome of care, in 76% of cases.

Austin, Balas, Mitchell and Ewigman (1994) published a meta-analysis of randomized clinical trials of physician reminder systems for preventive health care
measures in family or internal medicine clinics. The results indicate that reminder systems for cervical cancer screening and tetanus immunization effectively increase physician compliance with these preventive health care measures. The authors believed the evidence for the effectiveness of tetanus immunization reminders is so strong that further trials of such systems "would be unnecessary and probably unethical."

Shea, DuMouchel and Bahamonde (1996) also did a meta-analysis of randomized clinical trials of physician reminder systems in ambulatory care clinics. They concluded that the evidence supports the effectiveness of reminder systems for immunizations, breast cancer screening, and cardiovascular risk reduction. However, contrary to the findings of Austin et al. (1994), Shea et al. (1996) concluded the effectiveness of reminders for cervical cancer screening is not supported by the literature.

The most recent systematic review of reminder systems was by Hunt, Haynes, Hanna and Smith (1998), and it was an update of two earlier reviews by the same group (Haynes & Walker, 1994; Johnston, Langton, Haynes & Mathieu, 1994). The Hunt et al. (1998) review is not only the most recent, but it appears to have used the most comprehensive, thorough search strategy to identify relevant studies. As was true of the previously cited reviews and meta-analyses, Hunt et al. reviewed only prospective randomized clinical trials. They reviewed drug dosing systems, diagnostic aids, and reminder systems. Fourteen (74%) of the 19 reminder system studies reviewed improved physician performance. Only one of the reminder studies included an analysis of patient outcomes, and that study had negative findings (Barnett, et al., 1983).

Computerized reminder systems do not always have positive results. For example, Mendelson (1986) developed a reminder system to improve the documentation
of compliance with a guideline. However, physicians using the system did not believe
that improving documentation was an important goal, and so the reminders caused much
resentment. Overhage, Tierney and McDonald (1996) report that, despite the success of
computer reminders in increasing preventive care in outpatient settings, a similar
intervention in an inpatient setting was unsuccessful.

Several factors have been shown empirically to improve the effectiveness of
computerized reminder systems. For example, Tierney, Hui and McDonald (1986)
demonstrated that both immediate reminders and monthly feedback reports increased
physician compliance with suggested preventive care protocols, but the immediate
reminders were more effective. Lobach (1996) found a low guideline compliance rate
(6%) even with a computerized reminder system. However, sending physicians an e-mail
message biweekly that summarized their individual compliance with the guideline
increased compliance to 35%. Requiring physicians to respond to computerized
reminders has been shown to increase compliance with preventive care guidelines
describe a web-based clinical guideline support system that includes physician and
patient education components as well as a graphical representation of the clinical
guideline.

An alternative to computer decision support systems is manual systems in the
form of paper checklists, structured encounter forms and flow sheets (e.g., Madlon-Kay,
1987, 1998; Schreiner, Petrusa, Rettie & Kluge, 1988; Yarnall, Michener, Broadhead &
Tse, 1993). Such manual systems undoubtedly are more widely used than computer
systems because they are very inexpensive and easy to create and require little or no staff
training to implement. Typically, checklists in the paper medical record are intended to remind clinicians of required actions, guide them through an implicit algorithm, and document the action and the results.

Although manual reminder systems have not been evaluated as thoroughly as computerized systems, there is some experimental evidence that manual systems increase preventive health screening by physicians in comparison with screening rates obtained with no reminder or decision support system. Methodologically, these studies consist of randomized trials (Cheney & Ramsdell, 1987), pre- and posttest designs (Kohatsu, Cramer & Bohnstedt, 1994; Prislin, Vandenbark & Clarkson, 1986; Shank, Powell & Llewelyn, 1989), and a nonrandomized comparison of two clinics (Cohen, Littenberg, Wetzel & Neuhauser, 1982). On the other hand, Cowan, Heckerling and Parker (1992), in a randomized controlled trial, found that no significant improvement in compliance with health improvement recommendations resulted from placing the recommendations on the chart.

Given that both computerized and manual decision support systems are widely used and have empirical support, a direct comparison of the two is indicated. If integrated into the electronic medical record, the computerized reminder system would have the advantage of being able to read existing data from that record and of being able to write new data to that record. A manual system, on the other hand, would be less expensive and would be easier to develop. Do the benefits of computerized systems justify the time and effort required to implement them? The most thorough analysis of this question is found in the previously cited meta-analysis by Shea et al. (1996). They considered four intervention conditions in their analyses, viz., computer reminders alone,
manual reminders alone, both types of reminders combined, and no reminders. Both types of reminder systems alone resulted in higher compliance with preventive medicine recommendations than did the no reminder control condition. However, there was no significant difference between computer reminders alone and manual reminders alone across the studies included in the meta-analysis. There was a small advantage observed from combing both types of reminders systems over the use of either type of reminder alone. Contrary to the findings of the Shea et al. meta-analysis, two prospective, controlled trials of computer and manual reminder systems report significant benefits of computer reminders compared to the manual systems (Frame, Zimmer, Werth, Hall & Eberly, 1994; Tape & Campbell, 1993). The Frame et al. study was included in the Shea meta-analysis, but the Tape and Campbell study was not. The lack of compelling evidence of the superiority of computer reminders over manual reminders indicates that further study of their relative effectiveness is indicated.

In a review of both computer and manual reminder systems, Murrey, Gottlieb, and Schoenbaum (1992) stated that such systems “have met with significant success in almost all published studies of their effectiveness...” (p. 423). In spite of this overall conclusion, Murrey et al. pointed out that there is a great deal of variety in reminder systems in terms of their design, objective, and the targeted recipients of the reminders. Murrey et al. (1992) suggested that these differences result in substantial variability in compliance rates, but they did not characterize the nature of these effects. They did argue, however, that investigators need to be more specific about the nature of the reminder systems they are investigating, and they proposed a typology of reminder
systems. The five types of reminder systems they described are concurrent reports, intervisit reminders, registry reminders, enhanced lab reports, and statistical summaries.

*Veterans Health Administration Clinical Guideline for*

*Major Depressive Disorder*

VHA has developed a clinical practice guideline for the treatment of Major Depressive Disorder (MDD) (Veterans Health Administration, 1996). This guideline was designed for implementation in each of VHA’s 172 medical centers nationwide. The intent of the guideline was to increase the detection of MDD and to encourage the use of treatments for MDD that are supported by the scientific literature.

The MDD clinical practice guideline comprises five modules. The first module recommends assessment procedures to be followed with all patients and then addresses the treatment of mood disorder in medical patients who are followed in primary care settings. The next two modules outline the treatment of uncomplicated MDD in inpatient and outpatient mental health settings. The fourth module addresses the treatment of MDD with comorbid Posttraumatic Stress Disorder (PTSD), and the fifth module addresses the treatment of MDD with comorbid substance dependence. Although each of these modules contains material specific to the treatment setting or special population for which it was developed, all five share common algorithms for the screening, diagnosis, and treatment of MDD.

The MDD guideline was written in 1996 by a multidisciplinary group of 64 people. Forty-four were clinicians (27 from VHA field facilities, 11 from VHA Headquarters, and 6 from outside VHA) who were considered to be subject matter experts, and the rest were expert in guideline development, quality management, or staff
education. Of the 44 clinicians, 19 were psychiatrists, 12 were psychologists, 4 were social workers, 3 were nurses, 2 were primary care physicians, 2 were chaplains, 1 was a geriatrician, and 1 was a pharmacist.

The MDD clinical practice guideline is a rule-based algorithm. For example, Box 6 of the MDD/PTSD module (see Appendix A) may be stated as an “if-then” rule: If the patient meets Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM-IV) (American Psychiatric Association, 1994) criteria for PTSD, then go to Box 8, which is another rule concerning whether the patient needs to be hospitalized. If the patient does not meet DSM-IV criteria for PTSD, then go to another module of the guideline that covers the outpatient treatment of MDD patients without comorbid PTSD. Clinical issues that cannot be reduced to a simple set of “if-then” statements are covered in the MDD clinical practice guideline as narrative “annotations.” For example, Box 28 of the MDD/PTSD module is an “action” box (rather than a “decision” box) that states, “Initiate treatment for MDD/PTSD.” The “J” in the box is a reference to Annotation J, which is a multipage description of the use of psychotherapy and pharmacotherapy in the treatment of comorbid MDD and PTSD.

The MDD clinical practice guideline was only the second guideline that VHA attempted to implement throughout the agency, and initial implementation efforts relied heavily on education. The MDD guideline implementation effort was launched at a national training conference in Phoenix in September 1996, which was attended by over 200 clinicians from around the country. Within each of 22 geographical regions called Veterans Integrated Service Networks (VISNs), one medical center had been selected to be a pilot site for implementing the guideline, and staff from each of the pilot sites were
asked to attend the Phoenix conference. The conference stressed the clinical need for the guideline and reviewed the key components of the guideline in educational lectures given to the entire group. Then, attendees were divided into small groups to discuss implementation strategies and barriers to implementation at their local medical centers. The plan was that conference attendees would implement the guideline at their medical centers after the conference, supported by monthly conference calls at which they could share successful and unsuccessful implementation strategies. Also, MDD subject matter experts were made available to consult with local medical centers upon request. After successful implementation of the guideline at the pilot sites, pilot site staff members were to lead guideline implementation efforts at other medical centers within their VISN.

