Religious Involvement and Mortality: A Meta-Analytic Review

Michael E. McCullough
National Institute for Healthcare Research

David B. Larson
National Institute for Healthcare Research

William T. Hoyt
Iowa State University and University of Wisconsin—Madison

Carl Thoresen
Stanford University

A meta-analysis of data from 42 independent samples examining the association of a measure of religious involvement and all-cause mortality is reported. Religious involvement was significantly associated with lower mortality (odds ratio = 1.29; 95% confidence interval: 1.20–1.39), indicating that people high in religious involvement were more likely to be alive at follow-up than people lower in religious involvement. Although the strength of the religious involvement–mortality association varied as a function of several moderator variables, the association of religious involvement and mortality was robust and of the order of magnitude that has come to be expected for psychosocial factors. Conclusions did not appear to be due to publication bias.

Key words: religion, mortality, survival, longevity, meta-analysis

Substantial numbers of Americans engage in religious activity. More than 90% of American adults are affiliated with a formal religious tradition (Kosmin & Lachman, 1993). Nearly 96% of Americans believe in God or a universal spirit, 42% attend a religious worship service weekly or almost weekly, 67% are members of a local religious body, and 60% feel that religion is “very important” in their lives (Gallup, 1995).

Could such religious activities and beliefs confer physical health benefits? Some research suggests that religious involvement is favorably associated with measures of physical health such as high blood pressure (Levin & Vanderpool, 1989), cancer (Friedlander, Kark, & Stein, 1986), heart disease (Friedlander, Kark, & Stein, 1986), stroke (Colantonio, Kasl, & Ostfield, 1992), and suicide (Kark, Shemi, et al., 1996). Other studies suggest that religious involvement might help to buffer the impact of stress on physical and mental health (Kendler, Gardner, & Prescott, 1997; Krause & Van Tran, 1987; Pressman, Lyons, Larson, & Strain, 1990).

Hypothetically, these associations of religious involvement and health might lead to longer life. Several recent studies (Goldbourt, Yaari, & Medalie, 1993; Hummer, Rogers, Nam, & Ellison, 1999; Kark, Shemi, et al., 1996; Oxman, Freeman, & Manheimer, 1995; Strawbridge, Cohen, Shema, & Kaplan, 1997) have found that religious involvement—variously operationalized as religious attendance, membership in religious kibbutzim, finding strength and comfort from one’s religious beliefs, and religious orthodoxy—is associated with lower mortality.

Potential Moderators of the Association of Religious Involvement and Mortality

However, the association of religious involvement and mortality is unlikely to be unequivocal; it is probably influenced not only by the quality of research methods used to examine the association but also by several characteristics of the research samples under study in individual investigations. For example, a century of sociological theory and research suggests that the association of religious involvement and physical health might be more closely tied to the psychosocial resources that religion provides rather than any positive psychological states engendered specifically by more private forms of religious expression (Durkheim, 1912/1995; Idler & Kasl, 1997a). For this reason, measures of public religious involvement (i.e., religious attendance) may be more strongly related to health outcomes than are measures of private religiosity (e.g., self-rated religiosity, frequency of private prayer, or use of religion as a coping resource). However, this relation is complicated by a possible confound: Healthy persons might be more likely than unhealthy persons to attend public religious
activities. Thus, the association between religious involvement and mortality is likely to be stronger for measures of public as compared with private religiousness, and effect sizes for studies using public measures of religious involvement should be moderated also by statistical control of physical health.

Second, two studies of patients with cancer (Kune, Kune, & Watson, 1992; LoPrinzi et al., 1994) found that religious involvement was not associated with mortality, whereas many of the studies finding favorable associations of religious involvement and mortality involved basically healthy, community-dwelling adults (Goldbourt et al., 1993; Kark, Shemi, et al., 1996; Strawbridge et al., 1997). Because the health benefits of religiousness may be mediated in part by lifestyle choices and coping behaviors that have their effects over a number of years, the association of religious involvement and mortality might be stronger in basically healthy, community-dwelling samples than in samples of clinical patients.

Third, some data suggest that the association of religious involvement with mortality might be stronger in women than in men (House, Robbins, & Metzner, 1982; Strawbridge et al., 1997). If so, then studies with mostly female samples should yield more favorable associations of religious involvement and mortality than would studies with mostly male samples.

Finally, measures of religious involvement could be associated with, confounded with, or mediated by a variety of other demographic, psychosocial, and physiological variables, such as (a) age, (b) gender, (c) race-ethnicity, (d) general social support, (e) psychological well-being, (f) health practices such as exercise and smoking, and (g) physical health. To the extent that this is the case, the association of religious involvement with mortality would be more favorable in studies that controlled for fewer of these variables than in studies that controlled for large numbers of potential confounds and mediators (Idler & Kasl, 1997a, 1997b).

Although reviews of the relationship between denominational affiliation and mortality (Jarvis & Northcott, 1987; Troyer, 1988) and between religious involvement and physical health (Craigie, Liu, Larson, & Lyons, 1988; Levin & Vanderpool, 1989) have been published, no researchers to date have used meta-analytic methods to examine the association of religious involvement and all-cause mortality. To address this gap in the literature, we conducted a meta-analysis of the research on religious involvement and mortality.

Method

Literature Search

The literature search involved three steps. First, we searched six electronic databases relevant to medicine (Medline), psychology (PsycINFO), sociology (Sociofile), nursing (Cumulative Index of Nursing and Allied Health Literature [CINAHL]) and education (Education Resources Information Center [ERIC], Dissertation Abstracts) to find published and unpublished studies on religious involvement and mortality through June 1999. We crossed multiple search terms related to religious involvement (religion, religiousness, religiosity, religious) with multiple search terms related to mortality (mortality, fatality, death, survival) and leading causes of death (e.g., cardiovascular, cancer). Second, we examined reference sections of retrieved studies to identify additional studies. Third, we examined previous reviews of the literature and consulted with three experts in the field to identify fugitive studies. We excluded studies that used religious affiliation or denomination (e.g., Christian, Jewish) as the sole measure of religion.

