Depression Screening and Patient Outcomes in Cardiovascular Care
A Systematic Review

Brett D. Thombs, PhD
Peter de Jonge, PhD
James C. Coyne, PhD
Mary A. Whooley, MD
Nancy Frasure-Smith, PhD
Alex J. Mitchell, MSc, MRCPsych
Marij Zuidersma, MSc
Chete Eze-Nliah, MD, MPH
Bruno B. Lima
Cheri G. Smith, MLS
Karl Soderlund, BS
Roy C. Ziegelstein, MD

Major depressive disorder (MDD) is present in as many as 20% of patients with cardiovascular disease (CVD)1-3 and is associated with adverse cardiovascular outcomes, even after controlling for other risk factors.3-8 In addition, MDD is a chronic, disabling condition that is associated with poor quality of life,9 functional limitations,10 less favorable self-care behaviors,11 and higher health care costs among patients with CVD.12

Several clinical guidelines, including the American College of Cardiology and the American Heart Association's guidelines for ST-elevation myocardial infarction (MI),13 unstable angina/non–ST-elevation myocardial infarction (MI),13 unstable angina/non–ST-

Context Several practice guidelines recommend that depression be evaluated and treated in patients with cardiovascular disease, but the potential benefits of this are unclear.

Objective To evaluate the potential benefits of depression screening in patients with cardiovascular disease by assessing (1) the accuracy of depression screening instruments; (2) the effect of depression treatment on depression and cardiac outcomes; and (3) the effect of screening on depression and cardiac outcomes in patients in cardiovascular care settings.

Data Sources MEDLINE, PsycINFO, CINAHL, EMBASE, ISI, SCOPUS, and Cochrane databases from inception to May 1, 2008; manual journal searches; reference list reviews; and citation tracking of included articles.

Study Selection We included articles in any language about patients in cardiovascular care settings that (1) compared a screening instrument to a valid major depressive disorder criterion standard; (2) compared depression treatment with placebo or usual care in a randomized controlled trial; or (3) assessed the effect of screening on depression identification and treatment rates, depression, or cardiac outcomes.

Data Extraction Methodological characteristics and outcomes were extracted by 2 investigators.

Results We identified 11 studies about screening accuracy, 6 depression treatment trials, but no studies that evaluated the effects of screening on depression or cardiovascular outcomes. In studies that tested depression screening instruments using a priori-defined cutoff scores, sensitivity ranged from 39% to 100% (median, 84%) and specificity ranged from 58% to 94% (median, 79%). Depression treatment with medication or cognitive behavioral therapy resulted in modest reductions in depressive symptoms (effect size, 0.20-0.38; r², 1%-4%). There was no evidence that depression treatment improved cardiac outcomes. Among patients with depression and history of myocardial infarction in the ENRICHD trial, there was no difference in event-free survival between participants treated with cognitive behavioral therapy supplemented by an antidepressant vs usual care (75.5% vs 74.7%, respectively).

Conclusions Depression treatment with medication or cognitive behavioral therapy in patients with cardiovascular disease is associated with modest improvement in depressive symptoms but no improvement in cardiac outcomes. No clinical trials have assessed whether screening for depression improves depressive symptoms or cardiac outcomes in patients with cardiovascular disease.

JAMA. 2008;300(18):2161-2171

©2008 American Medical Association. All rights reserved.
Box. Systematic Review Questions

Key Question 1
What is the accuracy of screening instruments for depression in cardiovascular care populations?

Key Question 2
Is treatment of depression in cardiovascular care patients effective in improving depression?

Key Question 3
Is systematic screening for depression more effective than usual care in identifying patients with depression? Facilitating treatment of depression? Reducing depressive symptoms? Improving cardiac outcomes?

METHODS
Search Strategy
Articles for review were identified from the MEDLINE, PsycINFO, CINAHL, EMBASE, ISI, SCOPUS, and Cochrane databases, which were searched from inception to May 1, 2008. Two searches were conducted: (1) the first sought articles that compared a screening instrument with a valid major depressive disorder criterion standard or that assessed the effect of screening on depression identification and treatment rates, depression, or cardiac outcomes, and (2) the second search was for articles that compared the effects of depression treatment on depression or cardiac outcomes with placebo or usual care in a randomized controlled trial. Search terms are available from the corresponding author (B.D.T.). Manual searching was done on reference lists of included articles, cardiovascular care guidelines,13-16 systematic reviews, and 33 selected journals for the July 2007 to May 1, 2008, time frame. We tracked citations of included articles using Google Scholar24 and surveyed 36 experts, including members of the National Heart, Lung, and Blood Institute Working Group on the Assessment and Treatment of Depression in Patients with Cardiovascular Disease25 and authors of the included treatment studies, to seek unidentified published depression treatment trials and to determine if there were any unpublished trials.

Identification of Eligible Studies
Article eligibility criteria were established a priori. Eligible articles were studies with original data that evaluated patients in cardiovascular care settings that were published in any language. Abstracts, letters, editorials, and case series or case reports were excluded. Only published studies were eligible for the query seeking the accuracy of depression screening instruments (key question 1), since the inclusion of data from unpublished studies would have required additional analyses and interpretation outside the scope of this review. Cardiovascular care was defined based on diagnosis (eg, MI, congestive heart failure) or intervention (eg, coronary artery bypass graft surgery). Studies in which patient selection was based on clinical characteristics other than cardiovascular disease were excluded. When multiple articles were published on the same cohort, the most comprehensive article was included. Studies with mixed populations were included if CVD patient data were reported separately.