VHA has largely abandoned this educational approach to other guideline implementation because it was judged not to be cost-effective. The pilot sites, with few exceptions, did not implement the MDD guideline, a finding that is consistent with evaluations of other guideline educational efforts (e.g., Lomas, et al., 1989). Even if the educational efforts had been more successful, a reduction in money to support national conferences and the increasing number of clinical practice guidelines that were to be implemented within VHA made additional meetings such as the Phoenix conference impossible. Consequently, VHA has come to rely on incentives for managers to implement guidelines based on “performance measures.” For each clinical practice guideline, a small set of measures have been developed as indices of implementation of what are judged to be the most important components of the guideline. These measures then are included in the performance contract each Veterans Integrated Service Network (VISN) Director has with the VHA Undersecretary of Health. There are positive
financial incentives for the VISN Director to meet the performance measures and negative consequences (not necessarily financial) for not meeting them. VISN Directors, in turn, apply pressure on medical center directors to meet the performance measures, the medical center directors apply pressure on clinical program managers, and the clinical program managers apply pressure on line clinicians.

Three performance measures have been established for the MDD guideline. (1) All patients seen at least three times in Primary Care outpatient clinics within 1 year are to be screened annually for mood disorder. (2) MDD diagnoses given to patients who are followed in outpatient mental health clinics must specify which DSM-IV criteria for MDD that the patient meets. (3) Mental Health inpatients who receive an MDD diagnosis must be given a Global Assessment of Functioning (GAF) rating. The third performance measure is not very demanding because all mental health inpatients already are required to have a GAF rating included in their discharge summary regardless of diagnosis. The first performance measure has been given the most attention throughout VHA in the hope that the increased detection of MDD in Primary Care would have the greatest clinical impact. For each performance measure, a criterion of 80% of subject cases meeting the measure has been set for the VISN Director to get a “Fully Successful” rating, and a 90% criterion has been established for an “Outstanding” rating.

VHA’s performance measure based implementation strategy should be considered in terms of variables that previous research has shown to affect successful guideline implementation. The entire MDD guideline would be classified as “complex” as Grilli and Lomas (1994) defined the term, but the performance measures are not. The performance measures are simple to understand, and no resource changes are required to
implement them. Noncomplex guidelines are implemented more readily (Grilli & Lomas, 1994). The second variable related to successful implementation is “trialability,” i.e., “the extent to which the procedure can be experimented with on a limited basis before making a decision to adopt.” Theoretically, the guidelines operationalized in the performance measures are “trialable,” but the decision as to whether or not to implement them has been removed from the clinician. The fact that the decision to implement the guideline was removed from the clinician is related to the issue of whether clinicians are involved in guideline development and implementation (Horne, 1996; Karuza, et al., 1995; Tierney, et al., 1996). These studies indicate that clinician involvement enhances guideline implementation. Although a large number of clinicians were involved in developing the guideline, they constitute a very small percentage of the total number of VHA clinicians who are expected to implement the guideline. Thus, most VHA clinicians would be expected to see the guideline as something imposed from above. The fact that there is administrative pressure to implement the guideline may exacerbate the perception by clinicians that the guideline was being imposed on them. Administrative pressure may increase guideline compliance, but it will not convince staff that following the guideline has intrinsic merit. Thus, the Mendelson (1986) study would suggest the guideline might be resented rather than appreciated by clinicians. On the other hand, neither the guidelines nor the performance measures stipulate how the guideline is to be implemented, so the opportunity for clinician participation in this process exists. In summary, the lack of complexity in the performance measures might enhance their implementation, but the lack of local clinician participation in their development might diminish their implementation. Involving clinicians in the development of
implementation strategies might have offset the fact that they were not involved in the
development of the guideline.

Further, it should be noted that the performance measures reflect the expectation
that the guideline be followed in all cases. In other words, medical contraindications to
the guideline are not envisioned by the performance measures.

The present study was a comparison of two methods for increasing the screening
rate for mood disorders and for increasing the fidelity with DSM-IV criteria of the
documentation of MDD. One method was a computerized decision support system,
developed for this study, named the CaseWalker; and the other method was a paper
checklist. A secondary research objective was to study user acceptance of the
CaseWalker.

Although computer reminder systems have been studied extensively in medical
ambulatory care settings, no published randomized clinical trials of their use in mental
health settings were found. None of the systematic reviews or the either of the meta-
analyses of computerized reminder systems discussed earlier (Austin, et al., 1994; Balas
et al., 1996; Hunt et al., 1998; & Shea, et al., 1996) identified reminder studies conducted
in a mental health setting. A PubMed search using "decision support," "randomized
clinical trial," and either "mental health" or "psychiatry" yielded no relevant citations.
Thus, this study was a systematic replication of previous research in a novel clinical
setting. Further, the study provided additional data on the relative effectiveness of
computer and manual reminder systems, a question concerning which the literature is not
conclusive (e.g., Shea, et al., 1996; Tape & Campbell, 1993). Comparison of computer
and manual reminders is of practical significance for VHA because manual systems are so widely employed in the agency for meeting similar documentation requirements.

The MDD/PTSD module (see Appendix A) of the MDD clinical practice guideline was the target of the study, and the PTSD Clinical Team (PCT), an outpatient clinic for patients with PTSD, was chosen as the study site. Design, clinical and practical considerations all led to the decision to study the MDD/PTSD module in the PCT. The design consideration was to avoid confounding the effects of the experimental methods with administrative pressures to comply with performance measures. The performance measure for doing mood disorder screening was monitored by the medical center administration and quality management staff only in Primary Care clinics and not in the PCT. The performance measure that requires documentation of the DSM-IV MDD criteria met did apply to the PCT. However, no effort was made (outside of the present study) to comply with this requirement in any outpatient mental health clinic at the Salt Lake VA Medical Center during the study period.

The clinical consideration was that there is a high incidence of MDD in patients with PTSD (see review by Fairbank, Schlenger, Saigh & Davidson, 1995). In a community sample, Kessler, Sonnega, Bromet, Hughes and Nelson (1995) found that 48% of people with PTSD had comorbid MDD. In a nontreatment seeking sample of Vietnam veterans with PTSD, Kulka et al. (1991) found a 20% lifetime rate and a 16% current rate for MDD. Among treatment-seeking veterans, Keane and Wolfe (1990) found that 68% of veterans with PTSD also had MDD. Thus, the high incidence of MDD that was expected among PCT patients, all of whom are diagnosed as having PTSD prior to referral to the clinic, increased the opportunity to compare the two experimental
methods in their effectiveness in documenting the DSM-IV diagnostic criteria for MDD that these patients met.

Finally, the PCT was chosen as the study site for several practical reasons. First was the willingness of the PCT staff to participate in the study, especially the willingness of the Chief of the PCT to oversee the assignment of cases to experimental conditions. Second, the PCT has permanent staffing, so the logistical and statistical problems associated with the rotation of residents and interns were avoided. Third was the location of the clinic, which was within a few feet of the investigator and his staff. The fourth reason was the availability of hardware and a Local Area Network (LAN) to support the computer arm of the study.
CHAPTER II

METHODS

Research Participants

The four research participants were senior clinicians employed by the PCT at the Salt Lake VA Medical Center. One was a clinical psychologist, one was a registered nurse, one was a social worker, and one was an addiction therapist. They have an average of 16 years of professional experience and have worked with the PCT for an average of 2 years.

Study Site

The study was conducted at the PCT, which is an outpatient program within the Mental Health Care Center at the Salt Lake VA Medical Center. The PCT specializes in the assessment, diagnosis, and treatment of veterans who present with PTSD. Staffing consisted of one and a half psychiatrists, two psychologists, two social workers, one registered nurse, one addiction therapist, and various trainees in psychiatry, psychology, and social work. The PCT has a caseload of 400 patients, with about 150 new patients coming to the clinic annually. In 1998, the PCT had 8,300 outpatient visits.

Measures

The following variables were used to assess the relative efficacy of the CaseWalker and the paper checklist conditions:
1. The first measure was the proportion of cases screened for mood disorder. For cases in the CaseWalker condition, the occurrence of screening was determined by checking the CaseWalker database. In the paper condition, screening was documented on the paper checklist (see Appendix B).

2. The second measure was the proportion of cases for which the diagnosis of MDD was documented accurately according to DSM-IV diagnostic criteria. As with the screening measure, this variable was determined by checking the CaseWalker database for cases in that arm of the study. In the paper checklist condition, whether the patient met DSM-IV criteria was based on a computer scoring of the criteria checked on the paper checklist. Note that this measure focuses on the accuracy of documentation rather than on the relative incidence of MDD in the two arms of the study. The incidence of MDD was predicted to be equivalent in the two groups because patients were randomly assigned to the groups.

3. The third measure was the proportion of cases for whom the initial evaluation (i.e., mood disorder screening and the assessment of MDD) was documented in the medical record. In the CaseWalker condition, the requirement was that the CaseWalker generated progress note had to be located in either the paper or the electronic medical record. In the paper checklist condition, the checklist had to be located in the paper medical record.