Relevant Studies

We identified 41 research reports in which a measure of religious involvement was examined as a predictor of all-cause mortality. Of these reports, 5 (Berkman & Syme, 1979; Enstrom, 1975; Seeman, Kaplan, Knudsen, Cohen, & Guralnik, 1987; Strawbridge et al., 1997; Wingard, 1982) were based on the Alameda County data set, 5 (Comstock & Lundin, 1967; Comstock & Partridge, 1972; Comstock, Shah, Meyer, & Abbey, 1971; Comstock & Tonascia, 1977; Helsing & Szldo, 1981) were based on the Washington County data set, 2 (Idler & Kasl, 1991, 1992) were based on the Yale Health and Aging Project, 2 (Koenig, 1995; Koenig et al., 1998) were based on a cohort of male patients at a Veterans Administration Hospital, 2 (Bryant & Rakowski, 1992; Goldmar, Korenman, & Weinstein, 1995) were based on the National Health Interview Survey: Longitudinal Study of Aging, 70 Years and Over, 1984-1990 (Kovar, Fitti, & Chyba, 1990), and 2 (Ringdal, 1996; Ringdal, Gostestam, Kaasa, Kvissland, & Ringdal, 1995) were based on a cohort of cancer patients at the University Hospital of Trondheim, Norway. To satisfy the assumption of statistical independence that underlies meta-analytic research, effect size estimates for data sets yielding more than one report were based on the report that used (a) the longest observation period and (b) the largest number of cases, as is standard meta-analytic practice (e.g., Miller, Smith, Turner, Gijarmo, & Hallet, 1996). Thus, 42 effect sizes were extracted from 29 (noted in reference section by an asterisk) of 41 research reports.

Computation of Effect Size Estimates

Most studies reported the association of religious involvement and all-cause mortality in relative risk, relative hazard, or odds ratio metrics. Typically, these measures of association were adjusted for one or more covariates. Despite its ease of interpretability (Davies, Crombie, & Tuvakol, 1998; Laird & Mosteller, 1990), the relative risk (and by extension, the relative hazard) is not ideal for meta-analysis (Fleiss, 1994). Instead, most meta-analysis experts recommend using odds ratios as a standard measure of effect size for categorical data (Fleiss, 1994; Haddock, Rindskopf, & Shadish, 1998; Laird & Mosteller, 1990). The odds ratio for a fourfold table is the odds of a favorable outcome for a group of interest (i.e., the odds of survival at follow-up for highly religious individuals) divided by the odds for the comparison group (i.e., less religious individuals). For studies that included control variables (e.g., baseline physical health, alcohol or drug use), the odds ratios are likewise adjusted—they represent the relative odds of survival for religious and nonreligious individuals, controlling for the designated attributes. Odds ratios near 1.0 indicate weak or nonexistent associations between variables, whereas odds ratios greater than 3.0 (or less than 0.33, in the case of negative associations) represent strong associations between variables (Haddock et al., 1998).

For studies in which authors reported odds ratios, we used those as our effect size estimates. When only raw data (e.g., 2 × 2 cell frequencies) were available, we calculated odds ratios and variances using standard formulas (e.g., Fleiss, 1994). When study authors reported relative risks or relative hazards and measures of sampling variability (e.g., standard errors, variances, or 95% confidence intervals [CIs]), we estimated the corresponding odds ratios by reconstructing the implied fourfold tables. Odds ratios are always of slightly larger magnitude than their corresponding relative risks (Davies et al., 1998). As would be expected, our estimated odds ratios were also slightly larger (i.e., 6% larger on average) than their corresponding relative risk and relative hazard values.

Some authors (e.g., Janoff-Bulman & Marshall, 1982; Kune et al., 1992; Spiegell, Bloom, & Gottheil, 1983; Yates, Clahme, St. James, Foliansbee, & McKegney, 1981) reported effect sizes in other metrics (e.g., correlation coefficients, survival time). Details on how we derived odds ratio estimates for these effect sizes are available from Michael E. McCullough.
Log transformation. Because odds ratios are asymmetrical (negative associations can vary from 0 to 1.0, whereas positive associations can vary from 1.0 to +∞), they are customarily subjected to a natural log transformation for use in meta-analyses (Fleiss, 1994; Haddock et al., 1998). Log odds ratios are distributed around zero with a theoretical range of (−∞ to +∞). Negative values indicate negative associations, and positive values indicate positive associations. This transformation is ideal when within-study sample sizes are large (Shadish & Haddock, 1994), as was the case for the present meta-analysis. An additional advantage of using log odds ratios for meta-analysis is that their variances are independent of the magnitude of association between the variables and are easily estimated from the cell frequencies in the fourfold table (Fleiss, 1994). We present the results of the present study in log odds ratios and odds ratios (derived by taking the antilog of the log odds ratio) to facilitate interpretation.

Multiple effect sizes in a single study. Five studies (Janoff-Bulman & Marshall, 1982; Krause, 1998; Oxman, Freeman, & Manheimer, 1995; Idler & Kasl, 1992; Yates et al., 1981) examined the association of mortality with two or more measures of religious involvement. We computed the mean effect size across all measures of religious involvement for these five studies. Several studies also reported an effect size for the association of religious involvement and all-cause mortality both (a) before adjusting for other variables and (b) after adjusting for other variables. In such studies, we used the more stringently controlled effect size. Thus, each study contributed a single effect size to the meta-analysis, with the exception of nine studies in which we were able to derive independent effect sizes for multiple subsamples (e.g., men and women), yielding a total of 42 independent effect sizes for analysis.

Moderator Coding

Along with effect sizes, we coded each study for three classes of potential moderator variables: variations in research design, variations in sample characteristics, and variations in how religious involvement was operationalized. To understand the implications of research design, we coded each study for (a) statistical controls (i.e., number and types of variables for which the religious involvement—mortality association was adjusted) and (b) length of follow-up period in months. Sample characteristics of interest were (c) percentage of males, (d) whether the sample was drawn from a community or clinical population, and (e) mean age of participants at baseline. To examine the effects of variation in measurement practices, we created a categorical variable called (f) measure type (public, private, a combination of public and private, or missing—i.e., the authors indicated that religiousness was measured, but they did not indicate how). Interjudge agreement for the coding of the above-mentioned categorical variables was evaluated with Cohen’s kappas (κ > .85). Interjudge reliabilities for ratings of continuous variables were estimated using Shrout and Fleiss’s (1979) formula for the intraclass correlation coefficient (3, 1). The mean intraclass correlation coefficient for all coded variables was .97, with intraclass correlation coefficients ranging from .78 to 1.0.