Screening studies that assessed the accuracy of depression screening instruments (key question 1) were included if they compared a screening instrument with a valid standard, defined as a Diagnostic and Statistical Manual of Mental Disorders or International Classification of Diseases diagnosis of MDD based on a validated diagnostic interview procedure, and if they reported data allowing determination of sensitivity, specificity, positive predictive value, and negative predictive value. Examples of validated interviews include the Structured Clinical Interview for DSM-IV,26 the Composite International Diagnostic Interview,27 and the Diagnostic Interview Schedule.28

Articles assessing the effect of depression treatment on depression and cardiac outcomes (key question 2) were included randomized controlled trials with placebo or usual care controls that evaluated pharmacological, psychotherapeutic, or other interventions for MDD among patients in cardiovascular care settings. Only studies with MDD diagnosed using a validated psychiatric interview and Diagnostic and Statistical Manual of Mental Disorders or International Classification of Diseases criteria were included.

Eligible articles assessing the effect of screening on depression and cardiac outcomes in patients in cardiovascular care settings (key question 3) included randomized controlled trials and prospective studies that compared depression identification, depression treatment rates, depressive outcomes, or car-
diac outcomes between CVD patients who underwent depression screening and CVD patients who did not undergo screening.

Two investigators reviewed articles for eligibility independently. Translators were used to evaluate non-English titles, abstracts, and articles. If either reviewer deemed an article potentially eligible based on title or abstract review, then a full-text article review was completed. Disagreement between reviewers after full-text review was resolved by consensus. Chance-corrected agreement between reviewers was assessed using the Cohen κ statistic.

**Evaluation of Eligible Studies**

Investigators independently extracted and entered into a standardized spreadsheet relevant data and outcomes. Discrepancies were resolved by consensus. The authors of 1 screening26 and 1 treatment study30 provided data to correct minor inconsistencies in original publications. Authors of 2 studies31,32 provided data not included in the original report, which allowed investigators to calculate results not reported in the corresponding studies. Authors of 5 studies33-35,37 verified whether diagnoses of MDD were conducted blind to screening results. The author of 1 study clarified the threshold used to diagnose depression.36

Quality was based on methods developed by the US Preventive Services Task Force.19,38 Ratings reflected the quality of each study relative to our key questions rather than general quality per se. For the question of whether a study assessed the accuracy of depression screening instruments (key question 1), rated items included the relevance and availability of screening tests, the credibility of the reference standard, whether the diagnosis of MDD was made blind to screening results, the spectrum of patients included, sample size, and the screening test reliability. We rated sample size based on the number of patients with MDD in each study (poor, 0-24; fair, 25-99; good, 100 or more). Studies with fewer than 10 patients with depression were rated poor for overall quality regardless of ratings in other categories. Quality ratings of randomized controlled trials (key question 2) considered the establishment and maintenance of comparable groups, differential or high overall loss to follow-up, clarity of intervention definition, completeness of outcome variables, sample size, and analysis considerations. Study quality was assessed by 2 investigators with discrepancies resolved by consensus.

In studies that assessed the accuracy of depression screening instruments (key question 1), for each screening instrument, sensitivity, specificity, positive predictive value, and negative predictive value with 95% confidence intervals (CIs) are presented.39 Eligible studies were evaluated to determine whether data were sufficiently similar to warrant pooling of results. Substantial heterogeneity between studies was found with respect to cardiovascular diagnoses, criterion standards, screening instruments, and scoring thresholds for depression. Notably, some studies used a priori-defined scoring thresholds from the research literature, whereas others selected sample-specific thresholds based on receiver operating characteristic (ROC) curves. The latter methods tend to yield overly optimistic estimates of screening accuracy that do not replicate consistently.40 For these reasons, it was determined that data pooling was not appropriate.

For studies that assessed the effect of depression treatment on depression and cardiac outcomes (key question 2), in which multiple depression outcomes were reported, designated primary outcomes for each study were given highest priority. Then observed rates were prioritized over self-report measures. Treatment effect sizes are typically reported as the Cohen d statistic,41 which represent a standardized difference between 2 means. Rosenthal et al42,43 have pointed out, however, that these effect size metrics are nonintuitive and often misunderstood. They recommend using correlation (r) metrics that are more easily interpretable. Estimates of g or d and correlation estimates are essentially equivalent and readily converted from one to another.43,44 To facilitate interpretation of treatment effect sizes, both Hedges g and r², the percent of variance in depression change scores due to treatment, are reported. Eligible treatment studies were evaluated to determine the appropriateness of pooling results. Studies that assessed the effect of depression treatment on depression and cardiac outcomes (key question 2) consisted of a small number of heterogeneous studies, each of which used a different therapeutic intervention. Thus, it was determined that data pooling was not appropriate, and studies were reported qualitatively.

**RESULTS**

**Diagnostic Accuracy of Depression Screening Tools in Cardiovascular Care Settings (Key Question 1)**

The literature search for studies that assessed the accuracy of depression screening instruments and the effect of screening on depression and cardiac outcomes in patients in cardiovascular care settings (key questions 1 and 3) yielded 858 unique citations. Of these, 101 were selected for full-text review for key question 1. All but 10 studies33,34,36,37,45-49 were excluded because of ineligible patient populations; lack of a depression screening tool, a structured interview, or both; or lack of data permitting analyses of diagnostic accuracy. One additional study was identified from the references of a systematic review,30 which resulted in 11 articles (FIGURE 1). The κ for interrater agreement was 0.84.