4. The clinical practice guideline requires that patients with MDD be evaluated after six weeks of treatment to determine their response to therapy. To be subject to this requirement, two criteria had to be met. First, the patient had to have a diagnosis of MDD. Second, the patient had to have at least one visit to the PCT 6 weeks or more
after the second visit (i.e., the visit by which the diagnosis should have been made).

The fourth outcome measure was the proportion of criteria-concordant cases whose response to treatment was documented. This measure was determined from either the CaseWalker database or the paper checklist, depending on experimental condition.

5. Following completion of the study, the clinicians were interviewed individually by the principal investigator to elicit their assessment of the two methods and to evaluate their attitudes toward the clinical practice guideline. A semistructured interview form was used (see Appendix C).

**CaseWalker**

The decision support system for the MDD guideline was dubbed the “CaseWalker,” after the recommendation of the guideline developers that staff learn the logic of the guideline by “walking” cases through it. The CaseWalker leads users through the algorithm based on answers given to questions derived from the guideline rules. Although a probabilistic decision support model might support some of the material in the guideline annotations, most of the implementation emphasis in VHA has been on the performance measures that are derived from the rule-based flow charts of the guidelines.

The CaseWalker comprised four major components: (1) a database, (2) a reminder program, (3) a rule-based expert system, and (4) a graphical user interface (GUI) that stepped the user through the guideline. Each component will be described in more detail below. The application was designed to run on PC’s using the Windows NT operating system.
Database. An Interbase database (Inprise Corporation, 1998) contained three tables (see Appendix D). The patient table (tblPatient) contained a record for each PCT patient and was used by the PCT for administrative functions outside the scope of this study. One field (DOCOMPUTER) in this table identified the experimental condition to which the patient was assigned, and another (ADVOCATE) identified the counselor to whom the patient was assigned. The staff table (tblAdvocate) contained a record for each staff member, and a unique staff identifier (ABREV) in this table was linked to the other two tables. Status table (tblStatus) fields included patient (SSN) and staff (ABREV) identifiers linked to those respective tables as well as a binary large object (BLOB) field (STATUS) that contained the status of each hypothesis in the inference engine for a patient at the end of the previous CaseWalker session.

Reminder Program. The reminder program was always running. The Windows start-up menu initiated the program, but the program did not appear on the Windows task bar and so could not be terminated by the novice user. The reminder program initialization file set the time of day at which the reminder was to occur (the default value used in this study was 8:00 AM). The program read the system clock every minute to determine whether that time of day had passed. If so, and if no reminder had been given yet on that date, the reminder program initiated the CaseWalker GUI. Thus, the CaseWalker was launched under either of the following conditions: (1) if the PC was booted after the reminder time, or (2) if the reminder time occurred while the PC was on. In terms of the taxonomy of computerized reminder systems of Murrey et al. (1992), the CaseWalker was an intervisit reminder system in that it reminded the clinician of actions that were due and should be completed at the next visit.
Inference engine. Expert/C++ (formerly Nexpert), a rule-based expert system from Neuron Data, was used for this project. The knowledge base of the expert system consisted of a rule for each decision node of the MDD/PTSD module. For each rule, there was a hypothesis that could take on one of three conditions, viz., true, false, or not known. Rules could be linked by making the state of one hypothesis a condition (or "left-hand side condition") of another rule. The inference engine was an event-driven program that could process rules by either forward or backward chaining. In this project, forward chaining was always used to process rules. In other words, the user always began with hypothesis one. The result of the evaluation of the initial hypothesis then determined which rule would be tested next, and so on until a terminal hypothesis was reached.

Graphical User Interface (GUI). The GUI was a Delphi 1.0 16-bit program. At the time we began development of the program, PCT staff all had PC’s that ran Windows 3.11 for Workgroups. These machines were replaced before the study concluded with newer Pentium-processor machines running Windows NT. The GUI was linked to the expert system by means of object linking and embedding (OLE) automation, i.e., the expert system was not actually embedded in the GUI but was operated on by the GUI by means of function calls. Delphi function calls were used to pass information regarding the status of objects within the expert system to the GUI and to pass “events” from the GUI to the expert system.

When the GUI was initiated (either by the reminder program or directly), the first thing that was determined was the name of the user by checking the program’s initiation file on that user’s PC. Then, the program checked the Interbase database to determine
whether any patients assigned to that clinician required any action. Two types of action were possible, viz., either initial screening and MDD assessment, or 6-week progress review. If either action was required, a reminder window was displayed that contained a tabbed control. If any patient required initial screening and MDD assessment, the GUI began with the tab that listed those patients (see Figure 1). If there were no patients needing initial assessment but there were patients needing 6-week follow-up, the GUI began with the tab listing those patients (see Figure 2). If no action was required for any of the patients assigned to the clinician, he/she received a pop-up message to that effect.

![Patient lists for Bob Feldman](image)

Figure 1. Reminder for patients requiring initial screening and assessment.
Figure 2. Reminder list for patients requiring 6-week review.

The user had several options available on the “Start Patients” tab. If the patient had not yet been seen (often several weeks lapsed between the initial assignment of a patient to a clinician and the clinician’s first appointment with the patient), the user could click on the “Set a new start date” button. Clicking that button caused a calendar to appear, which could be used to easily reset the start date (see Figure 3). No further reminders for that patient would be shown until that date. Clicking on the “Postpone for 7 days” button removed the selected patient’s name from the list for a week. If the patient were no longer in treatment, clicking on the button with that label terminated reminders for that patient. Finally, if the user chose to do nothing at that time, he/she
could exit the CaseWalker GUI to terminate the session. With the exception of resetting the starting date, the same options were available on the “Review Patients” tab.

Clicking on the “Process now” button (or double-clicking on the highlighted patient’s name) on either the “Start Patients” tab or the “Review Patients” tab began guideline processing for the selected patient. If the patient had not been evaluated before using the CaseWalker, the session began with the “PTSD Assessment” tab of the tracking form (see Figure 4). This tab contained a slightly modified representation of the first page of the MDD/PTSD module (see Appendix A). The clinical practice guideline assumes that the patient has already been screened for mood disorder before being
Does the patient meet hospitalization criteria?

Figure 4. Initial assessment and evaluation form.

referred to the MDD/PTSD module, so the first node of the flow chart had to be modified to begin with the mood disorder screening (compare Appendix A, Box 1, with Figure 4).

Color was used to visually represent progress through the guideline. The current node of the guideline was yellow, previously traversed nodes were blue, and nodes not reached were white. Thus, in the example in Figure 4, the mood disorder screen for fictitious patient John Patient was positive and he met DSM-IV criteria for both MDD and PTSD. Currently, the clinician is being asked whether John Patient meets criteria for hospitalization. Note that on the left side of the form, the previous result and the current question were listed. Questions could be answered by clicking on the “Yes” or “No” buttons or by clicking on the “Y” or “N” next to the current decision box on the flow
chart. If there is an annotation for the current node of the guideline, that annotation could be accessed by clicking on the "Guideline Annotation" button. If there is a checklist for the current node, it could be accessed by clicking on the checklist button, the label of which changed dynamically to indicate the checklist currently available. In the example in Figure 4, clicking on the checklist button would bring up a checklist of hospitalization criteria. There were checklists for each decision node on the first tab, viz., for mood disorder screening (see Figure 5), for MDD diagnostic criteria (see Figure 6), for PTSD diagnostic criteria, and for hospitalization criteria.

![Assessment Tools](image)

**Figure 5.** Checklist for initial screening for mood disorder
Figure 6. Checklist for MDD DSM-IV criteria

The MDD diagnostic criteria checklist (Figure 6) popped up automatically when the user reached the MDD decision box. Thus, the user was required to use the checklist in making the diagnosis of MDD, and the diagnostic criteria were scored automatically by the Delphi program.

All of the steps on the “PTSD Assessment” tab were completed in the initial session for a patient. After six weeks, assuming the patient had both MDD and PTSD, a reminder that a follow-up evaluation was due would appear on the reminder form (see Figure 2). Clicking the “Process Now” button on that form would start a CaseWalker session for that patient that would begin on the “PTSD Outpatient” tab of the guideline tracking form. Figure 7 shows an example for the fictitious patient Jane Public. The user
Figure 7. Six-week review form.

indicated the patient’s response to treatment by clicking on one of the four response options, viz., clearly improved, somewhat improved, no improvement, or in remission. Selecting a response option then completed that session.

At the conclusion of a CaseWalker session, the clinician was given a chance to review the progress note that had been generated during the session (see Figure 8). Each user entry during the session concatenated appropriate text to the progress note. The content of these text strings came from the rules in the expert system. The progress note could be edited by the clinician. Then, the note could be printed and filed in the paper medical record. Clicking on the “Copy for TIU” button copied the progress note to the
7. Feelings of worthlessness or excessive or inappropriate guilt

8. Diminished ability to think or concentrate

9. Recurrent thoughts of death, recurrent suicidal ideation without a specific plan, or a suicide attempt or a specific plan for committing suicide

The symptoms do not meet criteria for a Mixed Episode. The symptoms cause clinically significant distress or impairment in social, occupational, or other important areas of functioning. The symptoms are not due to the direct physiological effects of a substance (e.g., a drug of abuse, a medication) or a general medical condition (e.g., hypothyroidism). The symptoms are not better accounted for by bereavement, i.e., after the loss of a loved one. The symptoms persist for more than 2 months or are characterized by marked functional impairment, morbid preoccupation with worthlessness, etc.