Analyses

To generalize beyond the sample of studies actually reviewed (i.e., to claim that their results reflect the likely magnitude of effects for other, future samples of studies in the research domain), meta-analytic researchers should use random-effects models to aggregate effect sizes and estimate the reliability of these aggregates (Hedges & Vevea, 1998). This strategy was clearly desirable for the present meta-analysis: Our belief that the variation among effect sizes was greater than expected by chance. Second, we examined the impact of the theoretically derived moderator variables on effect size. Third, we examined whether statistical control of specific demographic, psychosocial, and medical variables influenced effect size (to explore which variables might be confounds or mediators of the association of religious involvement and mortality). Fourth, we conducted sensitivity analyses to evaluate the validity of our meta-analytic findings and their tolerance to future null results.

Results

We computed a total of 42 independent effect sizes representing 125,826 participants. Effect size estimates (odds ratios) and characteristics associated with each effect size appear in Table 1.

Omnibus Analysis

In the omnibus analysis, no moderator variables were modeled, and the observed effect sizes were presumed to constitute a representative sampling of the study populations of interest. Effect size estimates were subject to both between-studies variance (because the true effect sizes differ for different classes of studies) and within-study variance (due to sampling error). The aggregate log odds ratio for the omnibus analysis (k = 42, N = 125,826) was γ0 = .26, SE = .036, p < .001. The γ0 of .26 corresponds to an odds ratio of 1.29 (95% CI: 1.21–1.39), indicating that across all studies, highly religious individuals had odds of survival approximately 29% higher than those of less religious individuals. These effect sizes were heterogeneous. Between-studies variance was significantly greater than zero: τ = .0206, χ²(41) = 91.62, p < .001. The corresponding Birge ratio (Haddock et al., 1998) was 2.23, suggesting that between-studies variation was 123% greater than expected due to sampling error alone. We therefore estimated other models that incorporated the moderator variables to determine the study characteristics to which between-studies variation in effect size could be attributed.

Moderator Analyses

Moderator analyses can be conducted in HLM using random-effects regression models with prediction equations of the form:

\[ E_{ij} = \gamma_0 + \gamma_1 W_{ij} + \gamma_2 W_{ij} + \ldots + \gamma_s W_{ij} + u_j + e_{ij} \]  