Details of the 11 studies are presented in TABLE 1. Participants in these studies were patients with a recent acute coronary syndrome or coronary revascularization (8 studies), coronary artery disease (1 study),33 hospitalized patients with congestive heart failure (1 study)34 and out-

©2008 American Medical Association. All rights reserved.
patients with congestive heart failure (1 study). Of the 8 studies of patients with recent acute coronary syndrome or revascularization, depression assessments were performed on hospitalized patients (3 studies) and in patients discharged from the hospital 1 to 3 months previously (5 studies). Six studies used prespecified thresholds to define depression, 1 used a threshold based on the upper tertile of scores, and 4 used ROC curve methods to identify thresholds that optimized accuracy. Frasure-Smith et al and Strik et al each reported sensitivity and specificity close to 80% for a Beck Depression Inventory score of 10 or greater in post MI patients. Freedland et al reported good sensitivity (88%), but poor specificity (58%) with the same threshold among hospitalized patients with congestive heart failure. Frasure-Smith et al and Low and Hubley reported good sensitivity (91% and 86%, respectively) and specificity (78% and 89%, respectively) for a Beck Depression Inventory-II score of 14 or greater in outpatients after acute coronary syndrome. Prespecified thresholds performed reasonably well for a 10-item version of the Center for Epidemiological Studies Depression Scale, the Patient Health Questionnaire-9, the Patient Health Questionnaire-2, a 2-item yes/no screening tool, and the Geriatric Depression Scale. However, found that recommended cutoffs from primary care for the Patient Health Questionnaire-2 (≥3) and Patient Health Questionnaire-9 (≥10) resulted in good specificity (92% and 90%, respectively), but poor sensitivity (39% and 54%, respectively) in patients with coronary artery disease.

Stafford et al used sample-specific ROC curve methods and found that a lower threshold of 6 or greater on the Patient Health Questionnaire-9 optimized sensitivity (83%) and specificity (78%) among coronary artery disease outpatients. Sensitivity in studies that used sample-specific ROC curve analyses ranged from 80% to 95%, and specificity from 74% to 85%. Two separate studies performed ROC curve analyses with the 14-item Hospital Anxiety and Depression Scale. The 2 studies reported similar sensitivity (90% and 88%) and specificity (84% and 85%), but an optimal Hospital Anxiety and Depression Scale score was reported to be 13 or greater in one and 17 or greater in the other.

The quality of 4 studies was good, 5 studies were fair, and 2 studies were poor. All studies were rated good for administering an appropriate screening tool and for using a credible reference standard except for 1 study that used the Mini International Neuropsychiatric Interview administered by telephone as its criterion standard. One study was rated poor for administration of a reliable screening test because patients returned screening questionnaires by mail. The delay between Structured Clinical Interview for DSM-IV assessments and questionnaire return was more than 2 weeks in some cases.

In summary, for tests of a priori screening thresholds (including studies in which more than 1 instrument was administered), sensitivity ranged from 39% to 100% (median, 84%), and specificity from 58% to 94% (median, 79%). There were few examples of demonstrated accuracy with screening tools or depression thresholds that were tested in more than 1 sample of cardiovascular care patients. No studies addressed potential harms of screening, including false-positive results, the cost and inconvenience of additional follow-up assessments, the adverse effects or costs associated with treating incorrectly diagnosed patients, or inappropriate labeling.
The Effects of Depression Treatment in Cardiovascular Care Patients (Key Question 2)

For assessing the effect of depression treatment on depression and cardiovascular outcomes (key question 2), there were 861 unique citations and 14 articles selected for full-text review. Six of these met inclusion criteria (Figure 2). No unpublished trials were identified. The κ for inter-rater agreement was 0.86.

Details of the 6 studies are presented in Table 2 and Table 3. There were 4 efficacy studies of antidepressant medications, including 1 each using fluoxetine, sertraline, citalopram, and mirtazapine. Strik et al compared the efficacy and safety of fluoxetine administered to patients after their first MI. The Sertraline Anti-depressant Heart Attack Randomized Trial (SADHART) tested the efficacy and safety of sertraline in patients with unstable ischemic heart disease. The Canadian Cardiac Randomized Evaluation of Antidepressant and Psychotherapy Efficacy Trial (CREATE) was a parallel-group, 2 x 2 factorial trial that compared citalopram to placebo and compared short-term interpersonal psychotherapy plus clinical management to clinical management alone in patients with coronary artery disease. Honig et al compared mirtazapine to placebo for 8 weeks in post-MI patients, then offered open treatment with citalopram in the case of insufficient response. The original report on the mirtazapine trial did not note that open-label treatment was introduced at 8 weeks and presented was 0.86.