The patient also meets the following criteria for the diagnosis of Major

Figure 8. Progress note generated by CaseWalker

Windows clipboard, from which it could be pasted to an electronic progress note in the hospital information system.

Software Developers

Writing the CaseWalker software was a team effort. Dale Cannon, Ph.D., did most of the design work. Delphi programming was begun by Jeff Sells, Ph.D., and was completed by Robert Feldman. The inference engine knowledge base was developed by all three (i.e., Drs. Cannon and Sells and Mr. Feldman).
**Procedure**

Consecutive admissions to the caseloads of the four participating clinicians between January 5, 1998, and October 7, 1998, were assigned randomly to one of two experimental conditions, viz., the CaseWalker or the paper checklist. The Chief of the PCT assigned 108 patients newly referred to the clinic to one of the four research participants on a rotational basis that was not random. One consideration in assigning a new patient was the current caseload of each clinician. Sometimes patient characteristics determined clinician assignment. For example, a female patient with PTSD secondary to sexual trauma would be assigned to the female therapist. Even though the assignment of cases across clinicians was not random, potential bias of comparisons of reminder system types was controlled for by the random assignment of cases to experimental condition within clinician. Randomization was based on a table of random numbers with two constraints intended to equalize the number of cases on each clinician’s caseload that was assigned to each experimental condition. The first constraint was that no more than three consecutive cases assigned to a given clinician could be assigned to the same experimental condition. The second constraint was that for every 10 cases assigned to a given clinician, five had to be assigned to each condition.

The number of cases per condition assigned to each research participant is shown in Table 1. As can be seen, three of the four clinicians were assigned 29 cases each, and the fourth was assigned 21. Within each clinician, no more than one more patient was assigned to one condition than to the other, and across clinicians exactly the same number of patients (i.e., 54) was assigned to each condition.
Table 1

Number of patients in each experimental condition by sample, experimental condition, and clinician.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Experimental Condition</th>
<th>Clinician A</th>
<th>Clinician B</th>
<th>Clinician C</th>
<th>Clinician D</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>All cases entered into the study</td>
<td>CaseWalker</td>
<td>15</td>
<td>14</td>
<td>10</td>
<td>15</td>
<td>54</td>
</tr>
<tr>
<td></td>
<td>Paper</td>
<td>14</td>
<td>15</td>
<td>11</td>
<td>14</td>
<td>54</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>29</td>
<td>29</td>
<td>21</td>
<td>29</td>
<td>108</td>
</tr>
<tr>
<td>Cases seen for at least two visits</td>
<td>CaseWalker</td>
<td>13</td>
<td>10</td>
<td>5</td>
<td>9</td>
<td>37</td>
</tr>
<tr>
<td></td>
<td>Paper</td>
<td>11</td>
<td>13</td>
<td>7</td>
<td>8</td>
<td>39</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>24</td>
<td>23</td>
<td>12</td>
<td>17</td>
<td>76</td>
</tr>
</tbody>
</table>

Because it was not always possible to do a complete initial assessment, including screening for mood disorder, during the first session with a patient, only those cases seen by the assigned clinician at least twice were included in the data analyses. A total of 83 cases met the two-visit-minimum criterion, but five cases that were supposed to be in the CaseWalker group were dropped from the data analyses because inadvertently the CaseWalker procedure was never initiated for these cases.\(^1\) Two additional cases in the paper checklist condition were dropped because their paper medical records could not be

---

\(^1\) The flag in the patient table of the CaseWalker database that triggers the reminders was never set for 5 cases in the CaseWalker condition. Two cases each had been assigned to Clinicians A and D, and one had been assigned to Clinician B.
found. A Chi Square analysis indicated no significant difference in the number of cases per clinician and experimental condition in the remaining sample of 76 cases that had at least two visits ($\chi^2 = .90, p = .83$).

Of the 76 cases included in the data analyses, 73 (96%) were male, and the mean age was 51.8 years (SD = 10.5, range = 25-75). Four (5%) served in the military during the World War II era, 6 (8%) during the Korean era, 50 (66%) during the Vietnam era, 2 (3%) during the Post-Vietnam era, and 14 (18%) during the Persian Gulf War era. Fifty-five (72%) were combat veterans, but none were former prisoners of war. Forty-eight (63%) were married, 22 (30%) were divorced, 4 (5%) had never married, 1 (1%) was separated, and the marital status of 1 (1%) was unknown.

**Data Analyses**

Analysis of variance (ANOVA) (Cohen & Cohen, 1975) was used to analyze the results. Independent variables were experimental condition (i.e., CaseWalker vs. paper checklist), clinician, and the experimental condition by clinician interaction. Dependent variables were the absence or presence of the indicated measure, scored as 0 (absent) or 1 (present). If the clinician effect were significant, post-hoc comparisons of clinicians were made using the Tukey test (Wilkinson, 1998). The effect size of significant group differences was determined by means of the $f$ statistic (Cohen, 1998) using a computer program (Bornstein & Cohen, 1988). Cohen (1988) designated an $f$ of .10 as small, .25 as medium, and .40 as large.

---

2 One case each had been assigned to Clinicians A and D.
ANOVA was used instead of nonparametric statistics such as chi square or the Fisher Exact Test for the following reasons. First, parametric statistics are intrinsically more powerful than nonparametric statistics (Cohen & Cohen, 1975). Second, the ANOVA factorial design partitions the variance due to clinician and clinician by method interaction from the variance due to method. Although the clinician effect and the interaction term are fixed effects that are not generalizable to other groups of clinicians, controlling for these effects further increases the power of the analysis of group effects over chi square analyses, in which all these effects would be confounded.

Power analyses (Cohen, 1988) were conducted using a computer program (Borenstein & Cohen, 1988) to estimate the effect size required to achieve power of .80 with the Type I error rate set to .05 and the sample size equal to 76. An f (i.e., effect size) of .335 would be required for the group effect to be significant at the .05 level and power equal to .80. This effect size would be between medium and large (Cohen, 1988).

Only 18 cases met the criteria for 6-week reviews. Both criteria for a 6-week review being required (i.e., a diagnosis of MDD and at least one visit to the PCT 6 weeks or more after the second visit) contributed to the small number of cases available for a 6-week review. As will be seen below, only 35 of the 76 cases with at least two visits to the PCT had a diagnosis of MDD. Of those 35, only 18 continued in treatment for 6 weeks or more. Even a large effect size (i.e., .40) would have a power of only .27 with alpha set to .05. Thus, the 6-week review data was not analyzed statistically. Likewise, the interview data, which were based on only four clinicians, were not analyzed statistically.
CHAPTER III

RESULTS

Screening for Mood Disorder

A total of 57 of the 76 cases seen for at least two visits were screened for mood disorder (i.e., 75%). An ANOVA in which the occurrence/non-occurrence of screening was the dependent variable was significant for both the group, $F(1, 68) = 5.90, p = .018$, and staff effects, $F(3, 68) = 3.02, p = .036$. The $f$ statistic for the group effect was .27, which indicates a medium effect size (Cohen, 1988). The group by clinician interaction was not significant. As shown in Figure 9, more cases in the CaseWalker condition (86.5%) were screened for mood disorder than in the paper checklist condition (64.1%). Tukey post-hoc tests indicate that Clinician C screened a smaller proportion of his cases across conditions than did either Clinician B or D ($p < .05$) (see Figure 10).

Documentation of the Diagnosis of MDD

The documentation of MDD diagnostic criteria met by experimental condition across all 76 cases is shown in Figure 11. The clinicians identified 46% of patients in each condition as having MDD (CaseWalker, 17/37; paper checklist, 18/39). In all 17 of the CaseWalker cases said to have MDD, the DSM-IV criteria met were documented. In the paper checklist condition, in only one of 18 cases (5.6%) said to have MDD were the criteria documented. An ANOVA across the 35 patients said to have MDD in which the dependent variable was whether the diagnosis of MDD was correctly documented was
Figure 9. Percent cases screened for mood disorder by experimental condition.

Figure 10. Percent cases screened for mood disorder by clinician, independent of method.
significant for the group effect, $F(1, 27) = 170.9, p < .001$. The $f$ statistic for the group effect was 2.83, which is substantially higher than the .40 that Cohen (1988) designates a large effect size.

**Medical Record Documentation**

The previous measures looked at all sources of evidence that mood disorder had been screened and the DSM-IV criteria for MDD had been followed. These sources included the CaseWalker database and paper checklists not filed in the medical record as well as the electronic and paper medical records. The present measure was limited to documentation in the electronic and paper medical records. Thus, the prior measures indicated whether the clinician complied with guideline requirements, and the present measure assessed whether these activities were documented in the official medical record.