where \( E_{ij} \) is the effect size for study \( j \), \( W_{ij} \) to \( W_{ij} \) are \( S \) predictor (moderator) variables, \( \gamma_1, \ldots, \gamma_s \) are regression weights associated with each of these predictors, \( u_j \) represents systematic variability in study \( j \) not captured by the \( S \) predictors, and \( e_{ij} \) represents sampling error for study \( j \). In this model, the intercept (\( \gamma_0 \)) is the estimated effect size for studies with a value of zero on all moderator variables, and the remaining regression weights indicate the amount of expected variation in this effect size for a one-unit change on each moderator. We centered continuous predictors around their means and coded the two categorical moderators so
<table>
<thead>
<tr>
<th>Reference</th>
<th>Sampling</th>
<th>N</th>
<th>% male</th>
<th>Mean age at baseline</th>
<th>Length of follow-up (months)</th>
<th>Measures of religious involvement</th>
<th>Variables controlled in effect size</th>
<th>Adjusted effect size (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abramson, Gofin, &amp; Peritz (1982)</td>
<td>83% sample of all men, age &gt; 60, from a Jewish neighborhood in Jerusalem</td>
<td>387</td>
<td>100</td>
<td>67.35</td>
<td>60</td>
<td>Missing</td>
<td>1, 3, 6, 8, 13</td>
<td>1.00</td>
</tr>
<tr>
<td>Baugher, Burger, Smith, &amp; Wallston (1989/1990)</td>
<td>Subsample of terminally ill people from National Hospice Study</td>
<td>300</td>
<td>48.5</td>
<td>68.10</td>
<td>0.25</td>
<td>Rating of religion as the greatest source of strength in one's life</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Comstock &amp; Tonascia (1977)</td>
<td>98% sample of all adults, age &gt; 16.5, in Washington County, MD</td>
<td>38,839</td>
<td>47.3</td>
<td>48</td>
<td>96</td>
<td>Frequency of church attendance</td>
<td>1, 2, 3, 5, 11, 16</td>
<td>1.38</td>
</tr>
<tr>
<td>Devins, Mann, Mandin, &amp; Paul (1990)</td>
<td>50% random sample of patients receiving hemodialysis or renal transplant in Calgary, Alberta, Canada</td>
<td>97</td>
<td>57.7</td>
<td>40.3</td>
<td>46</td>
<td>Missing</td>
<td>1, 3, 4, 5, 8, 11, 16, 17</td>
<td>1.00</td>
</tr>
<tr>
<td>Enstrom (1989)</td>
<td>Representative sample of community-dwelling adults in Alameda County, CA</td>
<td>364</td>
<td>100</td>
<td>39.8</td>
<td>168</td>
<td>Frequency of church attendance</td>
<td>1, 2, 3, 11</td>
<td>2.20</td>
</tr>
<tr>
<td>Enstrom (1989)</td>
<td>Representative sample of community-dwelling adults in Alameda County, CA</td>
<td>627</td>
<td>0</td>
<td>42.5</td>
<td>168</td>
<td>Frequency of church attendance</td>
<td>1, 2, 3, 11</td>
<td>1.26</td>
</tr>
<tr>
<td>Goldbourt et al. (1993)</td>
<td>Stratified random sample of male Israeli civil servants and municipal employees</td>
<td>9,245</td>
<td>100</td>
<td>50.3</td>
<td>276</td>
<td>Three-item measure of religious orthodoxy</td>
<td>1, 3, 17</td>
<td>1.16</td>
</tr>
<tr>
<td>Goldman et al. (1995)</td>
<td>National probability sample of noninstitutionalized adults, age &gt; 70</td>
<td>2,847</td>
<td>100</td>
<td>76.8</td>
<td>72</td>
<td>Religious service attendance in past 2 weeks</td>
<td>1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 16</td>
<td>1.43</td>
</tr>
<tr>
<td>Goldman et al. (1995)</td>
<td>National probability sample of noninstitutionalized adults, age &gt; 70</td>
<td>4,631</td>
<td>0</td>
<td>77.7</td>
<td>72</td>
<td>Religious service attendance in past 2 weeks</td>
<td>1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 16</td>
<td>1.42</td>
</tr>
<tr>
<td>House et al. (1982)</td>
<td>10% random sample of community-dwelling adults in Tecumseh Community Health Study</td>
<td>1,398</td>
<td>0</td>
<td>48</td>
<td>126</td>
<td>Frequency of church attendance in past year</td>
<td>1, 3, 8</td>
<td>1.15</td>
</tr>
<tr>
<td>House et al. (1982)</td>
<td>10% random sample of community-dwelling adults in Tecumseh Community Health Study</td>
<td>1,301</td>
<td>100</td>
<td>48</td>
<td>126</td>
<td>Frequency of church attendance in past year</td>
<td>1, 3, 6, 8, 11, 17</td>
<td>1.02</td>
</tr>
<tr>
<td>Hummer et al. (1999)</td>
<td>Nationally representative sample in National Health Interview Survey</td>
<td>21,204</td>
<td>47.5</td>
<td>43.2</td>
<td>102</td>
<td>Frequency of religious attendance</td>
<td>1, 2, 3, 4, 5, 6, 7, 9, 10, 11, 12, 13, 16</td>
<td>1.20</td>
</tr>
<tr>
<td>Idler &amp; Kasl (1992)</td>
<td>Stratified probability sample of noninstitutionalized adults, age &gt; 65, in New Haven, CT</td>
<td>1,166</td>
<td>100</td>
<td>73.9</td>
<td>72</td>
<td>Two-item measure of public religiosity (α = .63); two-item measure of private religiosity (α = .71)</td>
<td>1, 6, 8</td>
<td>1.00</td>
</tr>
<tr>
<td>Idler &amp; Kasl (1992)</td>
<td>Stratified probability sample of noninstitutionalized adults, age &gt; 65, in New Haven, CT</td>
<td>1,646</td>
<td>0</td>
<td>74.9</td>
<td>72</td>
<td>Two-item measure of public religiosity (α = .63); two-item measure of private religiosity (α = .71)</td>
<td>1, 2, 3, 4, 5, 6, 8</td>
<td>1.00</td>
</tr>
<tr>
<td>Janoff-Bulman &amp; Marshall (1982)</td>
<td>24% convenience sample of residents from a nursing home in western Massachusetts</td>
<td>25</td>
<td>32</td>
<td>75.4</td>
<td>30</td>
<td>Self-rated religiosity; perceived increase in religiosity since youth Membership in religious vs. secular kibbutz</td>
<td>1, 3, 5, 7, 8, 14, 19</td>
<td>0.42</td>
</tr>
<tr>
<td>Kark, Shemi, et al. (1996)</td>
<td>All adults from 11 religious and 11 secular kibbutz in Israel</td>
<td>1,996</td>
<td>100</td>
<td>53.6</td>
<td>192</td>
<td></td>
<td>1, 3</td>
<td>1.84</td>
</tr>
<tr>
<td>Study</td>
<td>Methodology</td>
<td>Sample Size</td>
<td>Mean Age</td>
<td>N</td>
<td>Measure Description</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------------------------------------</td>
<td>------------------------------------------------------------------------------</td>
<td>-------------</td>
<td>----------</td>
<td>---</td>
<td>---------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kark, Shemti, et al. (1996)</td>
<td>All adults from 11 religious and 11 secular kibbutzim in Israel</td>
<td>1,904</td>
<td>0</td>
<td>53.2</td>
<td>192</td>
<td>Membership in religious vs. secular kibbutz</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Koenig et al. (1998)</td>
<td>Consecutive sample of male admissions to Veterans Administration Medical Center in Durham, NC</td>
<td>1,010</td>
<td>100</td>
<td>67.8</td>
<td>109.5</td>
<td>Three-item measure of religious coping (α = .82)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Koenig et al. (1999)</td>
<td>Probability sample of community-dwelling adults, age &gt; 65, in North Carolina</td>
<td>1,387</td>
<td>100</td>
<td>74.35</td>
<td>75.5</td>
<td>Frequency of attending religious meetings or services</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Koenig et al. (1999)</td>
<td>Probability sample of community-dwelling adults, age &gt; 65, in North Carolina</td>
<td>2,581</td>
<td>0</td>
<td>75.1</td>
<td>75.5</td>
<td>Frequency of attending religious meetings or services</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Krause (1998)</td>
<td>Probability sample of adults, age &gt; 65, in Health Care Finance Administration Medicare Beneficiary Eligibility List</td>
<td>819</td>
<td>40.5</td>
<td>73.8</td>
<td>48</td>
<td>Three-item measure of organizational religiosity (α = .86); two-item measure of nonorganizational religiosity (α = .68); three-item measure of religious coping (α = .97)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kune et al. (1992)</td>
<td>Consecutive sample of new cases of adenocarcinoma of the colorectum in Melbourne, Australia</td>
<td>695</td>
<td>54.1</td>
<td>65</td>
<td>62</td>
<td>Self-rated religiousness</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LoPrinzi et al. (1994)</td>
<td>Consecutive sample of new patients with advanced cancer being treated by North Central Cancer Treatment Group</td>
<td>1,077</td>
<td>55.8</td>
<td>62.2</td>
<td>24</td>
<td>Personal involvement in organized religion; role of faith in personal life</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Musick, House, &amp; Williams (1998)</td>
<td>National probability sample of adults, age &gt; 25</td>
<td>3,617</td>
<td>47</td>
<td>47.1</td>
<td>93</td>
<td>Frequency of religious service attendance</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oman &amp; Reed (1998)</td>
<td>Community-dwelling residents, age &gt; 55, of Marin County, CA</td>
<td>1,931</td>
<td>42.2</td>
<td>74.4</td>
<td>59</td>
<td>Frequency of religious service attendance</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oxman et al. (1995)</td>
<td>Consecutive sample of adults, age &gt; 55, undergoing elective heart surgery by one of four Dartmouth surgeons</td>
<td>232</td>
<td>75</td>
<td>69.9</td>
<td>6</td>
<td>Frequency of attendance at religious services; strength and comfort derived from religion; number of people known in congregation; self-rated religiousness</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reynolds &amp; Nelson (1981)</td>
<td>Convenience sample of chronically ill veterans from two medical centers in Los Angeles, CA</td>
<td>180</td>
<td>100</td>
<td>70.1</td>
<td>41</td>
<td>Missing</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ringdal (1996)</td>
<td>Convenience sample of cancer patients, University of Trondheim, Norway</td>
<td>216</td>
<td>55</td>
<td>57</td>
<td>21.25</td>
<td>Two-item measure of private religiousness (α = .71)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rogers (1996)</td>
<td>National probability sample of adults, age &gt; 55, in National Health Interview Survey</td>
<td>15,938</td>
<td>43</td>
<td>69.42</td>
<td>90</td>
<td>Church or temple attendance in past 2 weeks</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rosenblatt (1996)</td>
<td>Type II diabetes patients undergoing dialysis in Dallas and Bexar County, TX</td>
<td>333</td>
<td>58.3</td>
<td>60.8</td>
<td>35.9</td>
<td>Four-item measure of religiousness (item factor loadings &gt; .73)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 1 (continued)