Table 1. Summary of Studies of Diagnostic Accuracy of Depression Screening Tools in Cardiovascular Care Settings

<table>
<thead>
<tr>
<th>Source; Setting</th>
<th>Instrument; Cutoff Score</th>
<th>No. of Patients (% Male)</th>
<th>Mean Age, y</th>
<th>Major Depression Criterion Standard</th>
<th>No. (%) With Major Depression</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>Positive Predictive Value</th>
<th>Negative Predictive Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inpatient</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Frasure-Smith et al, 1995, 1998; Canada</td>
<td>BDI ≥ 10&lt;sup&gt;d&lt;/sup&gt;</td>
<td>218 (78)</td>
<td>60</td>
<td>Modified DIS&lt;sup&gt;c&lt;/sup&gt;</td>
<td>33 (15)</td>
<td>82 (69-91)</td>
<td>78 (71-83)</td>
<td>40 (29-52)</td>
<td>96 (92-98)</td>
</tr>
<tr>
<td>Freedland et al, 2003; United States</td>
<td>BDI ≥ 10&lt;sup&gt;d&lt;/sup&gt;</td>
<td>613 (49)</td>
<td>66</td>
<td>Modified DIS&lt;sup&gt;d&lt;/sup&gt;</td>
<td>120 (20)</td>
<td>88 (80-92)</td>
<td>58 (54-63)</td>
<td>34 (29-39)</td>
<td>95 (92-97)</td>
</tr>
<tr>
<td>Dickens et al, 2004; Great Britain</td>
<td>HADS ≥ 17&lt;sup&gt;e&lt;/sup&gt;</td>
<td>314 (63)</td>
<td>58</td>
<td>SCAN</td>
<td>65 (21)</td>
<td>88 (78-94)</td>
<td>85 (80-89)</td>
<td>60 (50-69)</td>
<td>96 (93-98)</td>
</tr>
<tr>
<td>Huffman et al, 2006; United States</td>
<td>2-Items from BDI&lt;sup&gt;g&lt;/sup&gt;</td>
<td>131 (80)</td>
<td>62</td>
<td>SCID-IV</td>
<td>17 (13)</td>
<td>94 (73-92)</td>
<td>76 (68-83)</td>
<td>37 (24-52)</td>
<td>99 (94-100)</td>
</tr>
</tbody>
</table>

| Outpatient     |                           |                          |             |                                     |                             |             |             |                          |                          |
| Gutierrez, 1999; Canada | BDI ≥ 13<sup>b</sup> | 40 (50) | 70 | SCID-IV | 6 (15) | 83 (44-97) | 94 (81-98) | 71 (36-92) | 97 (85-99) |
| Strik et al, 2001; the Netherlands<sup>f</sup> | BDI ≥ 10<sup>a</sup> | 196 (77) | 60 | SCID-IV | 22 (11) | 82 (61-93) | 79 (72-84) | 33 (22-46) | 97 (93-99) |
| McManus et al, 2005; United States | CES-D-10 ≥ 10<sup>d</sup> & PHQ-9 ≥ 10<sup>d</sup> | 1024 (82) | 67 | DIS | 224 (22) | 76 (70-81) | 79 (76-82) | 50 (45-56) | 92 (90-94) |
| McManus et al, 2005; United States | HADS-D ≥ 4<sup>e</sup> & PHQ-2 ≥ 3<sup>e</sup> | 179 (77) | 60 | SCAN | 22 (11) | 96 (78-99) | 74 (67-83) | 32 (22-44) | 99 (96-100) |
| McManus et al, 2005; United States | HADS-D ≥ 4<sup>e</sup> | 179 (77) | 60 | SCAN | 22 (11) | 96 (78-99) | 74 (67-83) | 32 (22-44) | 99 (96-100) |
| Dunlop et al, 2006; the Netherlands | SAD4 ≥ 3<sup>a</sup> | 176 (76) | 60 | SCID-IV | 20 (11) | 95 (76-99) | 68 (60-75) | 28 (18-39) | 99 (95-100) |
| Low and Hubley, 2007; Canada | BDI-II ≥ 14<sup>b</sup> | 119 (75) | 63 | SCID-IV | 7 (6) | 86 (49-97) | 89 (81-93) | 33 (16-56) | 99 (94-100) |
| Stafford et al, 2007; Australia<sup>b</sup> | HADS-D ≥ 6<sup>e</sup> & PHQ-9 ≥ 6<sup>e</sup> | 193 (81) | 64 | MINI | 35 (18) | 80 (64-90) | 72 (68-80) | 49 (37-62) | 95 (90-97) |
| Frasure-Smith et al, 2008; Canada | BDI-II ≥ 14<sup>b</sup> | 804 (81) | 60 | SCID-IV | 57 (7) | 91 (81-96) | 78 (74-80) | 24 (19-30) | 99 (98-100) |

Abbreviations: BDI, Beck Depression Inventory; BDI-II, Beck Depression Inventory-II; CES-D-10, Center for Epidemiological Studies Depression Scale, 10-item version; CI, confidence interval; DIS, Diagnostic Interview Schedule; GDS, Geriatric Depression Scale; HADS, Hospital Anxiety and Depression Scale; MINI, Mini International Neuropsychiatric Interview; PHQ-2, Patient Health Questionnaire-2; PHQ-9, Patient Health Questionnaire-9; SAD4, Symptoms of Anxiety-Deression index; SCAN, Schedule for Assessment of Neuropsychiatric Disorders; SCID-IV, structured clinical interview for the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition; SCL-90-R, depression subscale of the Symptom Checklist 90.

<sup>a</sup> Diagnoses were post-acute myocardial infarction in Frasure-Smith et al; coronary artery disease in McManus et al; post-acute coronary syndrome in Stafford et al; coronary artery disease in Bach et al; congestive heart failure in Gutierrez; and Freedland et al; coronary artery disease in McManus et al; post-acute myocardial infarction, coronary artery bypass graft surgery, or percutaneous transluminal coronary angioplasty in Stafford et al.