**Figure 11.** Documentation of Major Depressive Disorder criteria by experimental condition.
Across all 76 cases, the required documentation appeared in the medical record only 44.7% of the time. Documentation rate by experimental condition is shown in Figure 12. Although the rate was higher in the CaseWalker condition (54%) than in the Paper condition (35.9%), this difference was not significant. An ANOVA in which the presence or absence of required documentation was the dependent variable was significant only for the clinician effect, $F(3, 68) = 4.88, p = .004$. Documentation rates by staff are shown in Figure 13. Tukey post-hoc tests of the clinician effect indicated Clinician B placed less documentation in the chart than did either Clinician A or D.³

![Figure 12](image.png)

**Figure 12.** Percent cases with documentation in medical record by experimental condition.

³ One research participant (Clinician B) kept 11 completed paper checklists in “shadow files” at his desk rather than in the medical record. He gave these checklists to the principal investigator at the end of the study. These cases were counted as being screened and the accuracy of MDD diagnoses was determined. However, they were not counted as being documented in the medical record. If these cases had been counted as meeting both documentation requirements, Clinician B would have met the requirement in 100% of his cases.
Figure 13. Percent cases with documentation in medical record by clinician, independent of experimental condition.

Six-week Reviews

Only 18 cases met both criteria for a 6-week review (i.e., a diagnosis of MDD and an outpatient visit at least 6 weeks after the second visit), which is too few to permit meaningful statistical analyses. However, a few observations may be made about the results, which are presented in Table 2. (Clinician C had no eligible cases in either experimental condition and so is omitted from this discussion.) Neither Clinician A nor Clinician D reviewed any of their eligible cases in the checklist condition, but Clinician B reviewed all of his (recall that Clinician B is the one who kept the checklists in his office, whereas the other clinicians kept their checklists in the medical record). However, Clinician B also reviewed all eligible cases in the CaseWalker condition.
Table 2

Six-week review of response to treatment by clinician and experimental condition.

<table>
<thead>
<tr>
<th>Clinician</th>
<th>Condition</th>
<th>Cases (N)</th>
<th>Number</th>
<th>% Reviewed Reviewed</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>CaseWalker</td>
<td>5</td>
<td>2</td>
<td>40</td>
</tr>
<tr>
<td></td>
<td>Paper</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>B</td>
<td>CaseWalker</td>
<td>2</td>
<td>2</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>Paper</td>
<td>6</td>
<td>6</td>
<td>100</td>
</tr>
<tr>
<td>C</td>
<td>CaseWalker</td>
<td>0</td>
<td>0</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>Paper</td>
<td>0</td>
<td>0</td>
<td>N/A</td>
</tr>
<tr>
<td>D</td>
<td>CaseWalker</td>
<td>0</td>
<td>0</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>Paper</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>All</td>
<td>CaseWalker</td>
<td>7</td>
<td>4</td>
<td>57</td>
</tr>
<tr>
<td></td>
<td>Paper</td>
<td>11</td>
<td>6</td>
<td>55</td>
</tr>
</tbody>
</table>

Interview Findings

Two of the four research participants correctly described the study as a comparison of the two experimental methods. The other two described it as a clinical study of the comorbidity of MDD and PTSD. All four correctly stated that the MDD clinical practice guideline required them to screen all patients for mood disorder and to document the MDD diagnostic criteria met. Three strongly agreed with the statement that “it is clinically worthwhile to screen patients for mood disorder.” The fourth agreed with the statement but expressed concern that screening might result in the over-diagnosis
of MDD. He did not believe that MDD was underdiagnosed in the PCT. All four strongly agreed with the statement that "it is clinically worthwhile to document the MDD criteria met." Three stated that the CaseWalker was definitely the easier method to use for screening for mood disorder, and the fourth said that the paper checklist was somewhat easier for this purpose. Three thought the CaseWalker was definitely the easier method to use for documenting the MDD criteria met, and one thought there was no difference between the two methods for this purpose. Three had a definite overall preference for the CaseWalker, and one had an equally strong overall preference for the paper checklist. The one who preferred the paper checklist explained that the CaseWalker was easier for her to ignore because she could just close the reminder notice without doing anything but that it was harder for her to ignore the paper forms on her desk. All four strongly agreed with the statement that the CaseWalker reminder system was helpful. One volunteered that he thought it was the program's strongest feature. Three agreed with the statement that the CaseWalker reminder system was annoying, but one strongly disagreed with the statement. Those who found it annoying said they did not like the fact that the reminder popped up while they were doing other things on the computer, but one also said this annoyance was probably necessary. Two of the research participants indicated they almost always used the pop-up form for screening mood disorder, and the other two said they never used it. All could correctly describe what they were to do with the CaseWalker generated progress note, and none thought it was difficult to either print the note and file it or to cut and paste it to the electronic medical record. Only one could describe correctly how to determine whether a patient required a 6-week progress review.
CHAPTER IV

DISCUSSION

Relative Effectiveness of the CaseWalker and the Paper Checklist

This study compared the effectiveness of an automated decision support system and a paper checklist in supporting the implementation of two primary objectives of the MDD clinical practice guideline, viz., to increase screening for mood disorder and to improve the documentation of the diagnosis of MDD. For both objectives, the automated decision support system was shown to be more effective than the manual system.

The difference in screening rates was of practical as well as statistical significance in that only the CaseWalker screening rate (i.e., 86.5%) was greater than the 80% rate required by the guideline performance measure. Although meeting the performance measure administrative standard may be of no clinical or scientific significance, it is of importance to VHA. This study suggests that computerized reminders may be more likely to achieve the standards that the agency has set for itself.

The CaseWalker also improved the completeness of the documentation of the diagnosis of MDD. In both experimental conditions, clinicians indicated that approximately the same percentage of cases (i.e., about 46%) met DSM-IV criteria for MDD. However, the diagnosis of MDD was documented fully in only one case in the paper checklist condition. The same documentation error was made in every case. The diagnosis of MDD requires first that one or more episodes of Major Depressive Episode (MDE) be documented, and second that alternative explanations of the MDE (e.g.,
Schizo-Affective Disorder, normal bereavement) be ruled out. In all but one case in the paper condition in which the clinician stated the patient had MDD, alternative explanations for MDE were not ruled out. The CaseWalker, on the other hand, automatically required the clinician to evaluate the presence/absence of the rule out conditions when MDE was established, so failure to document the rule out criteria never occurred. The 100% rate of accurately documenting the diagnosis in the CaseWalker condition is a very significant achievement from a quality assurance perspective. The finding that the diagnosis was fully documented in the CaseWalker condition strongly supports the conclusion that an automated decision support system improves the accuracy of following complex algorithms such as the MDD diagnostic criteria.

This study may be the first to observe in a mental health setting the superiority of a computer reminder system over a paper checklist placed in the medical record. Thus, the present findings are a systematic replication of the findings of Tape and Campbell (1993) as well as those of Frame et al. (1994) with preventive health reminders in ambulatory care settings.

The finding in this study that computer reminders are superior to manual reminders contrasts with the report of Shea et al. (1996) of no significant difference between computer and manual reminder systems. Failure to find a significant effect, of course, is not the same thing as demonstrating that the effect does not exist. There are now three independent randomized clinical trials, in quite different clinical settings, that have found an advantage for computer reminder systems. Although further research would be required to specify the conditions under which the two types of reminders do
and do not differ, it seems safe to conclude at this point that they do differ across a range of clinic settings and types of practice guidelines.

The present study does not demonstrate that manual reminders are ineffective, simply that they are less effective than computer reminders. The experimental design did not include a no-reminder control condition, which would have been necessary to assess the efficacy of the manual system. It was not possible to have a no reminder control condition because of the VHA mandate that the MDD guideline be implemented and the fact that a paper checklist was the accepted standard means of complying with the mandate.

From a statistical perspective, the effect on screening for mood disorder of the computer reminder, relative to the manual reminder, was of medium magnitude (Cohen, 1988). The effect size for complete documentation of DSM-IV criteria for MDD was of large magnitude. These medium and large effect sizes were observed in spite of an experimental design that may have attenuated group differences. Hunt et al. (1998) point out that a design in which individual clinicians are observed under multiple experimental conditions may underestimate experimental effects because the clinician’s experience with one arm of the study may generalize to the other. For example, in the present study, if the CaseWalker made clinicians more aware of the requirement to screen patients for mood disorder and that awareness increased their use of the checklist, then any differential effect the CaseWalker may have had on mood screening would have been attenuated.

A study of computer generated physician reminders by Chambers, Balaban, Carlson and Grasberger (1991) provides evidence that there are interactions between
experimental conditions when clinicians experience multiple arms of a study. In the Chambers et al. study, physicians were randomized to one of three groups that differed in the percentage of patients for whom the physicians received computer-generated reminders concerning annual influenza vaccination. The complete reminder group received reminders for all patients, the no reminder group received reminders for no patients, and the partial reminder group received reminders for half of the patients. Thus, physicians in the partial reminder group experienced the no reminder condition for half their patients and the complete reminder condition for the other half. Consistent with the findings of most studies, patients whose physicians were in the complete reminder group were vaccinated more often (51%) than those patients whose physicians were in the no reminder group (30%). Of those cases in the partial reminder group for whom reminders were given, 38% were vaccinated, which is more than the 30% in the control condition but less than the 51% in the complete reminder condition. Those cases in the partial reminder group for whom a physician reminder was not generated were vaccinated only 20% of the time, which is lower than the 30% in the control condition. Thus, it appears that partial reminders had a negative effect on physician compliance in the partial reminder group regardless of whether reminders were given or not. This finding supports the possibility that there may be some interaction between the CaseWalker and paper checklist conditions in the present study. The fact that there was a significant difference between the CaseWalker and the paper checklist in spite of the conservative bias of the design increases confidence that it is a replicable finding.

