HSR&D

Evidence Synthesis for Determining the Responsiveness of Depression Questionnaires and Optimal Treatment Duration for Antidepressant Medications

Executive Summary

October 2009

Prepared for:

Department of Veterans Affairs
Veterans Health Administration
Health Services Research
& Development Service
Washington, DC 20420

Prepared by:

Durham Veterans Affairs Medical Center/Duke Evidence-based Practice Center
Durham, NC

Investigators:

John W. Williams Jr., MD, MHS
Professor of Medicine and Psychiatry, Durham VA Medical Center and Duke University

Director, Evidence-Based Practice Center

Durham, NC

Monica Nora Slubicki, MD
Resident in Psychiatry,
VA Administrative Chief Resident 2009-2010,
Duke University Medical Center

Durham, NC

Damon S. Tweedy, MD
Medical Director, Primary-Care Mental Health Integration, Durham VA Medical Center

Consulting Associate, Duke University Medical Center

Durham, NC

Daniel W. Bradford, MD, MPH
Assistant Professor of Psychiatry, Durham VA Medical Center and Duke University Medical Center

Director, Psychosocial Rehabilitation and Recovery Center, Durham VAMC

Durham, NC

Ranak B. Trivedi, PhD
AHRQ Postdoctoral Fellow in Health Services Research,
Durham VA Medical Center and Duke University

Durham, NC

Dana Baker, MS
Clinical Research Coordinator
Duke University

Durham, NC

PREFACE

VA's Health Services Research and Development (HSR&D) Service works to improve the cost, quality, and outcomes of health care for our nation's veterans. Collaborating with VA leaders, managers, and policy makers, HSR&D focuses on important healthcare topics that are likely to have significant impact on quality improvement efforts. One significant collaborative effort is HSR&D's Evidence-based Synthesis Program (ESP). Through this program, HSR&D provides timely and accurate evidence syntheses on targeted health care topics. These products will be disseminated broadly throughout VA and will: inform VA clinical policy, develop clinical practice guidelines, set directions for future research to address gaps in knowledge, identify the evidence to support VA performance measures, and rationalize drug formulary decisions.

HSR&D provides funding for four ESP Centers. Each Center has an active and publicly acknowledged VA affiliation and also serves as an Evidence Based Practice Center (EPC) supported by the Agency for Healthcare Research and Quality (AHRQ). The Centers will each generate three evidence syntheses annually on clinical practice topics of key importance to VHA leadership and policymakers. A planning committee with representation from HSR&D, Patient Care Services (PCS), Quality Enhancement Research Initiative (QUERI), Office of Quality and Performance (OQP), and the VISN Clinical and Quality Management Officers, has been established to identify priority topics and key stakeholder concerns and to ensure the quality of final reports. Comments on this evidence report are welcome and can be sent to Susan Schiffner, ESP Program Manager, at Susan.Schiffner@va.gov.

This information is distributed solely for the purposes of pre-dissemination peer review. It has not been formally disseminated by the Department of Veterans Affairs. It does not represent and should not be construed to represent a Department of Veterans Affairs determination or policy.

Financial disclosure: No investigators have any affiliations or financial involvement (e.g., employment, consultancies, honoraria, stock ownership or options, expert testimony, grants or patents received or pending, or royalties) that conflict with material presented in the report.

EXECUTIVE SUMMARY

BACKGROUND

According to projections from the World Health Organization, depression will be the second leading cause of disability in the developed world by 2020. Primary care clinicians care for approximately two thirds of depressed individuals. In 2000, the U.S. economic burden of depressive disorders was estimated to be 83.1 billion dollars. This included 31% direct medical costs, 7% suicide-related mortality costs, and 62% workplace costs. A variety of strategies have been tested to improve patient outcomes. Among these, integrated care models have emerged as both effective and cost effective. A recent systematic review identifies symptom monitoring as a key element of these integrated care models. However, the review did not identify the standardized depression scales that are responsive to clinically important change.

A separate but important issue raised by Veterans Administration (VA) Stakeholders is how long to continue antidepressant medication for patients who respond to acute phase treatment. Clinical guidelines recommend continuation treatment for 4-6 months for uncomplicated major depression and some national performance measures are linked to these guidelines. However, clinical guidelines for longer-term maintenance phase treatment are more variable and performance indicators (e.g., Healthcare Effectiveness Data and Information Set, HEDIS) do not address maintenance phase treatment. A better understanding of the evidence for long-term treatment efficacy with antidepressants would inform guidelines and performance measurement.

The Key Questions (KQ) were:

KQ1: In patients with major depressive disorder treated in primary care settings, what assessment tools are responsive to change? This review should specifically address instruments that are feasible for the primary care setting.

KQ2: In primary care patients with major depressive disorder who remit with antidepressant medication, what is the minimum treatment duration to decrease the risk of relapse or recurrence? This review will focus on patients without comorbid substance abuse, PTSD, psychosis or other conditions where guidelines would recommend specialty based care.

METHODS

We searched PubMed from 1950-2009 using standard search terms; PsychInfo was also searched for key question one (KQ1). Additional citations were identified from reference lists. Titles, abstracts, and articles were reviewed in duplicate by physicians trained in the critical analysis of literature. For KQ1, we included primary literature comparing one of the 6 eligible depression symptom questionnaires to an interview-based reference standard. For key question two (KQ2), we searched for and identified a high quality systematic review, then searched for relevant randomized trials published since the original review (2007-2009). For eligible articles, data were extracted in duplicate. We evaluated study quality for the primary literature and the systematic review. All data were summarized narratively. An overall strength of evidence "GRADE" was assigned to the body of evidence for each key question.

RESULTS

For KQ1, we screened 743 titles, rejected 661, and performed a more detailed review on 82 articles. From these, we identified 3 unduplicated observational studies meeting eligibility criteria. For KQ2, we screened 154 titles, rejected 139, and performed a more detailed review on 15 articles. From these, we identified 1 recent high quality systematic review and 3 relevant randomized controlled trials (RCTs).

KEY QUESTION 1. In patients with major depressive disorder treated in primary care settings, what assessment tools are responsive to change?

We identified 3 studies evaluating the responsiveness of the Patient Heath Questionnaire-9 (PHQ-9), in primary care patients with depressive disorders; no studies for the other eligible questionnaires were identified. A total of 2,330 patients were evaluated, one study was limited to older adults and one included VA settings.

The most relevant study to VA settings and patients was a high quality secondary analysis from the IMPACT study, a randomized trial comparing collaborative care to usual care. In this study, participants were ≥ 60 years old, had a mean of 3.8 chronic diseases and a research-based diagnosis of major depression or dysthymia. Three of the eighteen primary care sites were VA. The analysis was limited to the 434 patients in the intervention arm with complete assessments at baseline, 3- and 6-month follow-up. Responsiveness was reported as the standardized response mean (SRM) which is calculated as: Mean (time 2) - Mean (time1)/standard deviation of score changes.

For the cohort overall, the mean change and standardized response mean at 3 months was: -7.5±5.8, SRM -1.3 (95% CI -1.4 to -1.2). At 6 months, the mean change and SRM were: -8.0±6.1, SRM -1.3 (95% CI –1.4 to -1.2). Responsiveness equaled or exceeded the longer Symptom Checklist-20 (SCL-20, self-administered 20-item questionnaire measuring depressive symptoms) at these two time points. Results were not significantly different when restricted to patients with major depressive disorder (MDD). For the 317 patients with MDD, an independent, structured diagnostic assessment was used to classify patients at six month follow-up as: persistent MDD, partial- or full remission. Greater clinical improvement was associated with larger reductions in PHQ-9 scores. The mean change and SRM for each group was: persistent MDD -5.6±6.6, SRM -0.8; partial remission -8.4±6.1, SRM -1.4; full remission -9.8±5.9, SRM -1.7. In this analysis, the SRM was again similar for the PHQ-9 and SCL-20. An analysis to determine the minimum clinically important difference (MCID), estimated this value conservatively at 4.78, meaning a 5-point or larger decline in the PHQ-9 indicates clinically meaningful improvement.

Two fair quality studies, conducted with a German language version of the PHQ-9 showed similar results at 3- and 12-month follow-ups. Standardized response means ranged from -1.42 to -2.15 for patients rated as responders by a structured interview. One study conducted subgroup analyses and found similar responsiveness for men and women, different age groups, depression diagnosis and presence or absence of comorbid physical illness.

These three studies differed in a variety of design features that could lead to heterogeneous results including: study quality, questionnaire language, follow-up timing, and participant characteristics. Despite these sources of potential variability, the overall results were consistent across studies. The PHQ-9 is responsive to clinically important changes in symptom status. Using the GRADE criteria, we judged the overall quality of evidence for this finding as moderate. For the finding that the minimum clinically important difference is 5, the quality of evidence is low based on a single, albeit, high quality studies.

A recent literature synthesis identified longitudinal assessment of depression symptoms with a standardized scale as a critical component of effective depression care. The PHQ-9 is the best validated scale in primary care populations, both for initial diagnosis and for detecting response to change. Its routine use for measuring response to treatment could improve patient care and outcomes, but logistical support to integrate the questionnaire into clinical practice would likely be needed to achieve successful implementation.