<table>
<thead>
<tr>
<th>Reference</th>
<th>Sampling</th>
<th>N</th>
<th>% male</th>
<th>Mean age at baseline (months)</th>
<th>Length of follow-up (months)</th>
<th>Measures of religious involvement</th>
<th>Variables controlled in effect size*</th>
<th>Adjusted effect size (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Schoenbach, Kaplan, Fiedman, &amp; Kleinbaum (1986)</td>
<td>Random sample of White males, ages 15–39, in Evans County, GA</td>
<td>603</td>
<td>100</td>
<td>51.3</td>
<td>144</td>
<td>Spending spare time in church activities</td>
<td>1, 2, 3, 4, 8, 9, 11, 13, 15, 16</td>
<td>1.43 (0.99–2.07)</td>
</tr>
<tr>
<td>Schoenbach et al. (1986)</td>
<td>Random sample of White females, ages 15–39, in Evans County, GA</td>
<td>719</td>
<td>0</td>
<td>52.5</td>
<td>144</td>
<td>Spending spare time in church activities</td>
<td>1, 2, 3, 5, 8, 9, 11, 13, 15, 16</td>
<td>1.11 (0.77–1.62)</td>
</tr>
<tr>
<td>Schoenbach et al. (1986)</td>
<td>Random sample of Black males, ages 15–39, in Evans County, GA</td>
<td>297</td>
<td>100</td>
<td>53.6</td>
<td>144</td>
<td>Spending spare time in church activities</td>
<td>1, 2, 3, 4, 8, 9, 11, 13, 15, 16</td>
<td>0.71 (0.44–1.15)</td>
</tr>
<tr>
<td>Schoenbach et al. (1986)</td>
<td>Random sample of Black females, ages 15–39, in Evans County, GA</td>
<td>438</td>
<td>0</td>
<td>53.5</td>
<td>144</td>
<td>Spending spare time in church activities</td>
<td>1, 2, 3, 5, 8, 9, 11, 13, 15, 16</td>
<td>1.40 (0.90–2.17)</td>
</tr>
<tr>
<td>Spiegel et al. (1983)</td>
<td>Consecutive sample of women with metastatic carcinoma of the breast</td>
<td>54</td>
<td>0</td>
<td>54.5</td>
<td>12</td>
<td>Four-item measure of family moral–religious orientation</td>
<td>1</td>
<td>0.70 (0.22–2.27)</td>
</tr>
<tr>
<td>Strawbridge et al. (1997)</td>
<td>Random sample of adults, ages 21–65, in Alameda County, CA</td>
<td>2,497</td>
<td>100</td>
<td>39.8</td>
<td>336</td>
<td>Frequency of religious attendance</td>
<td>1, 2, 3, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16</td>
<td>1.14 (0.91–1.42)</td>
</tr>
<tr>
<td>Strawbridge et al. (1997)</td>
<td>Random sample of adults, ages 21–65, in Alameda County, CA</td>
<td>2,789</td>
<td>0</td>
<td>39.8</td>
<td>336</td>
<td>Frequency of religious attendance</td>
<td>1, 2, 3, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16</td>
<td>1.58 (1.09–2.29)</td>
</tr>
<tr>
<td>Yates et al. (1981)</td>
<td>Sample of patients at Vermont Regional Cancer Center in Burlington, VT</td>
<td>36</td>
<td>39.4</td>
<td>59</td>
<td>12</td>
<td>Ten-item measure of religious beliefs; affiliation with a church (yes–no); importance of church; frequency of church attendance; feelings of closeness to God or nature</td>
<td>—</td>
<td>1.39 (0.37–5.17)</td>
</tr>
<tr>
<td>Zuckerman, Kasl, &amp; Ostfield (1984)</td>
<td>Men in poor physical health, age &gt; 62, in New Haven, CT</td>
<td>52</td>
<td>100</td>
<td>72</td>
<td>24</td>
<td>Three-item measure of public–private religiousness</td>
<td>1, 6, 8</td>
<td>3.07 (0.83–11.28)</td>
</tr>
<tr>
<td>Zuckerman et al. (1984)</td>
<td>Women in poor physical health, age &gt; 62, in New Haven, CT</td>
<td>151</td>
<td>0</td>
<td>72</td>
<td>24</td>
<td>Three-item measure of public–private religiousness</td>
<td>1, 6, 8</td>
<td>2.03 (0.72–5.73)</td>
</tr>
<tr>
<td>Zuckerman et al. (1984)</td>
<td>Men in good physical health, age &gt; 62, in New Haven, CT</td>
<td>77</td>
<td>100</td>
<td>72</td>
<td>24</td>
<td>Three-item measure of public–private religiousness</td>
<td>1, 6, 8</td>
<td>1.05 (0.26–4.24)</td>
</tr>
<tr>
<td>Zuckerman et al. (1984)</td>
<td>Women in good physical health, age &gt; 62, in New Haven, CT</td>
<td>118</td>
<td>0</td>
<td>72</td>
<td>24</td>
<td>Three-item measure of public–private religiousness</td>
<td>1, 6, 8</td>
<td>1.45 (0.20–10.57)</td>
</tr>
</tbody>
</table>

Note. Dashes indicate that no variables were controlled. CI = confidence interval.

*1 = sex; 2 = race; 3 = age; 4 = income–socioeconomic status; 5 = education; 6 = functional health; 7 = global subjective health appraisals; 8 = clinical or biomedical measures of physical health; 9 = social support; 10 = social activities; 11 = smoking; 12 = alcohol use; 13 = obesity–body mass; 14 = mental health, mood, or affect; 15 = exercise; 16 = marital status; 17 = employment status.
that zero represented the value for a typical study (0 = community sample, 1 = clinical sample) or a study whose measurement of religion would be expected to capture the most health-relevant variance (0 = public measure of religious involvement, 1 = other measures).

Study characteristics. Table 2 shows the regression coefficients and associated standard errors for the theory-derived moderators. The fact that the coefficient for the intercept \( \gamma_0 \) is significant \((p < .001)\) indicates that it is unlikely that the population effect size for our "typical" study is 0 (log odds). On the contrary, in a study with a score of zero on all moderator variables, we should expect to find a positive association between religiousness and longevity—the log odds of .3650 corresponds to an odds ratio of 1.44 (95% CI: 1.31-1.58), or a 44% higher odds of survival in the religious as compared with the less religious group.

The regression weights for the moderator variables indicate the extent to which each of these study characteristics would be expected to influence the observed effect size. Of the two study design characteristics, only the number of statistical adjustments was related to the size of the observed effect. Better-controlled studies (i.e., those including more covariates or copredictors) had smaller log odds ratios. This result is as predicted: Adjusted effect sizes (after controlling for mediators or confounds) are expected to be smaller than zero-order (unadjusted) effect sizes. Of the sample characteristics variables, the proportion of males in the sample was significantly related to effect size: As the proportion of males in a sample increased, the expected association between religiousness and mortality decreased. This result suggests that religious involvement might be a stronger protective factor for women than for men.

The type of measure used to assess religious involvement was also significantly associated with observed effect size. Because we regarded public measures of religious involvement as most likely to capture health-relevant variance in religiousness, we dummy coded this four-category variable so that public measures would fall into the 0 category on each dummy variable. All regression weights are negative, indicating that use of other measure types is likely to reduce the observed effect size. To clarify this relation, we repeated the analysis with a single indicator of measure type: a contrast between public measures (0) and all other measure types (1). All other theory-derived moderators were in the regression equation as before. The regression weight for measure type in this latter analysis was \( \gamma = -0.3179, SE(\gamma) = 0.1041, p = .005 \). A study using a nonpublic measure of religious involvement is predicted to have a substantially lower effect size, corresponding to an odds ratio of 1.04, compared with an odds ratio of 1.43 for studies indexing religious involvement by self-reports of public religious behaviors.

Substantial between-studies variance remained unaccounted for by the theoretical moderators, \( \tau = .0087, \chi^2(35) = 55.41, p = .015 \). This corresponds to a Birge ratio of 1.58 (i.e., 58% more between-studies variance than would be expected by chance in contrast to a Birge ratio of 2.23 for the omnibus model), indicating a substantial reduction in unexplained effect size variation. The chi-square difference test comparing this model with the omnibus model shows a significant increase in explanatory power, \( \Delta \chi^2(6) = 36.21, p < .001 \), with the moderators accounting for 58% of the random-effects variance among the 42 effect sizes.

Exploratory analyses on the effect sizes for public measures. The strong effect of type of religious measure in the preceding moderator analyses suggests that the positive association of religion and mortality is derived largely from (public) participation in religious organizations rather than from (private) religious attitudes and beliefs alone. To examine the association of public religious involvement and mortality more carefully, we conducted exploratory analyses with the \( (k = 21) \) effect sizes \((N = 107,910)\) involving public measures of religiousness. To avoid extremely high Type II error rates in these exploratory analyses, we chose to tolerate an increased risk of Type I errors and interpreted as marginally significant any moderator effect with a probability greater than or equal to .20. In an unconditional model involving the 21 effect sizes involving measures of public religiousness, the intercept was \( \gamma_0 = .3121, SE(\gamma_0) = .0404, p < .001, \) odds ratio = 1.37.

Then, we examined the moderating effects of study characteristics as we did with all 42 effect sizes. We excluded the dummy variable contrasting community and clinical samples because all of the studies using public measures of religious involvement involved community samples. For obvious reasons, we also excluded the three dummy variables representing the types of measures of religious involvement. The only study characteristic that was associated with effect size was percentage of males in the sample, \( \gamma = -0.0020, SE(\gamma) = .0009, p = .046 \). For a study with a gender breakdown typical of these samples (i.e., 56% males), the intercept was \( \gamma_0 = .3045, SE(\gamma_0) = .0359, p < .001, \) odds ratio = 1.36.

Given the diversity of covariates and copredictors of mortality included in the primary studies, we set out to compare the effect sizes from studies that controlled for each of 15 variables (race, income, education, employment status, functional health, global health appraisals, clinical or biomedical measures of physical health, social support, social activities, marital status, smoking, alcohol use, obesity—body mass index, mental health or affective distress, and exercise) with effect sizes from studies that did not control for each respective variable (0 = controlled, 1 = not

Table 2
Random-Effects Regression Weights for Design Characteristics Associated With 42 Effect Sizes

<table>
<thead>
<tr>
<th>Parameter</th>
<th>( \gamma )</th>
<th>( SE(\gamma) )</th>
<th>( p )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>.3650</td>
<td>.0470</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Length of follow-up (months)</td>
<td>.0006</td>
<td>.0005</td>
<td>.252</td>
</tr>
<tr>
<td>No. of statistical adjustments</td>
<td>-.0180</td>
<td>.0085</td>
<td>.041</td>
</tr>
<tr>
<td>% male</td>
<td>-.0018</td>
<td>.0008</td>
<td>.043</td>
</tr>
<tr>
<td>M age at baseline</td>
<td>.0043</td>
<td>.0029</td>
<td>.149</td>
</tr>
<tr>
<td>Community (0) vs. clinical (1)</td>
<td>-.0010</td>
<td>.1737</td>
<td>.995</td>
</tr>
<tr>
<td>Measurement of religiousness*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Private (1) vs. others (0)</td>
<td>-.1435</td>
<td>.2053</td>
<td>.489</td>
</tr>
<tr>
<td>Mixed (1) vs. others (0)</td>
<td>-.3077</td>
<td>.1070</td>
<td>.007</td>
</tr>
<tr>
<td>Missing (1) vs. others (0)</td>
<td>-.4369</td>
<td>.2238</td>
<td>.059</td>
</tr>
</tbody>
</table>

* Each religion measure was coded into one of four categories (public, private, mixed, and missing). For the regression analyses, these four categories were converted into three dummy variables (measures of private religious involvement, measures that combined public and private measures of religious involvement, and measures that were insufficiently described) so that public measures would fall into the 0, or other, category for each dummy variable.
controlled. We conducted 15 separate moderator analyses. In these analyses, we entered the percentage male variable simultaneously with individual control variables into a series of moderator models. Among the 21 effect sizes, obesity—body mass index was the only control variable that was associated even marginally with effect size, $\gamma = .156, SE(\gamma) = .0706, p = .118$. A study that controlled for obesity—body mass index in a sample that was 56% male would be expected to yield an odds ratio of 1.26, whereas a similar study that did not control for obesity—body mass index would be expected to yield an odds ratio of 1.42.