The algorithm for diagnosing MDD is more complex than the requirement that one screen new patients for mood disorder. This difference in guideline complexity may
account for the greater differential effect of the CaseWalker for the two measures (Grilli & Lomas, 1994). It seems as though the requirement to screen new patients is about as simple as a guideline gets, and thus it is a guideline for which a paper checklist may have some beneficial effect. However, guidelines can be much more complex than even the algorithm for diagnosing MDD. For example, VHA has just published a performance measure for the recently-released psychoses clinical practice guideline that reads as follows: “Percentage of patients who have a Global Assessment of Functioning score of 40 or less, which is unchanged or worsened during the past 12 months, who have had a change in antipsychotic pharmacological therapy.” It is unlikely that a paper checklist would help clinicians very much with this performance measure, but a computerized decision support system might. Thus, as the number of guideline performance measures in VHA and their complexity increase, computerized decision support systems will be essential. In fact, the VHA committee that sets software development priorities for the agency has made automated decision support for clinical practice guidelines one of its top priorities (R. M. Kolodner, personal communication, January 26, 1999).

Neither experimental method resulted in satisfactory documentation of mood screening and the diagnosis of MDD in the medical record. There was a nonsignificant difference in favor of the CaseWalker (i.e., 54% vs. 35.9%). Even if this difference were significant, it still would not be impressive because even 54% is woefully inadequate. Clearly, screening and diagnoses should be documented in the medical record 100% of the time. The low documentation rate in the CaseWalker condition may have been a consequence of the difficulty in getting the computer-generated progress note into the medical record. When the study began, staff members were not entering any progress
notes in the electronic medical record. Midway through the study, electronic progress notes began to be required throughout the Mental Health Care Center. It was possible to “cut and paste” the CaseWalker progress note into the electronic medical record, but only one clinician did so with any consistency. The other three all could describe how to do it, but they almost never did so in spite of saying it was not too hard to do. The other option with CaseWalker progress notes was to print them and file them in the paper medical record. The three staff who did not “cut and paste” the note into the electronic medical record printed and filed them in less than half the cases.

Two problems diminished the documentation rate in the paper checklist condition. First, staff simply never filled out the form. The clinic clerk always placed the form in the paper chart when a patient was assigned to the checklist condition, but staff frequently failed to fill it out (e.g., only 64% filled out the initial screening form). The second problem, which occurred with one clinician, was the presence of “shadow” medical files. This clinician completed the paper checklist for almost all of the cases in the checklist condition that were assigned to him, but he kept the checklists in files in his desk rather than in the patients' medical records. He explained that the paper medical record was too difficult to locate, so he just kept his own files in his office. The existence of such “shadow” medical records is a common occurrence with paper charts. The practical problem with them, of course, is that data in the shadow record are unavailable to other clinicians treating the patient. Further, shadow files are not part of the system of legal medical records.

There was no difference between the CaseWalker (57%) and the checklist (55%) in compliance with the 6-week review of response to treatment for patients with MDD.
(see Table 2). This guideline requirement is one for which the CaseWalker was expected to have a higher compliance rate than the paper checklist because of the challenge of remembering when the review was due. Even though too few cases met the criteria for the 6-week review to make reliable statistical inferences, some speculation may be offered. First, the method of reminding clinicians of the need for a 6-week review in the CaseWalker probably was ineffective. That reminder appeared on the second tab of a tabbed control, and most staff could not describe where to find that reminder. There probably should be one list that shows both the patients needing screening and the patients needing 6-week reviews, with some visual representation of which requirement was due. Second, the effectiveness of clinical reminders might be enhanced if the system were integrated with the appointment system for the electronic medical record. Finally, the only clinician to do any reviews of cases in the paper checklist condition was the one who kept the checklists in his office. It is likely that a paper checklist with a follow-up reminder on the back page is pretty ineffective, especially if the checklist is in a paper chart that is hard to obtain.

Although the sample of clinicians is too small to generalize the results to other clinicians, the results suggest user acceptance of the CaseWalker was good. Three of the four clinicians in the study reported a strong overall preference for the CaseWalker over the checklist. The one clinician who preferred the checklist gave an unexpected explanation, i.e., that the checklist was harder to ignore. This clinician is, by her own admission and by informal observation, the clinician who is least comfortable with using a computer. For example, in spite of the fact that she has a new, high-performance PC on her desk with the latest word processing software, she prefers to use an old electric
typewriter. However, her preference for the paper checklist apparently had little to do with her aversion to using computers. Rather, she found it too easy to reschedule CaseWalker sessions by clicking on the “Postpone for 7 days” button. Uncompleted paper checklists sitting on her desk served as more effective reminders for this clinician.

Critique of the CaseWalker

The results of this study encourage further development of an automated decision support system for clinical practice guidelines. However, the study also suggests ways in which such a system might be improved.

First, the scope of the decision support system should be limited to those aspects of the clinical practice guideline considered to be most important by clinicians and guideline developers. When the CaseWalker was first designed, the intent was to support every decision node in the entire guideline. Thus, the entire flowchart of the MDD/PTSD module was represented graphically, and the clinician had to respond to a question for each decision node. Extension of the CaseWalker to the rest of the MDD guideline would require substantial additional programming to represent the other four modules of the guideline, and the clinician data entry burden would increase proportionately. However, since the original design of the CaseWalker, VHA has operationally defined what are considered the most important aspects of the guideline by publishing a brief set of performance measures. Simply stated, all that is going to be monitored with respect to the MDD clinical practice guideline is whether patients are screened for mood disorder, whether the diagnosis of MDD is justified properly, and whether patients with MDD are given Global Assessment of Functioning ratings. Presumably, these performance measures index the aspects of the guideline with greatest clinical significance. Whether
that assumption is correct or not, the performance measures are the aspects of the
guideline that command greatest administrative attention because of the incentives (both
positive and negative) associated with them. Supporting these three measures would be
much simpler for both program developers and for clinicians than trying to support the
entire guideline.

Further, limiting the decision support system to performance measures would
facilitate supporting other clinical practice guidelines as well. Since publication of the
MDD guideline, VHA has published a clinical practice guideline for the treatment of
schizophrenia and other psychoses, and a guideline covering the treatment of substance
use disorders is almost ready for release. It would be extremely difficult from an
application development standpoint to include every decision node in all of these
guidelines in a decision support system. Even if we were successful in doing so, the
number of reminders that would be generated and the data entry burden would be so
aversive that clinician compliance probably would be very low.

The second major lesson learned from the CaseWalker is that the insertion into
the electronic medical record of progress notes generated by the decision support system
needs to be simplified. The infrastructure to support the direct insertion of progress notes
into the electronic medical record now exists within VHA. VHA is installing personal
computers in all patient care areas because a GUI to electronic records is being adopted
for clinicians nationwide. At the Salt Lake VA Medical Center, PC's already are in each
mental health clinician's office. Finally, mental health clinicians now are expected to
write all their progress notes using the electronic medical record, and the new GUI
facilitates progress notes writing.
The third major lesson learned is that a database in the hospital information system should replace the Interbase database on the LAN. The data generated by the decision support system need to be in the electronic medical record in computable form for medico-legal and quality assurance reasons as well as for clinical reasons. For example, there is a diagnostic code for indicating that a screening for mood disorder has been done (i.e., ICD-9 V Code 79.0) that could be stored in the file in which diagnoses are filed. If a screening for mood disorder using the decision support system resulted in a diagnosis of V79.0 being sent to this diagnoses file, the occurrence of the screening would be known to other applications that generate statistical reports of guideline compliance based on information available in the electronic medical record. The current Interbase database is, and always would be, an electronic “shadow” file outside the official electronic medical record. Using such a database expedited the initial development of the CaseWalker, but a database that is part of the hospital information system must be used in the future.

**Future Development using V^ISTA Clinical Reminders**

The clinical information system in use throughout VHA is known as the Veterans Health Information Systems and Technology Architecture, or V^ISTA. V^ISTA includes a Clinical Reminder utility that is used to provide reminders regarding health maintenance schedules. Reminders generated by this utility are passed to the V^ISTA Computerized Patient Record System (CPRS), which is the GUI for the V^ISTA electronic medical record (see Figure 14).
The Clinical Reminder utility is a good vehicle for implementing the lessons learned from the CaseWalker project. It is capable, with planned extensions described below, of providing reminders for mental health clinical practice guidelines. Also, it is the official VHA decision support system, so deployment throughout VHA is not a problem and it does not utilize “shadow” electronic medical records. Clinical reminders can be modified easily by individual medical centers to reflect local conditions and practice patterns.