- The PHQ-9 is the best validated instrument for detecting clinically important response to treatment. Quality of Evidence = Moderate
- A 5 point change on the PHQ-9 is estimated as the minimum clinically important difference. Quality of Evidence = Low

KEY QUESTION 2. In primary care patients with major depressive disorder who remit with antidepressant medication, what is the minimum treatment duration to decrease the risk of relapse or recurrence?

We included 1 applicable high quality systematic review and 3 RCT's with 4 comparisons published since the systematic review. A total of 9,024 patients in 26 RCT's were evaluated. None of the studies included a VA setting; three were restricted to patients age ≥ 65 years old.

The systematic review evaluated 23 fair quality RCT's comparing second-generation antidepressant to placebo in fully- or partially-remitted patients. Patients with comorbid psychiatric or serious medical conditions were generally excluded. Twelve took place in unspecified outpatient clinics, four in primary care and psychiatry clinics, and the remaining seven did not specify the setting. Relapse or recurrence was generally defined using a predefined score on the Hamilton Depression Rate Scale (HDRS), a validated, interview-administered depression severity measure. The authors stratified the studies according to treatment duration: less than 1 year after acute phase treatment remission (continuation) and 1 year or more after acute phase treatment duration (maintenance). The unadjusted frequency of relapse for continuation phase (12 studies) was 22% for active treatment and 42% for placebo. In a pooled analysis the relative risk of relapse was 0.54 (95% CI 0.46 to 0.62); heterogeneity was moderate (I2 =47%). The unadjusted frequency of recurrence for maintenance phase (11 studies) was similar to shorter duration studies, 26% with active treatment and 48% with placebo. The relative risk of recurrence was 0.56 (95% CI 0.48 to 0.66); heterogeneity was moderate (I2 =30%). Loss to follow up due to adverse events was not significantly different between antidepressant and placebo. Only one study out of the 23 RCT's randomized patients in remission to varying

durations (14, 38 or 50 weeks) of continuation phase antidepressant or placebo. In that study, relapse rates were significantly lower for patients on active treatment at 14 weeks (26% vs. 49%), and 38 weeks (9% vs. 23%) but not at 50 weeks (11% vs. 16%). In meta-regression analyses, the duration of treatment prior to and after randomization were not associated with the magnitude of treatment effect, suggesting a constant reduction in relative risk.

Of the three additional RCT's identified, the PREVENT study was the most informative. This multi-phase, double-blind, placebo-controlled study evaluated 12 and 24 month treatment with venlafaxine ER versus placebo. It found that venlafaxine ER was associated with a statistically significantly lower recurrence rate at 12-month follow-up (23.1% vs. 42.0%). Using an expanded definition of recurrence, freedom from recurrence at 24 month follow up was 67% for venlafaxine vs. 41.0% for placebo. The 24 month PREVENT follow up phase did not report on patients lost to follow up. Another good quality RCT reported the results of a 24 week RCT of escitalopram (10-20mg/day) versus placebo in older adults who had responded to acute treatment with escitalopram for MDD. Escitalopram was associated with a significantly lower relapse rate compared with placebo (9% vs. 33%, $p<0.001$). The last RCT evaluated was a small, fair quality trial that did not find a significant difference between antidepressant and placebo for prevention of relapse.

The high quality systematic review and 2 of the most recent relevant RCT's provide moderately strong evidence that continued antidepressant treatment decreases the risk of subsequent relapse for patients with MDD who achieve partial- or full-remission. The moderate strength of evidence grade is based on RCT's with some important methodological limitations, generally consistent results, and a precise estimate of effect. Of note, none of these studies were performed in a VA population. The magnitude of risk reduction was similar for shorter- and longer-term trials and maintained for up to 2.5 years. However, these trials do not directly address the question about the minimum duration of continued antidepressant treatment since they report the average risk reduction over these time periods. At the individual patient level, the decision for how long to continue antidepressant treatment should be based on effectiveness, adverse effects and patient preferences. Additional studies that could include decision analyses and randomized trials that stratify treatment duration based on risk factors are needed to inform clinical guidelines and performance measures for maintenance phase treatment.

- A high quality systematic review and 2 of the most recent relevant RCT's provide moderately strong evidence that continued antidepressant treatment decreases the risk of subsequent relapse for patients with MDD who achieve partial- or full-remission. Continued treatment for 1 to 2 years after achieving partial- or full-remission with second-generation antidepressants decreases the risk of relapse or recurrence by almost 50%. The number needed to treat to prevent one relapse was 5. Quality of Evidence = Moderate.

- The magnitude of risk reduction was similar for shorter- and longer-term trials and maintained for up to 2 years. However, these trials do not directly address the minimum duration of continued antidepressant treatment since they report the average risk reduction over these time periods.