Depression and Diabetes Treatment Nonadherence: A Meta-Analysis

Jeffrey S. Gonzalez, PhD1,2
Mark Peyrot, PhD3,4
Lauren A. McCarty, MA1
Erin Marie Collins1

Luís Serpa1
Matthew J. Mimiaga, ScD, MPH1,2
Steven A. Safren, PhD1,2

OBJECTIVE — Depression is common in patients with diabetes and is associated with worse treatment outcomes. Its relationship to treatment adherence, however, has not been systematically reviewed. We used meta-analysis to examine the relationship between depression and treatment nonadherence in patients with type 1 and type 2 diabetes.

RESEARCH DESIGN AND METHODS — We searched MEDLINE and PsycINFO databases for all studies published by June 2007 and reviewed references of published articles. Meta-analytic procedures were used to estimate the effect size r in a random effects model. Significance values, weighted effect sizes, 95% CIs, and tests of homogeneity of variance were calculated.

RESULTS — Results from 47 independent samples showed that depression was significantly associated with nonadherence to the diabetes treatment regimen (z = 9.97, P < 0.0001). The weighted effect size was near the medium range (r = 0.21, 95% CI 0.17–0.25). Moderator analyses showed that the effect was significantly larger in studies that measured self-care as a continuous versus categorical variable (P = 0.001). Effect sizes were largest for missed medical appointments and composite measures of self-care (r values = 0.31, 0.29). Moderator analyses suggest that effects for most other types of self-care are also near the medium range, especially in studies with stronger methodologies.

CONCLUSIONS — These findings demonstrate a significant association between depression and treatment nonadherence in patients with diabetes. Studies that used stronger methodologies had larger effects. Treatment nonadherence may represent an important pathway between depression and worse diabetes clinical outcomes.

Diabetes Care 31:2398–2403, 2008

Patients with diabetes are more likely to experience depression than the general population, and the presence of depression is associated with poorer quality of life (1), increases in hyperglycemia (2), health care utilization (3), risk of complications (4,5), functional impairment (6), and risk of mortality (5). The relationship between depression and worse outcomes in diabetes could be explained, in part, through depression’s relationship to poorer self-care and treatment adherence. Meta-analysis has been used to document a consistent relationship between depression and hyperglycemia (2) and depression and diabetes complications (4). One meta-analysis has documented a consistent relationship between depression and nonadherence to medical treatment in 12 studies of diverse patient populations, but this meta-analysis did not include any studies of patients with diabetes (7).

We surveyed the growing literature on depression and treatment nonadherence in patients with diabetes and performed a meta-analysis to evaluate the strength of this association. We examined whether the association between depression and worse self-care varied as a function of the type of self-care assessed (i.e., composite measure of multiple self-care behaviors, medical appointment attendance, diet, medication adherence, exercise, self-monitoring of blood glucose, foot care). We also examined whether this association differed by study population because the behaviors involved in self-care for type 1 diabetes are significantly different from type 2 diabetes because children and adolescents are developmentally different from adults and may be differentially sensitive to the effects of depression. Finally, we sought to determine whether the depression-adherence relationship differed depending on the strength of study methodology, i.e., is the relationship attenuated in more rigorous studies or does stronger methodology reveal a larger relationship? Therefore, we evaluated a set of potential moderators of this relationship focusing on strength of study methodology (i.e., different measurement strategies and longitudinal versus cross-sectional designs).

RESEARCH DESIGN AND METHODS

Sources of data and literature review strategy

We searched the MEDLINE and PsycINFO databases from 1950 through 1 June 2008 to identify all studies examining the relationship between depression and self-care in patients with diabetes. We combined the keywords “diabetes” or “diabetes mellitus”, various synonyms for compliance, adherence, self-care, and self-management; and various terms related to depression or depressive symptoms. The reference lists from relevant manuscripts were reviewed for additional citations. In addition, we contacted the corresponding author for each manuscript retained for analysis to solicit additional articles. Corresponding authors were also asked to identify other researchers in the field who might have applicable articles; we then queried these researchers with the same requests as noted above.
Inclusion and exclusion criteria
Studies were limited to those 1) involving children, adolescents, or adults with type 1 or type 2 diabetes where findings were reported in English; 2) reporting sufficient data on the strength of the relationship between depression and treatment adherence to calculate an effect size; and 3) not involving any intervention that could possibly affect the relationship between treatment adherence and depression because, as in the meta-analysis by DiMatteo et al. (7), we sought to examine a naturally occurring correlate of nonadherence. Treatment adherence was broadly operationalized to include adherence to medication, dietary recommendations, exercise, glucose monitoring, foot self-care, scheduled medical appointments, or any overall adherence composite measure. Depression was operationalized to include studies that used any measure that had been developed specifically to assess current depressive symptoms or diagnosis.

Study procedures
After the search was run, titles and abstracts were reviewed by three coders (E.M.C., L.A.M., and L.S.). Each coder identified any abstract that he or she felt might have assessed depression and diabetes self-care. Full-text manuscripts were then obtained for these abstracts. Thus, each abstract was reviewed at least three times by independent coders, and final determinations about inclusion/exclusion were made in consultation with the primary author (J.S.G.) who reviewed each manuscript in question. In addition to the information needed to compute effect sizes, each manuscript was coded according to self-care domain (medication, diet, exercise, glucose monitoring, foot care, attendance at scheduled medical appointments, or an overall adherence composite measure); diabetes type (type 1 or type 2); life stage (adult versus child/adolescent); research design (longitudinal versus cross-sectional); level of adherence measurement (categorical versus continuous); level of depression measurement (categorical versus continuous); use of validated measures versus nonvalidated measures for depression; and use of objective versus nonobjective measures of self-care.

For effects to be independent, each study could contribute only one effect size for each meta-analysis. For the overall analysis, when multiple types of diabetes self-care were measured, we took the average of the effects. However, when several different measures of the same variable were available, we selected only one effect. The general criterion we used was to select the effect obtained by the stronger method, e.g., we selected continuous measures of depression or adherence if data on both continuous and dichotomized relationships were reported in a single study, since it has been demonstrated that dichotomizing continuous variables causes the estimation of the population correlation to be systematically reduced in magnitude and reduces power (8,9). If sample sizes differed for different measures, we used the measures that had a larger sample size (10,11). One study reported both self-reported and caregiver reported data for depression and self-care; we used self-reported data for the analysis (12). Where two papers were published from samples that were not independent, we included the article with the stronger design: longitudinal over cross-sectional (13,14) or larger sample (15,16). Finally, one study measured depression as both a diagnosis based on structured interview and as a comparison based on the total score from the Centers for Epidemiological Studies Depression Scale (CESD) using two different cutoffs (16 or 22) (17). Because an average of the effect sizes would not be appropriate, and there was no intrinsic difference in strength of methodology, we chose the data based on the CESD cut point of 22. Using data based on the CESD cutoff of 16 or the diagnosis did not alter any of our results (data not shown).

Statistical analysis
We followed the metaanalytic procedures of Lipsey and Wilson (18). The effect size \( r \) was used. When \( r \) or phi statistics were not provided, we computed \( r \) from means and SDs, odds ratios, \( t \) tests, \( \chi^2 \), \( F \), contingency table data, or exact \( P \) values. When means and SDs were reported for more than two groups (e.g., adherence scores for low, intermediate, and high depression), we based our effect size calculations on comparisons between the two most extreme groups. When studies presented effect sizes that had more than 1 d.f., we (J.S.G., M.M.) contacted authors to obtain bivariate data. However, because it has been demonstrated that \( r \) can be accurately estimated from \( \beta \)-coefficients in multiple regression, even when covariates are present, we estimated \( r \) in one case (i.e., 19) using the formula \( r = 0.98\beta + 0.05\lambda \), where \( \lambda \) is an indicator variable that equals 1 when \( \beta \) is nonnegative and 0 when \( \beta \) is negative (20).

We report effect sizes based on the random effects model and used a fully random effects analysis to examine moderators. Effect sizes are weighted by the inverse variance of each study, which is determined primarily by sample size but also takes into account other factors that affect the precision of the effect size (18). All computations were based on Fisher \( z \) transformations of \( r \). Meta-analytic software (Comprehensive Meta-Analysis 2.0) was used to calculate average \( z \) scores and \( P \) values, weighted effect size \( r \) values, and the 95% CI around the collective effect size values.

We evaluated the significance of variability among effect sizes by calculating the Q statistic and quantified the degree of dispersion of effect sizes with an \( I^2 \) squared index, which serves as a ratio of the variance between studies to total variance (21). We also used the software to examine whether the effect size differed significantly across levels of the potential moderators of the depression-adherence relationship. We performed tests of moderation for the overall analysis and within each type of self-care behavior, if the results displayed significant heterogeneity and the number of studies was adequate. Because examination of moderation in meta-analysis requires that the effects in each group be independent, we were unable to conduct moderation analyses for type of self-care behavior because most studies measured more than one type of self-care. However, we present results of separate meta-analyses for each type of self-care behavior. For these analyses, the same study could be counted in multiple categories of self-care because the effects were independent within each category.

Because “the file drawer problem,” or the selective publishing of significant findings and the disproportionate exclusion of nonsignificant findings from publication, can be a problem in generalizing the results of meta-analysis, we also calculated the “fail-safe n” for each grouped effect size (22), reflecting the number of unpublished studies necessary to reduce the obtained effect size to nonsignificance. We also provide the number of unpublished studies with a mean effect size of \( r = 0.00 \) that would be required to reduce the effect size below the \( r = 0.05 \) level, chosen as an effect size that would be trivial (23). Cohen (24) has suggested that, in behavioral science research, effects can be considered small when \( r =
**RESULTS**

**Overall analysis**
A total of 47 independent study samples, including 17,319 participants from 43 published reports (see online appendix at http://dx.doi.org/10.2337/dc08–1341), met our inclusion criteria. The characteristics and findings of these studies are summarized in the online appendix. The studies used a variety of measures, methods, samples, and analytical approaches. Meta-analysis revealed a significant association between depression and poorer self-care ($P < 0.0001$) and a weighted effect size ($r = 0.21, 95\% \text{ CI} 0.17–0.25$), which places the effect near the benchmark of $r = 0.25$ for a medium-sized effect (24). The “fail-safe $n$” for the number of unpublished studies with null findings required to bring this effect to non-significance is 4,997. Orwin’s fail-safe $n$ for missed medical appointments and exercise were not heterogeneous, moderators were not examined. Because only two studies measured foot care, we could not perform a moderator analysis for this variable. We found no significant moderators for composite measures.

As can be seen in Table 2, the only effect moderation for type of study population was that studies of children and adolescents obtained significantly larger effects for glucose monitoring than those of adults; although the pattern in the overall analysis was similar, it did not reach statistical significance ($P = 0.064$). In contrast, there were multiple methodological moderators of effect sizes. Longitudinal studies found significantly larger effects for diet than cross-sectional studies. Studies that measured and analyzed adherence as a continuous variable found significantly larger effects in the overall analysis, for medication adherence, and for glucose self-monitoring than those that measured adherence as a categorical

**Analyses by type of self-care**
We also ran our analyses aggregating studies by type of diabetes self-care and performed separate meta-analyses for each grouping. For these analyses, 88 effects were distributed across seven categories of diabetes self-care (Table 1). The effect of depression and self-care varied across types of self-care behaviors, with the strongest effect size found for missed medical appointments ($k = 4, r = 0.31, 95\% \text{ CI} 0.29–0.34$) and the smallest effect size found for foot care ($k = 2, r = 0.07, -0.08–0.21$), which was nonsignificant.

**Moderation analyses**
We examined potential moderators of the effects for the overall analysis and specific self-care domains when significant heterogeneity was present. Because the effects

### Table 1—Meta-analysis results aggregated by type of self-care

<table>
<thead>
<tr>
<th>Subcategory</th>
<th>Domain (P)</th>
<th>$n$</th>
<th>$z$ ($P$)</th>
<th>Weighted $r$</th>
<th>$95% \text{ CI}$</th>
<th>Heterogeneity $Q$ (df and $I^2$)</th>
<th>Fail-safe $n$ ($r = 0.05$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall analysis</td>
<td>Diet (0.001)</td>
<td>47</td>
<td>9.81 (&lt;0.0001)</td>
<td>0.21</td>
<td>0.17–0.25</td>
<td>217.66 (46); $P &lt; 0.001$; $I^2 = 78.87$</td>
<td>149</td>
</tr>
<tr>
<td>Appointment keeping</td>
<td>Composite measures</td>
<td>4</td>
<td>21.58 (&lt;0.0001)</td>
<td>0.31</td>
<td>0.29–0.34</td>
<td>1.79 (3); $P = 0.617$; $I^2 = 0.00$</td>
<td>22</td>
</tr>
<tr>
<td>Diet</td>
<td>18</td>
<td>9.66 (&lt;0.0001)</td>
<td>0.29</td>
<td>0.23–0.34</td>
<td>49.51 (24)</td>
<td>37</td>
<td></td>
</tr>
<tr>
<td>Medication</td>
<td>18</td>
<td>5.15 (&lt;0.0001)</td>
<td>0.14</td>
<td>0.09–0.20</td>
<td>16.86 (22)</td>
<td>22</td>
<td></td>
</tr>
<tr>
<td>Exercise</td>
<td>13</td>
<td>7.89 (&lt;0.0001)</td>
<td>0.14</td>
<td>0.10–0.17</td>
<td>49.73 (16); $P &lt; 0.001$; $I^2 = 65.82$</td>
<td>24</td>
<td></td>
</tr>
<tr>
<td>Glucose monitoring</td>
<td>Foot care</td>
<td>2</td>
<td>3.50 (&lt;0.0001)</td>
<td>0.10</td>
<td>0.04–0.16</td>
<td>31.00 (14); $P = 0.006$; $I^2 = 54.82$</td>
<td>4</td>
</tr>
</tbody>
</table>

0.10, medium when $r = 0.25$, and large when $r \geq 0.40$.

### Table 2—Moderation analyses: significant results

<table>
<thead>
<tr>
<th>Moderators</th>
<th>Subcategory analyses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Methodology: longitudinal (L) vs. cross-sectional (C)</td>
<td>Diet (0.021); Overall analysis (0.001); Medication (0.011); Glucose monitoring (0.036); Glucose monitoring (0.023)</td>
</tr>
<tr>
<td>Self-care continuous (Con) vs. categorical (Cat)</td>
<td>Overall analysis (0.001); Con: ($k = 32; r = 0.24; 0.21–0.28$); Cat: ($k = 14; r = 0.13; 0.07–0.18$)</td>
</tr>
<tr>
<td>Objective measure of self-care (O) vs. non-objective measure (N)</td>
<td>Overall analysis (0.064); Y: ($k = 10; r = 0.29; 0.20–0.37$); A: ($k = 37; r = 0.19; 0.15–0.24$)</td>
</tr>
<tr>
<td>Population: youth (Y) vs. adults (A)</td>
<td>Overall analysis (0.001); Y: ($k = 2; r = 0.22; 0.09–0.35$); N: ($k = 13; r = 0.07; 0.02–0.12$)</td>
</tr>
</tbody>
</table>
variable. Studies that used objective measures of self-care (i.e., electronically downloaded data) found significantly stronger effects for monitoring than those that used other assessment methods.

**CONCLUSIONS** — The results of this meta-analysis, based on 47 independent samples totaling over 17,000 patients, suggest that depression is significantly associated with nonadherence to diabetes self-care with an effect of moderate strength relative to the range of effect sizes typical in the social sciences. The size of this overall effect \((r = 0.21)\) is identical to the effect \((r \text{ weighted by } n - 3 = 0.21)\) obtained by DiMatteo et al. \((7)\) based on a meta-analysis of 12 studies of patients adhering to treatment regimens of other chronic illnesses. The effect is also similar in magnitude to the effects obtained from meta-analyses of the relationship between depression and hyperglycemia \((r = 0.17)\) \((2)\) and between depression and diabetes complications \((r = 0.25)\) \((4)\). To the extent that self-care is causally related to hyperglycemia and diabetes complications, our results suggest that impairment in self-care may be one pathway through which depression is associated with these negative health outcomes in patients with diabetes.

We found that the relationship between depression and worse self-care differed depending on the type of self-care measured. The effect was strongest and heterogeneous for studies of missed medical appointments. Adherence to this type of self-care is unique in that it requires interpersonal behavior. Clinically, depression is associated with impairments in interpersonal behavior such as social withdrawal, disengagement from important activities, avoidance, and often with disruption of interpersonal relationships. Patients with increased levels of depression have been shown to report more dissatisfaction with their providers \((27)\). Considering that missed appointments are also often associated with increased provider frustration \((28)\), decreased empathy and patient-provider communication \((29)\), and less continuity of care \((29)\), this relationship may have important ramifications for depressed patients with diabetes.

Studies that used composite measures of diabetes self-care that tapped into more than one aspect of self-care also found stronger effects. This may be because comprehensive measures are more robust methodologically, or because the effects of depression are more global and are better captured by a more global measure of self-care. Although we found somewhat weaker relationships between depression and diet and medication adherence, studies that used stronger methodology found larger effects. If these effects \((r \text{ values } = 0.20-0.28)\) are considered to be the best estimates of the true effect, then it appears that the effects on diet and medication adherence are also in the moderate range. Surprisingly, the relationship between depression and exercise was relatively weaker and homogenous. The effect for self-monitoring of blood glucose was smaller, but studies that measured monitoring as a continuous variable found significantly stronger effects. Also, the two studies \((25,26)\) that used objective measures of glucose monitoring found significantly stronger effects than those that did not, and the weighted average of these effects \((r = 0.22)\) was near the medium effect size benchmark. Finally, foot care was not significantly related to depression overall. However, only 2 of 47 effects measured foot care, which suggests this aspect of self-care may not be given sufficient attention in research or in providers’ instructions to patients.

With regard to population moderators, we examined differences between type 1 and type 2 diabetes and between children/adolescents and adults. We focused on these contrasts because it was plausible to expect that self-care routines are sufficiently different between these groups so as to affect the relationship between depression and self-care. We found no evidence to suggest that the relationship between depression and self-care varied as a function of type of diabetes. We did find that studies of children and adolescents tended to report larger effects than studies of adults, although effects in both populations were significant. This finding could also support more research on the impact of depression on diabetes self-care in children and adolescents, particularly for self-monitoring of blood glucose. We are unaware of depression treatment interventions that have targeted children or adolescents with diabetes to examine impacts on self-care or control. Our results suggest that, if the relationship between depression and worse self-care is causal, interventions in youth may have more of an impact on diabetes self-care and control than interventions with adults.

The results of this meta-analysis also provide important information on the methodological moderators of the effect between depression and diabetes self-care. We did not find any evidence to suggest that studies using stronger methodology found weaker effects; to the contrary, the effects were significantly larger when stronger methods were used. This was especially true for studies that analyzed self-care as a continuous variable: larger effects were obtained in the overall analyses, for medication adherence, and for glucose monitoring. Significantly larger effects for diet were also found in studies that used longitudinal designs. Also, studies that used objective measures of glucose monitoring found significantly stronger effects than those that relied on self-report or other methods more vulnerable to bias. Further, in the analyses aggregated by type of self-care, the strongest effects were found for appointment keeping, and each of the studies that measured this type of self-care used objective methods to measure it (i.e., medical records data). Thus, assuming that studies that use more rigorous methods will provide more accurate estimates of the true effect, it appears that our overall effect of \(r = 0.21\) may actually underestimate the true association between depression and poorer self-care.

Future studies could be improved by using more rigorous measures and avoiding dichotomization of variables. The question of dichotomization is not only a statistical consideration that reduces power and accuracy \((8)\); it also reflects a limitation in the conceptualization of depression and adherence as purely categorical constructs. For example, in a previous article based on a large sample of primary care patients with type 2 diabetes, we showed that the relationship between depressive symptoms and worse diabetes self-care was not limited to patients likely to meet diagnostic criteria for major depressive disorder. We found the same magnitude of effects between symptoms of depression and worse self-care in a subsample of patients who did not meet screening criteria for major depressive disorder as we did in the overall sample \((13)\). These findings, along with the results of this meta-analysis, suggest that it would be inaccurate to think that the relationship between symptoms of depression and poorer diabetes self-care is limited to those who have clinically significant levels of depression. Rather, it appears that increases in depressive symptoms (measured by a variety of methods) are incrementally associated
with decreases in diabetes self-care. We would also argue, both on conceptual grounds and based on the results of our moderation analyses, that it is also inaccurate to conceptualize self-care adherence as a categorical construct and to think of “adherent” versus “nonadherent” patients.

The results of this meta-analysis may have implications for patient interventions. It should be stated at the outset that none of the studies reviewed here provide conclusive evidence that the relationship between depression and poorer diabetes self-care is causal. However, one study used a longitudinal design and found significant relationships between depressive symptoms at baseline and increases in symptoms of depression over time and a variety of self-care behaviors assessed 9 months later, even when baseline levels of self-care were controlled (14). Another study of an open-label treatment trial of bupropion hydrochloride in patients with type 2 diabetes and major depressive disorder found that depression severity, BMI, total fat mass, and A1C decreased during the acute phase of treatment, whereas adherence to diet and exercise improved significantly (30). The pattern of findings from these studies is consistent with a causal relationship but does not provide conclusive proof. The current meta-analysis excluded randomized controlled trials, which would provide the strongest causal evidence, because we sought to evaluate the naturally occurring relationship between depression and poorer self-care, without the influence of intervention. Yet, evidence from randomized controlled trials suggesting that treating depression in patients with diabetes has positive effects on diabetes self-care has been lacking. Trials of antidepressants (31,32), cognitive behavioral therapy (33), and stepped-care case management (34–36) have had positive effects on depression but have generally failed to have a positive impact on self-care behaviors. However, these trials have avoided integration of diabetes self-management training with the treatment of depression to isolate the effect of treating depression alone on diabetes outcomes.

The limitations of the available intervention literature suggest there may be an opportunity to maximize effects on diabetes control by developing comprehensive interventions to improve both depression and self-care (37). It may be necessary but not sufficient to treat depression to improve self-care in patients who are struggling with depression and problems with diabetes self-management. Comprehensive interventions that address both self-care and depression management are likely to be more successful, and guidelines for how such interventions may be implemented in practice have been proposed (38). Because even nonclinical levels of depressive symptoms can be associated with nonadherence (13) and are associated with significant increases in risk for complications, functional impairment, and death (3), it is likely that interventions oriented toward self-care or adherence should include a component to address psychological/emotional distress. Previous research has shown that such integrated programs can have positive effects on both self-care and emotional distress (39,40). Taken together with the findings of this meta-analysis, this work could guide the design of well-powered randomized controlled trials of such interventions.

Acknowledgments—The efforts of J.S.G. and S.A.S. were supported by a grant from the National Institute of Mental Health (1R01MH078571).

References