A current limitation for support of any clinical practice guideline by the Clinical Reminder utility is that it is simply a reminder system, and actions directly based on reminders are not possible. For example, clicking on the flexible sigmoidoscopy
reminder in Figure 14 would not enable the user to order the procedure, enter the results, or write a progress note. Instead, the user would simply see a pop-up message with a very brief explanation of the reminder. In the flexible sigmoidoscopy example, the message says, “No flexisigmoidoscopy "CPT" or "ICD O/P" on file. Final Frequency and Age Range used: 5 years for ages 50 and older.” However, the Clinical Reminder utility and the CPRS GUI are being modified to permit actions from reminders, including the generation of a progress note.

A second limitation of the Clinical Reminder utility for supporting mental health clinical practice guidelines is that data in the V*STA Mental Health Package (MHP) are not available to the Clinical Reminder program. The Mental Health Informatics Section is working with Chief Information Office Technical Services developers to develop application program interfaces (API’s) to make MHP data available to the Clinical Reminder package. These APIs will make it possible, for example, to develop clinical reminders based on whether psychological tests have been administered within a specified time frame or to develop reminders based on psychological test scores.

An application based on these extensions of the Clinical Reminder utility is being developed to support the MDD clinical practice guideline. The planned application will replace the CaseWalker with a better product that can be deployed easily throughout VHA. The planned application will work as follows: A reminder will search several V*STA files to determine whether a patient has been screened for mood disorder within the previous year. It will look for the mood disorder screening diagnosis (ICD-9 V79.0) in the diagnoses file, and it will look for the existence of any of a set of mood disorder screening tests in the MHP psychological test file. If none of these conditions are met, a
reminder will appear on the CPRS cover sheet. Clicking on this reminder will bring up a form containing the mood disorder screening test recommended by the clinical practice guideline. The clinician then can enter the patient’s responses to the four screening questions and click the “Done” button. Then, three things will happen: (1) ICD-9 V79.0 will be stored in VISTA as a diagnostic code associated with that visit, (2) the results of the screening test will be stored in the MHP psychological testing file, and (3) a progress note will be generated that the clinician can edit and sign while in CPRS. A second reminder will be based on a positive screening result. Clicking on this reminder will bring up a form with DSM-IV diagnostic criteria for MDD. When the clinician has completed this evaluation, the criteria will be scored automatically and a positive finding will be saved to the VISTA diagnoses file. Additional text, automatically generated based on the findings, will be inserted in the progress note.

A further advantage of using the VISTA Clinical Reminder utility is that it contains a report utility based on reminders. For any given reminder, it is possible to generate a list of patients in a given clinic or assigned to a given clinician who have or have not had the reminder satisfied. Also, it is possible to generate summary reports that indicate the number and percentage of patients who have satisfied the reminder. These reports can be generated at the clinician, clinic, or medical center level. Thus, they should be useful not only to individual clinicians as retrospective feedback on their guideline compliance but as quality assurance and management tools to administrators.

Clinician Attitudes toward the MDD Guideline

It has been established in the literature on guideline compliance that clinician involvement in the development of guidelines increases their compliance with those
guidelines (e.g., Horne, 1996; Karuza, et al., 1995; Tierney, Overhage, & McDonald, 1996). Further, Mendelson (1986) found that even a computerized reminder system was not effective if physicians did not agree with the guideline it was supporting. In this study, the participating clinicians had no involvement in the development of the MDD guideline, the performance measures were imposed from VHA Headquarters, and it was not known whether the clinicians would agree with the guideline. In spite of these potentially negative factors, the post-study interviews with the clinicians indicated they did agree with the value of screening patients for mood disorder and with the need to document diagnoses accurately. This buy-in on the part of the clinicians undoubtedly was critical to the high rate of compliance with the screening and documentation requirements.
Major Depressive Disorder
Post Traumatic Stress Disorder with MDD
Assessment

1. Patient referred with positive results on brief screen for both PTSD and Mood disorder symptoms. (Pt. is negative for SWA)

2. Pt. meets DSM-IV criteria for MDD? [A]
   - Y
       - Y
           - Y
             7. Go to: Box #11 Mental Health Specialty Module
           - N
             8. Meet criteria for hospitalization? [C]
               - Y
                 10. Initiate outpatient therapy for MDD/PTSD [E]
               - N
                 9. Identify and prioritize target symptoms in outpatient setting [D]

     - N
       4. Treat for PTSD w/o aggregate features

   - N
     5. Re-Evaluate Return to appropriate level for treatment of mood disorder or mild PTSD

   - Y
       - Y
         7. Go to: Box #11 Mental Health Specialty Module
       - N
         8. Meet criteria for hospitalization? [C]
           - Y
             10. Initiate outpatient therapy for MDD/PTSD [E]
           - N
             9. Identify and prioritize target symptoms in outpatient setting [D]

8. Meet criteria for hospitalization? [C]
   - Y
     10. Initiate outpatient therapy for MDD/PTSD [E]
   - N
     9. Identify and prioritize target symptoms in outpatient setting [D]

Continued on page 2
Major Depressive Disorder
Post Traumatic Stress Disorder with MDD
Outpatient Setting

11. Monitor treatment every 1-2 weeks

12. Assess response (week 6) [F]

13. Clearly improved?
   - Y: Continue for 6 more weeks
   - N: Somewhat improved?
      - Y: Continue treatment with appropriate adjustment
      - N: Still willing to treat at current setting?
         - Y: Augment or change treatment
         - N: Refer to General Psychiatry

20. Monitor every 1-2 weeks [G]

21. Assess response at 12 weeks

22. Clearly improved?
   - Y: Follow-up treatment maintenance, prevention of recurrence and patient & family education
   - N: Continue in appropriate setting [H]

24. Refer to General Psychiatry
Major Depressive Disorder
Post Traumatic Stress Disorder with MDD
Inpatient Treatment

Continued from page 1

26 Admit to inpatient care

27 Identify and prioritize target symptoms

28 Initiate treatment for MDD/PTSD

29 Assess for improvement in target symptoms and GAF after 1-2 weeks

30 Does patient meet discharge criteria?

Y

31 Reassess and modify treatment as necessary. If not improved in 4 weeks reassess by case conference or consultation

N

32 Does patient meet discharge criteria after 2 more weeks?

Y

33 Initiate pt. and family education

34 Discharge to outpatient status with follow-up appointment within 2 weeks at Vet center or Mental Health Clinic

N

Consider referral to appropriate alternate level of inpatient/residential care

M

L

N
Brief Screen for Mood Disorder for Patients under Age 60

1. Have you felt depressed for one or more days during the past week?
   a. Yes
   b. No

2. In the past year, have you had two consecutive weeks or more during which you felt sad, blue or depressed; or when you lost all interest or pleasure in things that you usually cared about or enjoyed?
   a. Yes
   b. No

3. Have you had two years or more in your life when you felt depressed or sad most days even if you felt okay sometimes?
   a. Yes
   b. No

4. Have you felt depressed or sad much the time in the past year?
   a. Yes
   b. No

Office use only

This mood disorder screen was positive/negative.

Staff Signature: ______________________________

Patient Name: ______________________________

SSN: ___________________________ Date: ________________
**Brief Screen for Mood Disorder for Patients Aged 60 or above**

For each of the following, please indicate how often you felt that way during the past week, using the following ratings. Circle the letter (A, B, C, or D) that best describes how you felt.

1. I felt that I could not shake off the blues even with help from my family or friends.

<table>
<thead>
<tr>
<th>A Rarely or none of the time (less than 1 day)</th>
<th>B Some or a little of the time (1-2 days)</th>
<th>C Moderately or much of the time (3-4 days)</th>
<th>D Most or almost all the time (5-7 days)</th>
</tr>
</thead>
</table>

2. I felt depressed.

<table>
<thead>
<tr>
<th>A Rarely or none of the time (less than 1 day)</th>
<th>B Some or a little of the time (1-2 days)</th>
<th>C Moderately or much of the time (3-4 days)</th>
<th>D Most or almost all the time (5-7 days)</th>
</tr>
</thead>
</table>

3. I felt fearful.

<table>
<thead>
<tr>
<th>A Rarely or none of the time (less than 1 day)</th>
<th>B Some or a little of the time (1-2 days)</th>
<th>C Moderately or much of the time (3-4 days)</th>
<th>D Most or almost all the time (5-7 days)</th>
</tr>
</thead>
</table>

4. My sleep was restless.

<table>
<thead>
<tr>
<th>A Rarely or none of the time (less than 1 day)</th>
<th>B Some or a little of the time (1-2 days)</th>
<th>C Moderately or much of the time (3-4 days)</th>
<th>D Most or almost all the time (5-7 days)</th>
</tr>
</thead>
</table>

5. I felt hopeful about the future.

<table>
<thead>
<tr>
<th>A Rarely or none of the time (less than 1 day)</th>
<th>B Some or a little of the time (1-2 days)</th>
<th>C Moderately or much of the time (3-4 days)</th>
<th>D Most or almost all the time (5-7 days)</th>
</tr>
</thead>
</table>

Office use only

This mood disorder screen was positive/negative.

Staff Signature: ____________________________

Patient Name: ________________________________

SSN: ____________________________ Date: ____________________
DSM-IV Diagnostic Criteria for Major Depressive Disorder

Initial all criteria that apply.