<sup>b</sup> Cut-off derived from the literature.

<sup>c</sup> The modified DIS did not require that symptoms be of at least 2 weeks’ duration and did not apply the criteria of seeking medical help and experiencing impairment.

<sup>d</sup> The depression section of the modified DIS starts with somatic rather than cognitive or mood-related symptoms and focuses on current rather than lifetime symptoms.

<sup>e</sup> Cut-off derived from receiver operating characteristic curve.

<sup>f</sup> The number of patients administered each screening tool and diagnostic data were provided by the authors of the original study to correct minor inconsistencies in the published manuscript.

<sup>g</sup> Cut-off derived from upper tertile.

<sup>h</sup> Diagnoses for Stafford et al also include CABG or PTCA.
results for both 8 and 25 weeks. Only data at 8-week follow-up were included for this review. The Honig et al trial was conducted as part of the treatment group of the Myocardial Infarction and Depression-Intervention Trial (MIND-IT). MIND-IT was an effectiveness study because rather than testing the efficacy of a single treatment under optimal conditions, it investigated whether implementing any active treatment strategy resulted in better outcomes compared with usual care. Treatment options included mirtazapine or placebo as part of the double-blind Honig et al trial, open treatment with citalopram, and tailored treatment at the discretion of the treating psychiatrist. The Enhancing Recovery in Coronary Heart Disease Patients (ENRICHD) trial examined the effects of cognitive behavioral therapy plus adjunctive sertraline treatment in the case of insufficient response on depression and cardiac outcomes in post-MI patients.

Only the ENRICHD and MIND-IT studies were designed to assess cardiovascular outcomes, although the MIND-IT study had very low statistical power. Neither found evidence that depression treatment affects cardiac outcomes.

---

### Table 2. Diagnoses and Timing for Randomized Controlled Trials of Pharmacological or Psychotherapeutic Treatment of Depression in Cardiovascular Care Settings

<table>
<thead>
<tr>
<th>Source</th>
<th>Setting</th>
<th>Diagnosis</th>
<th>Assessment</th>
<th>Treatment Duration, wk</th>
<th>Cardiovascular Follow-up</th>
</tr>
</thead>
<tbody>
<tr>
<td>Efficacy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Strik et al, 2000</td>
<td>The Netherlands</td>
<td>post-AMI</td>
<td>3-12 mo</td>
<td>25</td>
<td>25 wk</td>
</tr>
<tr>
<td>Glassman et al, 2002</td>
<td>United States, Canada, Europe, Australia</td>
<td>post-ACS</td>
<td>1st ≤30 d followed by 2nd assessment after 2-wk placebo run-in</td>
<td>24</td>
<td>24 wk</td>
</tr>
<tr>
<td>Honig et al, 2007</td>
<td>The Netherlands</td>
<td>post-AMI</td>
<td>3-12 mo</td>
<td>8</td>
<td>24 wk</td>
</tr>
<tr>
<td>Lesperance et al, 2007</td>
<td>Canada</td>
<td>CAD</td>
<td>NA</td>
<td>12</td>
<td>12 wk</td>
</tr>
<tr>
<td>Effectiveness</td>
<td>van Melle et al, 2007</td>
<td>The Netherlands</td>
<td>post-AMI</td>
<td>3-12 mo</td>
<td>24</td>
</tr>
<tr>
<td>Cardiovascular outcomes</td>
<td>Berkman et al, 2003</td>
<td>United States</td>
<td>post-AMI</td>
<td>≤28 d</td>
<td>26</td>
</tr>
</tbody>
</table>

Abbreviations: ACS, acute coronary syndrome; AMI, acute myocardial infarction; CAD, coronary artery disease; NA, not applicable.

- Denotes timing after occurrence of acute event.
- Denotes duration for cardiovascular events following randomization.
- Denotes the Sertraline Antidepressant Heart Attack Randomized (SADHART) trial.
- Outcomes at 8 weeks were reviewed instead of 24-week results because 8 weeks open treatment with citalopram was offered in the case of refusal or insufficient treatment response.
- Denotes the Canadian Cardiac Randomized Evaluation of Antidepressant and Psychotherapy Efficacy (CREATE) trial.
- Denotes the Myocardial Infarction and Depression-Intervention (MIND-IT) trial.
- Denotes the Enhancing Recovery in Coronary Heart Disease Patients (ENRICHD) trial.
- Maximum duration of the cognitive behavioral therapy intervention was 6 mo. Group therapy could extend 12 additional weeks and adjunctive sertraline treatment for up to 12 mo.
outcomes. Among patients with depression and history of MI in the ENRICHD clinical trial, there was no difference in event-free survival between participants treated with cognitive behavioral therapy supplemented by an antidepressant vs usual care (75.5% vs 74.7%). Cardiac event-free survival in the MIND-IT trial was 86.2% for patients in the treatment group and 87.3% for patients in the control group.

### Table 3. Outcomes for Randomized Controlled Trials of Pharmacological or Psychotherapeutic Treatment of Depression in Cardiovascular Care Settings

<table>
<thead>
<tr>
<th>Source</th>
<th>No. Randomized</th>
<th>Depression Remission</th>
<th>Depression Response</th>
<th>Cardiovascular With Outcome (%)</th>
<th>Depression Primary</th>
<th>Depression Secondary</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Efficacy</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Strik et al,2000</td>
<td>Fluoxetine 27</td>
<td>7 (26)</td>
<td>13 (48)</td>
<td>1 (4)</td>
<td>0.38 (−0.16 to 0.92)</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Placebo 27</td>
<td>4 (15)</td>
<td>7 (26)</td>
<td>6 (22)</td>
<td>0.13 (0.02-1.21)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>2.01 (0.51-7.90)</td>
<td>2.65 (0.84-8.34)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glassman et al,2002</td>
<td>Sertraline 186</td>
<td>NA</td>
<td>125 (67)</td>
<td>32 (17)</td>
<td>0.20 (0.00 to 0.41)</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Placebo 183</td>
<td>NA</td>
<td>97 (53)</td>
<td>41 (22)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.82 (1.19-2.77)</td>
<td>0.72 (0.43-1.21)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Honig et al,2007</td>
<td>Mirtazapine 47</td>
<td>16 (34)</td>
<td>27 (57)</td>
<td>8 (17)</td>
<td>0.35 (−0.06 to 0.77)</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Placebo 44</td>
<td>7 (16)</td>
<td>18 (41)</td>
<td>10 (23)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>2.73 (1.00-7.48)</td>
<td>1.96 (0.85-4.49)</td>
<td>0.70 (0.25-1.97)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lespérance et al,2007</td>
<td>Citalopram 142</td>
<td>51 (36)</td>
<td>75 (53)</td>
<td>6 (4)</td>
<td>0.33 (0.10 to 0.56)</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Placebo 142</td>
<td>32 (23)</td>
<td>57 (40)</td>
<td>6 (4)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.93 (1.14-3.25)</td>
<td>1.67 (1.04-2.67)</td>
<td>1.00 (0.32-3.18)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interpersonal</td>
<td></td>
<td>40 (28)</td>
<td>61 (43)</td>
<td>9 (6)</td>
<td>0.23 (−0.46 to 0.00)</td>
<td>1</td>
</tr>
<tr>
<td>psychotherapy and</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>clinical</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>management 142</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clinical management only</td>
<td></td>
<td>43 (30)</td>
<td>71 (50)</td>
<td>3 (2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>only 142</td>
<td></td>
<td>0.90 (0.54-1.51)</td>
<td>0.75 (0.47-1.20)</td>
<td>3.14 (0.83-11.83)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Effectiveness</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>van Melle et al,2007</td>
<td>Active treatment 209</td>
<td>91/132 (69)</td>
<td>27 (14)</td>
<td>0.12 (−0.15 to 0.39)</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Usual care 122</td>
<td>56/86 (67)</td>
<td>15 (13)</td>
<td>1.10 (0.56-2.16)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.11 (0.62-1.99)</td>
<td>1.10 (0.56-2.16)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cardiovascular</td>
<td>Cognitive behavioral therapy 925</td>
<td>NA</td>
<td>NA</td>
<td>227 (25)</td>
<td>0.22 (0.11 to 0.33)</td>
<td>1</td>
</tr>
<tr>
<td>outcomes</td>
<td>Usual care 901</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>230 (29)</td>
<td>0.96 (0.81-1.17)</td>
<td>0.31 (0.20 to 0.42)</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>128 (14)</td>
<td>129 (14)</td>
<td>0.97 (0.76-1.24)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Abbreviations:** BDI, Beck Depression Inventory; BDI-II, Beck Depression Inventory-II; CGI-I, Clinical Global Impression-Improvement; CGI-S, Clinical Global Impression-Severity; CI, confidence interval; HAMD-17, 17-item Hamilton Depression Rating Scale; HAMD-24, 24-item Hamilton Depression Rating Scale; NA, not applicable; OR, odds ratio; SCL-90-D, depression subscale of the Symptom Checklist 90.

**Point at which treatment and control groups are defined.**

**Baseline HAMD-17 scores, mean (SD) are:** 21.6 (3.6) for Strik et al; 19.6 (5.4) for Glassman et al; 17.8 (not reported) for Honig et al; 22.8 (5.1) for Lespérance et al; and 17.7 (6.4) for Berkman et al.

**Remission defined by Strik et al as a HAMD-17 score lower than 7, by Honig et al as a HAMD-17 score of 7 or lower, by Lespérance et al as a HAMD-24 score of 8 or lower, and van Melle et al did not describe remission criteria. Results shown for van Melle et al reflect the percentage of patients who no longer met International Classification of Diseases, 10th Revision criteria for depressive disorder among those who were assessed (132 patients [83%] in the intervention group and 86 patients [85%] in the usual care group [calculated from original data]).

**Response defined by Strik et al as a 50% or greater reduction in HAMD-17 scores, by Glassman et al as a CGI-I score of 1 or 2 [very much or much improved], by Honig et al as a 50% or greater reduction in HAMD-17 scores or a HAMD-17 score of 9 or lower, by Lespérance et al as a 50% or greater reduction in HAMD-24 scores.**

**Cardiovascular outcomes are cardiac hospitalization for Strik et al; major adverse cardiac events (those involving death or requiring hospitalization) for Glassman et al; hospitalization for Honig et al; cardiovascular serious adverse events (myocardial infarction, congestive heart failure, worsening angina, stroke, or other cardiovascular events) for Lespérance et al; total cardiac events (cardiac death, recurrent myocardial infarction, revascularization, heart failure, myocardial ischemia, and ventricular arrhythmia) [17 patients were lost to follow-up: treatment group, 19; usual care, 11] for van Melle et al; recurrent myocardial infarction or death from any cause and death for Berkman et al.**

**Denotes the Sertraline Antidepressant Heart Attack Randomized (SADHART) trial.**

**Patients were assessed with HAMD-17 at 16 weeks, but not 24 weeks.**

**Denotes the Canadian Depression Randomized Evaluation of Antidepressant and Psychotherapy Efficacy (CREATE) trial.**

**Denotes the Myocardial Infarction and Depression-Intervention (MIND-IT) trial.**

**No. randomized patients in Lespérance et al denotes a 2×2 factorial design with 284 total patients.**

**Denotes the Enhancing Recovery in Coronary Heart Disease Patients (ENRICHD) trial.**

©2008 American Medical Association. All rights reserved.

(Reprinted) JAMA, November 12, 2008—Vol 300, No. 18
Among the 6 treatment studies, 3 were rated as good\textsuperscript{31,60,61} and 3 were rated as fair-good\textsuperscript{30,32,59}. Characteristics that limited quality were nonblind- ing of patients for psychotherapeutic treatments,\textsuperscript{31,32} unequal distribution of confounders,\textsuperscript{30} differential loss to follow-up,\textsuperscript{30,31,59} use of last observation carried forward to impute missing data,\textsuperscript{30,59} and small sample sizes.\textsuperscript{30,59} The MIND-IT study,\textsuperscript{31} which was rated as fair-good overall, was an effectiveness study and therefore not designed to evaluate treatment efficacy.

In summary, effect sizes for treatment of depression in the 4 efficacy trials and the ENRICHD trial were modest. All studies reviewed met or exceeded the 6- to 8-week duration that is typical in acute phase trials of antidepressant agents.\textsuperscript{33,60} However, none continued treatment or assessed follow-up long enough to determine whether antidepressant use reduced the risk of depression relapse or recurrence in cardiac patients who responded to acute treatment.\textsuperscript{65} Only 2 studies had follow-up periods that were long enough to assess cardiac outcomes.\textsuperscript{31,32} Neither found evidence of an effect of depression treatment. Two studies reported that selective serotonin reuptake inhibitors did not affect cardiac function,\textsuperscript{39,60} but no studies assessed other potential harms such as medication adverse effects.\textsuperscript{38}

**Figure 3. Comparison of Effect Sizes for Pharmacological Interventions in Cardiovascular Care and in Studies Registered With the FDA\textsuperscript{68}**

<table>
<thead>
<tr>
<th>Drug</th>
<th>No. of Studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fluoxetine</td>
<td></td>
</tr>
<tr>
<td>Strik et al.,\textsuperscript{59} 2000</td>
<td>5</td>
</tr>
<tr>
<td>Sertraline</td>
<td></td>
</tr>
<tr>
<td>Glassman et al.,\textsuperscript{60} 2002</td>
<td>5</td>
</tr>
<tr>
<td>Mirtazapine</td>
<td></td>
</tr>
<tr>
<td>Hong et al.,\textsuperscript{31} 2007</td>
<td>10</td>
</tr>
<tr>
<td>Citalopram</td>
<td></td>
</tr>
<tr>
<td>Lespérance et al.,\textsuperscript{59} 2007</td>
<td>5</td>
</tr>
</tbody>
</table>

FDA indicates US Food and Drug Administration. The numbers of FDA studies are from published and unpub- lished studies reported by Turner et al\textsuperscript{68} and registered with the FDA.

**Effect of Depression Screening on Outcomes in Cardiovascular Care Patients (Key Question 3)**

Of the 858 unique citations identified in the search for studies that assessed the accuracy of depression screening instruments (key question 1) or the effect of screening on depression and cardiac outcomes in patients in cardiovascular care settings (key question 3), 66 articles were selected for full-text review for key question 3. No articles met eligibility criteria, and no unpublished studies were identified.

**COMMENT**

Several clinical guidelines for cardiovascular care\textsuperscript{13,14,16} recommend that depression be evaluated or that screening for depression be considered in patients with CVD. Whether depression screening is of benefit to patients with CVD is unknown. Our systematic review of the evidence shows that depression screening tools are reasonably accurate in patients with CVD, but there are few examples of screening tools or screening tool thresholds with demonstrated accuracy in more than 1 sample of patients with CVD. There is evidence that depression treatment in patients with CVD improves depression, but the effects on depression are modest with only minimal benefit compared with usual care or placebo. There is no evidence that depression treatment reduces cardiovascular events. No studies have examined whether screening for depression in patients with CVD improves access to depression care or outcomes.

Among the studies that tested depression screening instruments using a priori thresholds, the ranges of sensitivity (39%-100%; median, 84%) and specificity (58%-94%; median, 79%) were similar to those reported in a systematic review of case-finding instruments in primary care in which sensitivity ranged from 50% to 97% (median, 85%) and specificity from 51% to 98% (median, 74%).\textsuperscript{66} However, given the high false-positive rate of screening tools, a clinical interview is necessary to establish a diagnosis of depression. Based on the 15% median prevalence of MDD we identified in the depression screening studies (Table 1), along with the median sensitivity (84%) and specificity (79%) of the depression screening tools, 1000 depression screenings would result in 304 patients needing further evaluation of whom 126 (41% of those who screen positive) would have MDD. Thus, the adoption of depression screening in patients with CVD would consume substantial resources and might identify problems that are not highly amenable to intervention in a cardiovascular care setting.\textsuperscript{67} If antidepressant therapies are prescribed by cardiologists based on a screening tool alone without a follow-up clinical interview to establish a diagnosis of MDD, then potentially dangerous overtreatment and mislabeling could occur.

We found that effect sizes for drug treatment trials were consistently positive, but generally small (0.20-0.38; 1%-4% of the variance in depression change scores). The mean baseline 17-item Hamilton Depression Rating Scale scores in drug trials we reviewed were generally several points lower than baseline 17-item Hamilton Depression Rating Scale scores of trials submitted to the US Food and Drug Administration for licensing, and lower baseline scores have been shown to be associated with smaller responses to
DEPRESSION AND CARDIOVASCULAR CARE

The greatest challenges to ensuring accurate diagnosis and treatment of depression for cardiac patients may not lie in the accuracy of available screening tools, but in the effectiveness with which treatment can be delivered, given the competing demands when treating patients with CVD. One-quarter to one-third of patients in primary care settings discontinue depression treatment within 1 month of initiation and as many as one-half discontinue treatment within 3 months.73-77 There is little reason to think that the situation would be better in cardiovascular care settings. Collaborative care has been proposed as a potential solution to management barriers that may improve both short- and longer-term depression outcomes.74,78,79 Collaborative care is a multifaceted organizational intervention based on chronic disease management principles that involves a greater role of nonmedical specialists (eg, nurse practitioners or case managers) working with mental health specialists and other clinicians to provide optimal disease management and treatment follow-up.78,80 A recent meta-analysis by Gilbody et al,79 however, found only modest effects on depression outcomes for collaborative care interventions at 6-month (effect size 0.25) and up to 5-year follow-up (effect size 0.15). Although 1 study82 found that collaborative care was cost-effective for patients with depression and diabetes when total health service costs were considered, more favorable evidence is needed.

In summary, this systematic review of the evidence did not find evidence for or against the recommendations that depression be evaluated or that screening for depression be considered as part of standard care in patients with CVD.13,14,16 There was not enough evidence to assess potential harms related to screening or treatment. The high prevalence of depression in patients with CVD, the adverse health care outcomes associated with depression, and the availability of easy-to-use case-finding instruments make it tempting to endorse widespread depression screening in cardiovascular care. However, the adoption of depression screening in cardiovascular care settings would likely be unduly resource intensive and would not likely be beneficial to patients in the absence of significant changes in current models of care. More research on the impact of depression screening in the context of different care models is needed. If collaborative care models were to show, for instance, that the cost of improving depression outcomes could be offset by increased productivity and decreased absenteeism,81 then depression screening in collaborative care may be justifiable. Finally, although routine depression screening is not supported by the evidence, physicians should be aware that there are patients with serious and potentially life-threatening depression in most cardiovascular care settings. For these patients, physicians should provide appropriate treatment, referral, or both.

Author Affiliations: Departments of Psychiatry (Drs Thombs and Frasure-Smith) and Nursing (Dr Frasure-Smith), McGill University, Montréal, Québec, Canada; Department of Psychiatry, Sir Mortimer B. Davis Jewish General Hospital, Montréal, Québec, Canada (Dr Thombs); Interdisciplinary Center for Psychiatric Epidemiology, Departments of Psychiatry (Dr de Jonge and Ms Zuidersma), and Internal Medicine, University Medical Center Groningen, University of Groningen, Groningen, the Netherlands (Dr de Jonge); CoRPS-Center of Research on Psychology in Somatic Diseases, Department of Medical Psychology, Tilburg University, Tilburg, the Netherlands (Dr de Jonge); Department of Psychiatry, University of Pennsylvania School of Medicine, Philadelphia (Dr Coyne); Departments of Medicine, Epidemiology, and Biostatistics, University of California, San Francisco, and Veterans Affairs Medical Center, San Francisco, California (Dr Whooley); Research Centre, Centre Hospitalier de l’Université de Montréal, Research Centre, Montréal Heart Institute, Montréal, Québec, Canada (Dr Frasure-Smith); Department of Cancer and Molecular Medicine, Leicester Royal Infirmary, Leicester, England (Dr Mitchell); Interfaith Medical Center, Brooklyn, New York (Dr Eze-Nliam); Federal University of Ceará School of Medicine, Fortaleza, Brazil (Mr Lima); Harold E. Harrison Medical Library, Johns Hopkins Bayview Medical Center (Ms Smith) and Department of Medicine, Johns Hopkins University School of Medicine (Mr Soderlund and Dr Ziegelstein), Baltimore, Maryland.

Critical revision of the manuscript for important intellectual content: Thombs, de Jonge, Zuidersma, Eze-Nliam, Lima, Smith, Soderlund, Ziegelstein.

Statistical analysis: Thombs, Mitchell, Zuidersma.

Administrative, technical, or material support: Thombs, de Jonge, Eze-Nliam, Lima, Ziegelstein.

Funding/Support: Thombs is supported by the Etablissement de Jeunes Cher-
DEPRESSION AND CARDIOVASCULAR CARE

references


32. Low GD, Hubble AM. Screening for depression after cardiac events using the Beck Depression Inventory-II and the Geriatric Depression Scale. Soc Indic Res. 2007;82:527-543.


DEPRESSION AND CARDIOVASCULAR CARE