At a reviewer's request we also examined the aggregate effect size when all 15 control variables were controlled simultaneously. The purpose of these analyses was to address whether the relation between public religious involvement and mortality could be attributed to some combination of sociodemographic differences, initial health status differences, differences in health behaviors, and differences in social support between religious and nonreligious groups.

We conducted a series of four regression models in which classes of control variables (i.e., sociodemographics, physical health, health behaviors, and social support) were added systematically. We encountered problems with multicollinearity among these control variables, but we included as many control variables within each class as was empirically possible. The predictor-to-case ratio increased threefold (i.e., from a 4-to-21 to a 12-to-21 ratio) from the first to the fourth model. As a result, each successive model yielded coefficients with larger standard errors and, consequently, lower statistical power. Nevertheless, these analyses are helpful for modeling how the association of public religious involvement and mortality might change as greater numbers of possible confounds and mediators of the association are controlled statistically.

The intercept ($\gamma_0$) in each model reflects the expected log odds ratio for a study with 56% males, controlling for all included moderators. The first model, including percentage male, race, income, and education, yielded $\gamma_0 = .2650, SE(\gamma_0) = .0623, p = .001$, corresponding to an odds ratio of 1.30. No sociodemographic control variable was associated with effect size (all $p > .20$). The second model including (a) the sociodemographic variables entered in the previous model and (b) functional and clinical-biomedical measures of physical health yielded $\gamma_0 = .2298, SE(\gamma_0) = .0870, p = .020$, corresponding to an odds ratio of 1.26. None of the control variables was associated with effect size (all $p > .20$). The third model including (a) the sociodemographic control variables and health variables included in the previous model and (b) smoking, alcohol use, and obesity—body mass yielded $\gamma_0 = .1886, SE(\gamma_0) = .0990, p = .083$, corresponding to an odds ratio of 1.21. In this model, control for smoking ($\gamma = -.2700$) and alcohol use ($\gamma = -.2833$) were marginally associated with effect size ($p = .144$ and .104, respectively). Studies that did control for smoking and alcohol use yielded larger effect sizes than studies that did not control for smoking and alcohol use. This finding is counterintuitive and probably reflects sampling variation rather than any substantive effects. The fourth model including (a) the sociodemographic, health, and health behavior control variables included in the previous model and (b) social support, social activities, and marital status yielded $\gamma_0 = .2031, SE(\gamma_0) = .1853, p = .306$, corresponding to an odds ratio of 1.23.

Although the power of the significance tests in these analyses was low due to the small number of effect sizes, it appears that these general classes of variables account for part of the religion—mortality association. A study that controlled sociodemographics, physical health, health behaviors, and social support would be expected to demonstrate a smaller, but still substantial, association between public religious involvement and mortality.

Publication Bias and Sensitivity Analyses

The studies that are practically available for inclusion in a meta-analysis (i.e., those studies obtainable by the meta-analysts) may not be a representative sample of the studies conducted in the research domain. Indeed, the most easily obtained studies (i.e., those available in journals) tend to be biased toward positive results (Becker, 1994). This creates the potential for publication bias, also called the file drawer problem (Begg, 1994; Rosenthal, 1979).

We used several methods for evaluating the possible impact of publication bias on our findings. First, we examined a graphical display of the effect sizes as a function of their sample size. A roughly funnel-shaped display suggests that the meta-analytic data points represent an unbiased, representative sample from the population of relevant studies (Begg, 1994). The funnel-shaped distribution should occur because studies with small sample sizes have greater sampling variability, and thus, greater interstudy variability in their estimates of the population effect size, whereas studies with larger sample sizes have less sampling variability and, thus, should estimate more accurately the population effect size. By contrast, a graph that is skewed (to the right) toward more positive effect sizes for smaller sample studies suggests bias due to overreliance on published studies; the presumption here is that a number of small-sample studies that exist with less favorable effect sizes are missing from the meta-analytic sample. The display of effect sizes (log odds ratios) as a function of sample size conformed to a funnel shape (see Figure 1).

Second, we used the formulas presented in Begg (1994) to examine the correlation between the ranks of standardized effect sizes and the ranks of their sampling variances. Using the Spearman rank correlation coefficient, $r(42) = -.07, p > .30$, one-tailed. Using Kendall's rank correlation coefficient, $\tau(42) = -.06, p > .25$, one-tailed. These near-zero rank correlations also suggest little or no publication bias.

Third, we calculated Rosenthal's (1979) fail-safe $N$, which estimates the number of file drawer studies, averaging null results, that would be required to overturn an observed pattern of meta-analytic results (i.e., if the file drawer studies had been included). We calculated a fail-safe $N$ for the omnibus analysis ($k = 42$ effects) based on formulas given in Begg (1994), which is a function of the $z$ values associated with each of the effect sizes included in the meta-analysis. This revealed that 1,418 effect sizes with a mean odds ratio of 1.0 (i.e., literally no relationship of religious involvement and mortality) would be needed to overturn the significant overall association of religious involvement and mortality (i.e., to render the resulting mean effect size nonsignificant, $p > .05$, one-tailed) that we found in our omnibus analyses.

Begg (1988) also noted that publication bias is most likely in meta-analyses of research domains that consist of many studies with small sample sizes. In contrast, our search for relevant studies
yielded only 42 effect sizes with a mean sample size of 2,996. These converging lines of evidence suggest that our conclusions are relatively safe from publication bias. However, readers are invited to send unpublished or published study results that were not included in the present review to Michael E. McCullough. Submitted data will be included in a future update to the present review and will help in ruling out publication bias as an explanation for the present results.

**Discussion**

In the course of an extensive literature search, we identified 42 independent effect sizes based on samples of nearly 126,000 people that represented the association of religious involvement and all-cause mortality. Most ($k = 23$) of these effect sizes were based on single-item measures of religious attendance or subjective religiousness with limited reliability, even though superior tools for assessing religious involvement are widely available (Critt & Hood, 1999). Unreliability attenuates the association of the measured variable with other variables of interest (e.g., mortality), yielding smaller effect sizes than would be observed had variables been measured without error (Hunter & Schmidt, 1990). Thus, the effect sizes reported here should be considered conservative estimates of the association of religious involvement and mortality.

**Association Between Religious Involvement and All-Cause Mortality**

Despite such psychometric limitations, the meta-analysis indicated that the odds of survival for people who scored higher on such measures of religious involvement (after statistical control) were 129% of the odds of survival for people who scored lower on such measures. An odds ratio of this size is equivalent to a tetrachoric correlation of .10 (Davidoff & Goheen, 1953). This effect size is considered small by Cohen's (1988) rules of thumb for the behavioral sciences. Nonetheless, the religious involvement–mortality association may have considerable practical significance given the importance of the criterion variable (i.e., mortality) and the number of people in the population who are potentially exposed to religion (Rosenthal, 1990). Although the strength of the association varied as a function of several moderator variables, the basic finding was robust: Religious involvement is associated with higher odds of survival (or conversely, lower odds of death) during any specified follow-up period. These findings could not be attributed to publication bias.

**Moderator Variables: Explaining the Association of Religious Involvement and Mortality**

Our moderator analyses helped to clarify the nature of the relation between religious involvement and mortality. The following explanations are offered with circumspection, however, because they are derived by interpreting multivariate correlational data gleaned from a fairly small sample of studies (Hedges, 1994; Hunter & Schmidt, 1990).

**Study characteristics.** As expected, studies exerting the greatest statistical control yielded the least favorable associations of religious involvement and mortality. This finding suggests that the association of religious involvement and mortality can be explained in part as a function of other demographic, psychosocial, or health-related variables. For example, studies that failed to control for obesity–body mass yielded more favorable effect size estimates than did those that did control for obesity–body mass. There is some evidence that people with high levels of religious involvement are less obese (Baecke, Burema, Frijters, Hautvast, & van der Wiel-Wetzels, 1983), suggesting that people who are religious might avoid early death in part via lower obesity (but cf. Strawbridge et al., 1997). Therefore, researchers should include obesity–body mass index in their models to estimate the extent to which religious involvement obtains its association with mortality through obesity–body mass.

**Sample characteristics.** The percentage of males in the study sample was the only characteristic we examined that was related to
effect size. Every 1% increase in males within a study sample is expected to yield a reduction of 0.0018 in the observed log odds ratio. Thus, a sample with 100% males (44 percentage points higher than the mean of 56%) would be expected to yield an effect size of 0.3650 - (44 × 0.0018) = 0.2858, or an odds ratio of 1.33, compared with a sample of 100% females, with a predicted effect size of 0.3650 + (56 × 0.0018) = 0.4658, or an odds ratio of 1.59. Thus, the favorable association of religious involvement and mortality appears to be considerably greater for women than for men. This gender difference might be due to differences in the psychosocial resources that men and women receive from religious involvement. Because women live longer than men and tend to be more religious than men (Levin & Chatters, 1998; Levin & Taylor, 1997), researchers should control for sex statistically or estimate models separately for men and women to prevent confounding.

**Measures of religious involvement.** Studies using public measures of religious involvement yielded larger effect sizes than did those using other types of measures of religious involvement. This finding is consistent with speculations that the health-related effects of religious involvement are due partially to the psychosocial resources derived from frequent attendance at religious services, membership in religious groups, or involvement with other (religious) people (Goldbourt et al., 1993; Idler & Kasl, 1997a).

The particularly favorable association of public religious involvement and mortality might also be, in part, due to what Levin and Vanderpool (1987) identified as a proxy effect (i.e., a confounding of public religious involvement with physical functioning). Although we found no evidence that the association of religious involvement and mortality was stronger in studies that did not control for physical health, researchers should take care to control baseline physical health functioning in future research, lest the true association of religious involvement and mortality be overestimated. Indeed, researchers who investigate religion and mortality in the future should endeavor to control for all of the sociodemographic, social, and health variables that are known to be risk factors for early death. Some of these variables (e.g., race, gender, age, and probably physical health status) are confounds of the relationship between religious involvement and mortality. Others (including social support, social activities, and health behaviors) could be confounds or mediators of the religion—mortality relationship. In either case, researchers will paint an accurate picture of the religion—mortality association only when they are careful to measure and model these potential confounds and mediators adequately.

**Conclusion**

Although the correlational nature of the data prohibit causal inferences, religious involvement has a nontrivial, favorable association with all-cause mortality. This association is stronger in studies in which women constitute the majority of participants, there is inadequate control of other covariates of mortality, and measures of public religious involvement are used. Although part of the religious involvement—mortality association may be a product of confounding, much of the association may be substantive, perhaps mediated by health-promotive behaviors, such as maintaining a healthy body mass.

Given these conclusions—based on a meta-analytic sample representing nearly 126,000 participants—future researchers interested in these issues should probably not focus exclusively on exploring whether an association exists but should also explore the mechanisms through which religious involvement obtains a favorable association with mortality. To advance this research agenda, researchers should use more reliable measures of multiple dimensions of religious involvement (e.g., public religious involvement, private religious activities, religious beliefs, religious motivations, and religious coping). In addition, more sophisticated statistical methods (i.e., structural equation modeling) should be used to model the mechanisms (including substantive mechanisms, such as psychosocial or physiological pathways, as well as methodological mechanisms such as confounding) by which religious involvement could obtain its associations with mortality. Potential confounds that should be modeled include age, race, gender, and physical health. Potentially substantive pathways might include reductions in risky behaviors such as smoking, drug use, alcohol use, obesity, and unsafe sexual practices (e.g., see Benson, 1992); improvements in social support and marital–family stability (Ellison & George, 1994); and positive attitudes and emotions that are associated both with physical health and with religious involvement (e.g., Kark, Carmel, Sinnreich, Goldberger, & Friedlander, 1996; Myers & Diener, 1995; Witter, Stock, Okun, & Haring, 1985).

Given the large numbers of people who are religiously active, the favorable association of religious involvement and mortality is a health phenomenon with some relevance for a substantial proportion of the American population. Elucidating the nature of this robust but poorly understood association could be a fruitful topic for future research at the interface of psychology and health.

**References**

References marked with an asterisk indicate studies included in the meta-analysis.


*Kune, G., Kune, S., & Watson, L. (1992). The effect of family history of