**Major Depressive Disorder**

- Presence of either a single (296.2) or multiple (296.3) Major Depressive Episode(s), respectively (see below).
- The major depressive episode(s) is/are not better accounted for by Schizoaffective Disorder and is/are not superimposed on Schizophrenia, Schizophreniform Disorder, Delusional Disorder, or Psychotic Disorder Not Otherwise Specified.
- There has never been a Manic Episode, a Mixed Episode, or a Hypomanic Episode. Note: This exclusion does not apply if all of the manic-like, mixed-like or hypomanic-like episodes are substance or treatment induced or are due to the direct physiological effects of a general medical condition.

**Major Depressive Episode**

- At least five of the following symptoms have been present during the same 2-week period, nearly every day, and represent a change from previous functioning. At least one of the symptoms must be either (1) depressed mood or (2) loss of interest or pleasure:
  - depressed mood most of the day, nearly every day, as indicated by self or others.
  - markedly diminished interest or pleasure in all, or almost all, activities.
  - significant weight loss or weight gain (5%/mo.) or loss/gain in appetite nearly every day.
  - insomnia or hypersomnia nearly every day.
  - psychomotor agitation or retardation nearly every day (as noted by others).
  - fatigue or loss of energy nearly every day.
  - feelings of worthlessness or excessive or inappropriate guilt nearly every day.
  - diminished ability to think or concentrate or indecisiveness nearly every day.
  - recurrent thoughts of death, recurrent suicidal ideation without a specific plan, or a suicide attempt or a specific plan for committing suicide.

Patient Name: __________________________ SSN: ____________ Date: ________
DSM-IV Diagnostic Criteria for Major Depressive Disorder (cont.)

_____ The symptoms caused clinically significant distress or impairment in social, occupational, or other important areas of functioning.

_____ Symptoms are not better accounted for by a Mixed Episode, Mood Disorder Due to a General Medical Condition, a Substance-Induced Mood Disorder, or Bereavement (normal reaction to the death of a loved one).

_____ Symptoms are not better accounted for by a Psychotic Disorder (e.g., Schizoaffective Disorder).

The patient meets/does not meet DSM-IV criteria for Major Depressive Disorder.

Staff Signature: ___________________________ Date: ____________

Patient Name: ________________ SSN: ____________ Date: ___________
Clinician Interview
CaseWalker Study

Clinician: ____________________________________________

Years professional experience: ____________________________

Years with PCT: ____________________________

1. What is your understanding of the study?

2. What is your understanding of the MDD CPG? What part applies to you?

3. In your opinion, it is clinically worthwhile to screen patients for mood disorder.

<table>
<thead>
<tr>
<th>Strongly disagree</th>
<th>Disagree</th>
<th>No opinion</th>
<th>Agree</th>
<th>Strongly agree</th>
</tr>
</thead>
</table>

4. In your opinion, it is clinically worthwhile to document the MDD criteria met.

<table>
<thead>
<tr>
<th>Strongly disagree</th>
<th>Disagree</th>
<th>No opinion</th>
<th>Agree</th>
<th>Strongly agree</th>
</tr>
</thead>
</table>

5. Which study method was easier to use for screening?

<table>
<thead>
<tr>
<th>CW definitely</th>
<th>CW somewhat</th>
<th>No difference</th>
<th>Paper somewhat</th>
<th>Paper definitely</th>
</tr>
</thead>
</table>

6. Which study method was easier to use for documenting MDD criteria?

<table>
<thead>
<tr>
<th>CW definitely</th>
<th>CW somewhat</th>
<th>No difference</th>
<th>Paper somewhat</th>
<th>Paper definitely</th>
</tr>
</thead>
</table>
7. Which study method did you prefer overall?

<table>
<thead>
<tr>
<th>CW definitely</th>
<th>CW somewhat</th>
<th>No difference</th>
<th>Paper somewhat</th>
<th>Paper definitely</th>
</tr>
</thead>
</table>

8. The CW reminder system was helpful.

<table>
<thead>
<tr>
<th>Strongly disagree</th>
<th>Disagree</th>
<th>No opinion</th>
<th>Agree</th>
<th>Strongly agree</th>
</tr>
</thead>
</table>

9. The CW reminder system was annoying.

<table>
<thead>
<tr>
<th>Strongly disagree</th>
<th>Disagree</th>
<th>No opinion</th>
<th>Agree</th>
<th>Strongly agree</th>
</tr>
</thead>
</table>

10. How did you screen for mood disorder in the CW condition (did you use the pop-up form?)

11. What are you supposed to do with the PN that CW generates? How do you do it?

12. Using the CW, how do you find out when a patient is due for a 6-week progress review?

13. What suggestions do you have for improving the CW?
APPENDIX D

CASEWALKER DATABASE
### Structure of the CaseWalker Interbase database.

<table>
<thead>
<tr>
<th>Table</th>
<th>Field</th>
<th>Type</th>
<th>Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>tblPatient</td>
<td>SSN</td>
<td>VARCHAR</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>LNAME</td>
<td>VARCHAR</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td>FNAME</td>
<td>VARCHAR</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td>MI</td>
<td>VARCHAR</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>GENDER</td>
<td>CHAR</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>ADVOCATE</td>
<td>CHAR</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>MD</td>
<td>CHAR</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>DD214</td>
<td>SHORT</td>
<td></td>
</tr>
<tr>
<td></td>
<td>PSF</td>
<td>SHORT</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ADMITDATE</td>
<td>DATE</td>
<td></td>
</tr>
<tr>
<td></td>
<td>DOB</td>
<td>DATE</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SC</td>
<td>VARCHAR</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td>TESTS</td>
<td>SHORT</td>
<td></td>
</tr>
<tr>
<td></td>
<td>POW</td>
<td>SHORT</td>
<td></td>
</tr>
<tr>
<td></td>
<td>WAR</td>
<td>VARCHAR</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>REFDATE</td>
<td>DATE</td>
<td></td>
</tr>
<tr>
<td></td>
<td>COMMENTS</td>
<td>VARCHAR</td>
<td>50</td>
</tr>
<tr>
<td></td>
<td>CHARTED</td>
<td>SHORT</td>
<td></td>
</tr>
<tr>
<td></td>
<td>DOCOMPUTER</td>
<td>SHORT</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>VETCNTR</td>
<td>SHORT</td>
<td></td>
</tr>
<tr>
<td>Table</td>
<td>Field</td>
<td>Type</td>
<td>Length</td>
</tr>
<tr>
<td>------------------</td>
<td>----------</td>
<td>----------</td>
<td>--------</td>
</tr>
<tr>
<td>tblPatient (cont)</td>
<td>ACTIVESTATUS</td>
<td>SHORT</td>
<td></td>
</tr>
<tr>
<td></td>
<td>STARTED</td>
<td>SHORT</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>APPTDATE</td>
<td>DATE</td>
<td></td>
</tr>
<tr>
<td>tblAdvocate</td>
<td>ABREV</td>
<td>VARCHAR</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>NAME</td>
<td>VARCHAR</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>STATUS</td>
<td>SHORT</td>
<td></td>
</tr>
<tr>
<td>tblStatus</td>
<td>SSN</td>
<td>VARCHAR</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>DOB</td>
<td>DATE</td>
<td></td>
</tr>
<tr>
<td></td>
<td>STATUS</td>
<td>BLOB</td>
<td></td>
</tr>
<tr>
<td></td>
<td>MODULE</td>
<td>VARCHAR</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>REVIEWDATE</td>
<td>DATE</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ABREV</td>
<td>VARCHAR</td>
<td>5</td>
</tr>
</tbody>
</table>
REFERENCES


Veterans Health Administration. (1996). *Clinical Guideline for Major Depressive Disorder (MDD), MDD with Posttraumatic Stress Disorder (PTSD), and MDD with Substance Abuse (SA).* Washington, DC: Author.


CURRICULUM VITAE

Name: Dale Sherman Cannon

Birthplace: Upland, California

Birthdate: September 24, 1942

Education:

<table>
<thead>
<tr>
<th>Institution</th>
<th>Degree</th>
<th>Year</th>
<th>Major</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pepperdine University</td>
<td>B.A.</td>
<td>1964</td>
<td>Psychology</td>
</tr>
<tr>
<td>Pepperdine University</td>
<td>M.A.</td>
<td>1966</td>
<td>Psychology (Clinical)</td>
</tr>
<tr>
<td>University of Utah</td>
<td>Ph.D.</td>
<td>1970</td>
<td>Psychology (Clinical)</td>
</tr>
</tbody>
</table>

Internship: 1968-1969: Salt Lake VA Medical Center

Licensure: Licensed Psychologist, State of Utah

Professional Experience:

1994-Present        Associate Chief for Informatics, Mental Health Strategic Health Group
1980-1994           Chief, Psychology Service, Dallas VA Medical Center
1978-1980           Chief, Alcohol Dependence Treatment Program, Salt Lake VAMC
1972-1980           Coordinator, Psychology Training Committee, Salt Lake VAMC
1971-1978           Clinical Research Psychologist, Alcohol Dependence Treatment Program, Salt Lake VAMC
1969-1971           Project Director, Behavior Systems Corporation

Publications:


Abstract:


Books and Book Chapters:
