Prospective, observational studies from many settings have shown that depression is independently associated with an increased risk for death (1–5). However, few studies have evaluated whether an intervention focused on depression can modify this risk. Intervention studies that have reported on death in patients with depression have serious limitations: They lacked randomization to the intervention condition (6–8), focused only on patients after cardiovascular events (9–12), or adjusted for treatment in examining the association of depression with death rather than studying the effect of treatment (6, 13).

In our study, we analyzed the relationship between a depression care management intervention and the risk for death among older primary care patients during a 5-year interval. We used data from the Prevention of Suicide in Primary Care Elderly: Collaborative Trial (PROSPECT), which we supplemented with data from the National Death Index (NDI). The PROSPECT was an effectiveness study designed to assess the effect of care management on reducing risk factors for late-life suicide (14). It integrated a randomized trial with a population-based, public health model. The study intervention was implemented at the practice level and involved a depression care manager working with physicians to provide algorithm-based care (14). The primary care practice was the unit of randomization, and practices that were not randomly assigned to the intervention were expected to provide usual care. Patients were not assigned to receive specific treatments, so patients and their physicians in the intervention and usual care practices decided whether patients would receive depression treatment. Our research question focused on whether this practice-level intervention influenced a patient-level outcome, namely survival. We hypothesized that older adults with depression in practices randomly assigned to a depression management intervention would experience an attenuated risk for death compared with those patients in usual care practices. Guided by published criteria for performing and reporting subgroup analyses (15, 16), we were particularly interested in whether the effects of care management were specific to patients meeting standard criteria for major depression (Diagnostic and Statistical Manual for Mental Disorders, fourth edition [DSM-IV] [17]).
Persons with depression are more likely to die, but studies have not shown that treatment of depression reduces mortality.

Investigators observed a 45% reduction in the hazard of death among patients with major depression cared for in primary care practices that were randomly assigned to a depression care management program.

The reduction in deaths occurred almost exclusively among patients who died of cancer. The mechanism for the effect is unclear and might be due to misclassification of cause of death or vital status.

A practice-level depression care management program seemed to reduce deaths due to cancer in older patients with major depression.

---

Methods

Study Sample

The PROSPECT was conducted in 20 primary care practices located in greater New York, New York, and Philadelphia and Pittsburgh, Pennsylvania, from May 1999 to August 2001, with individual patients followed for 2 years. After being paired by urban location, academic affiliation, size, and population type, practices were randomly assigned to the intervention or to usual care by coin flip (cluster randomization by practice). Patients were recruited from an age-stratified (60 to 74 years and ≥75 years), random sample of patients with upcoming appointments. Research associates confirmed study eligibility (age ≥60 years, Mini-Mental State Examination score ≥18 [18], and English-speaking) of consenting patients and screened patients for depression by using the Centers for Epidemiologic Studies Depression scale (19).

Eligibility Criteria

All patients scoring greater than 20 on the Centers for Epidemiologic Studies Depression scale were invited to enroll. Patients from a 5% sample with lower scores were also invited for assessment of false-negative results on screening. To increase the sensitivity of screening, the PROSPECT investigators also recruited patients with scores less than 20 and positive responses to questions about previous depression episodes or treatment. Research associates met patients at the practice, obtained written consent, and administered a baseline interview (20). The sampling strategy yielded a cohort that approximated a representative sample of attendees of the PROSPECT practices, with oversampling of patients with depression symptoms.

Depression Assessment

We conducted a Structured Clinical Interview for DSM-IV Disorders Axis I (SCID-I) and allowed for the full range of DSM-IV diagnoses for depression disorders (17, 21). Our investigation focused on patients targeted by the intervention, that is, patients who met DSM-IV criteria for major depression (17) or who had clinically significant minor depression, defined by DSM-IV criteria for minor depression that we modified by requiring 4 depression symptoms, Hamilton Depression Rating Scale (HDRS) score of 10 or more, and duration of 4 weeks or more (22). To minimize variation in applying criteria for depression, the Cornell Advanced Center for Intervention and Services Research, White Plains, New York, conducted regular teleconferences with research associates to review diagnostic practices and conduct reliability assessments. Ongoing monitoring indicated excellent reliability within and across sites for SCID-I assessments (intraclass correlation, 0.78 to 1.00).

Assessment of Patient Characteristics

We obtained baseline information on age, sex, marital status, self-reported ethnicity, educational attainment, and smoking status (according to tobacco use within the past 6 months). We included ethnicity in describing our sample because it has been associated with patterns of mental health service use (23, 24). Patients self-reported having a medical comorbid condition according to the Charlson Comorbidity Index (25). The 24-item HDRS measured depression severity (26), and the Scale for Suicidal Ideation indicated the presence of suicidal ideation (27).

Description of Usual Care and Intervention Groups

Practices randomly assigned to usual care received educational sessions for primary care physicians and notification of the depression status of their patients. No specific recommendations were given to physicians about individual patients, except for psychiatric emergencies. The following were made available to practices randomly assigned to the intervention: educational sessions for primary care physicians, education for patients’ families, and a depression care manager who worked within the practice. The care manager implemented the intervention by reviewing patients’ depression status, medical history, and medication use and subsequently worked with the primary care physician to recommend treatment according to standard guidelines.

Care managers had psychiatric backup, including on-demand consultation, weekly supervision by psychiatrist-investigators, and monthly interpersonal therapy cross-site supervision. They were introduced to patients by research associates immediately after the baseline interview. The 15 care managers included social workers, nurses, and psychologists who interacted with patients in person or by telephone at scheduled intervals and as necessary. Care managers focused efforts on depression treatment (not on care for medical conditions or preventive services) by mon-
Monitoring symptoms, adverse effects of medications, and treatment adherence.

All patients with depression received citalopram for first-line treatment. Citalopram therapy was initiated at 10 mg before bedtime on the first day, 20 mg/d for the next 6 days, and 30 mg/d subsequently. After 6 weeks, the target dose was sustained for 8 weeks.

Figure 1. Study flow diagram.

Primary practice units (n = 20)

Randomly assigned to Intervention (n = 10)

Active patients (n = 9822)

Sampled for screening (n = 8048)

Screened (n = 4243)

Assessed for enrollment (n = 936)
  Positive results on CES-D screening: 524
  Random sample of patients with negative results on CES-D screening: 255
  Negative results on CES-D screening, positive history of depression: 157

Excluded (n = 320)
  Ineligible: 8
  Declined: 312

Enrolled (n = 616)

Insufficient baseline data (n = 7)

Major depression at baseline (n = 214)
  Died (n = 39)

Minor depression at baseline (n = 106)
  Died (n = 21)

Not depressed at baseline (n = 289)
  Died (n = 52)

Randomly assigned to usual care (n = 10)

Active patients (n = 11363)

Sampled for screening (n = 8660)

Screened (n = 4829)

Assessed for enrollment (n = 952)
  Positive results on CES-D screening: 537
  Random sample of patients with negative results on CES-D screening: 250
  Negative results on CES-D screening, positive history of depression: 165

Excluded (n = 330)
  Ineligible: 13
  Declined: 317

Enrolled (n = 622)

Insufficient baseline data (n = 5)

Major depression at baseline (n = 182)
  Died (n = 38)

Minor depression at baseline (n = 97)
  Died (n = 17)

Not depressed at baseline (n = 338)
  Died (n = 56)

CES-D = Centers for Epidemiologic Studies Depression scale; NDI = National Death Index.
Table 1. Baseline Characteristics*

<table>
<thead>
<tr>
<th>Variable</th>
<th>Intervention Group (n = 320)</th>
<th>Usual Care Group (n = 279)</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Demographic characteristics</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean age (SD), y</td>
<td>71 (7.8)</td>
<td>70 (8.1)</td>
<td>0.144</td>
</tr>
<tr>
<td>Women, n (%)</td>
<td>221 (69)</td>
<td>208 (75)</td>
<td>0.139</td>
</tr>
<tr>
<td>Ethnic minority, n (%)†</td>
<td>93 (29)</td>
<td>103 (37)</td>
<td>0.69</td>
</tr>
<tr>
<td>Mean education (SD), y</td>
<td>13 (3.3)</td>
<td>13 (5.1)</td>
<td>0.37</td>
</tr>
<tr>
<td>Married, n (%)</td>
<td>116 (36)</td>
<td>104 (38)</td>
<td>0.53</td>
</tr>
<tr>
<td>Habit, n (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Current smoker</td>
<td>35 (11)</td>
<td>15 (5)</td>
<td>0.051</td>
</tr>
<tr>
<td>Medical conditions, n (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cardiovascular disease</td>
<td>150 (48)</td>
<td>115 (42)</td>
<td>0.131</td>
</tr>
<tr>
<td>Stroke</td>
<td>78 (25)</td>
<td>55 (20)</td>
<td>0.164</td>
</tr>
<tr>
<td>Diabetes</td>
<td>70 (23)</td>
<td>93 (19)</td>
<td>0.40</td>
</tr>
<tr>
<td>Gastrointestinal tract disease</td>
<td>88 (28)</td>
<td>74 (27)</td>
<td>0.91</td>
</tr>
<tr>
<td>Cancer</td>
<td>41 (13)</td>
<td>48 (17)</td>
<td>0.068</td>
</tr>
<tr>
<td>Respiratory diseases</td>
<td>42 (14)</td>
<td>43 (16)</td>
<td>0.36</td>
</tr>
<tr>
<td>Cognition and depression†</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean MMSE score for cognitive function (SD)</td>
<td>27 (2.9)</td>
<td>27 (2.5)</td>
<td>0.82</td>
</tr>
<tr>
<td>Major depression disorder, n (%)</td>
<td>214 (67)</td>
<td>182 (65)</td>
<td>0.74</td>
</tr>
<tr>
<td>Mean HDRS score for depression severity (SD)</td>
<td>19 (6.1)</td>
<td>18 (5.8)</td>
<td>0.22</td>
</tr>
<tr>
<td>Suicidal ideation (SSI score &gt;0, n (%)</td>
<td>94 (29)</td>
<td>56 (20)</td>
<td>0.010</td>
</tr>
</tbody>
</table>

* Data from the Prevention of Suicide in Primary Care Elderly: Collaborative Trial (1999 to 2004) (14). DSM-IV = Diagnostic and Statistical Manual of Mental Disorders, fourth edition; HDRS = Hamilton Depression Rating Scale; MMSE = Mini-Mental State Examination; SSI = Scale for Suicidal Ideation.
† Defined as an ethnicity other than non-Hispanic white (total, n = 159; Hispanic, n = 26; non-Hispanic black, n = 159; Asian, n = 4; and other non-Hispanic, n = 6).
‡ Major depression disorder defined as DSM-IV major depression in contrast to clinically significant minor depression (defined as 4 DSM-IV symptom groups, HDRS score ≥10, and ≤4-week duration). Score range for MMSE is 0 to 30 (inclusion criteria limited the range to 18 to 30), with high scores indicating less cognitive impairment. Score range for HDRS is 0 to 76, with high scores indicating greater depressive symptoms. Score range for SSI is 0 to 38, with high scores indicating greater suicidal ideation.

Dosage was maintained if the patient exhibited a substantial improvement (≥50% reduction in the HDRS score) (26) and was increased if the patient exhibited a partial improvement (30% to 50% reduction in the HDRS score). Nonresponders, for whom guidelines called for switching to another antidepressant, were defined as patients who did not demonstrate either minimal improvement after 6 weeks of treatment at the target dosage or substantial improvement after the dose was increased to the maximum recommended dose after 12 weeks of treatment (28). For patients who had not responded at 12 weeks, the health specialist followed guidelines for switching antidepressants (28). Health specialists informed patients and their family members about the possible occurrence of specific side effects and to contact them if side effects occurred. When side effects occurred, health specialists provided support and, if warranted, asked primary care physicians to adjust doses or time of administration or to institute symptomatic treatment (28). The PROSPECT treatment algorithm also took account whether patients were already being treated for depression. For example, for patients already receiving pharmacotherapy who remained symptomatic, the care manager optimized the current antidepressant before switching patients to another antidepressant (28).

Interpersonal therapy could be used alone or as an augmentation strategy, depending on whether the patient tolerated the antidepressant therapy and on the presence or absence of a partial response. Care managers from each site, all of whom had some experience in psychotherapy for depression, participated in interpersonal therapy training at the University of Pittsburgh Medical Center, Pittsburgh. In both study groups, physicians were informed by letter if patients reported any suicidal ideation and immediately when patients were identified as being at high risk for suicide according to prespecified guidelines. Other sources detail the PROSPECT treatment algorithm and implementation, including the role of the care manager (29), strategy for pharmacotherapy (28), management of suicidal ideation (30, 31), and types and proportions of treatment received over time by patients in practices randomly assigned to the intervention or to usual care (14, 32).

Ascertainment of Vital Status

We used the National Center for Health Statistics NDI Plus (33) to assess vital status. The underlying causes of death that we obtained from NDI Plus are similar to codes assigned by trained nosologists (33, 34). Because querying the NDI required that we provide the National Center for Health Statistics with personal identifiers (for example, Social Security numbers), confidentiality safeguards warrant discussion. We did not transmit any study data along with identifying data or transmit identifying data via e-mail. The 3 PROSPECT sites verified the vital status information obtained from the NDI by physician report of death and match of identifying information for each individual and sent the final version, which was indexed by unique study identifier and stripped of personal identifiers.
identifiers, to the School of Medicine of the University of Pennsylvania Data Core, Philadelphia, for production of the analytic data set. We obtained written consent, including permission to obtain death certificate information, from each patient. Our study received approval from the institutional review boards at the University of Pennsylvania, Philadelphia; University of Pittsburgh Medical School, Pittsburgh; and Weill Medical College of Cornell University, White Plains, and from independent review at the National Center for Health Statistics.

**Statistical Analysis**

Our analysis involved sorting patients into 6 groups according to baseline depression status (major, minor, or nondepressed) and practice randomization assignment (intervention or usual care) to allow for examination of the main effects and the interaction of depression status and study group by using the mortality experience of each group. We compared baseline patient characteristics across the groups by using linear and logistic regression with random effects to account for patient clustering by practice. For the death outcome, we used the Cox model, adjusting SEs for within-practice clustering (35). We began by exploring potential confounding variables by using univariate models with baseline characteristics as predictors of time to death. Our final model included influential covariates identified by their association ($P < 0.05$) with the outcome of interest, time to death. The final adjusted model included terms for baseline age, sex, education, smoking status, cardiovascular disease, stroke, diabetes, cancer, cognition, and suicidal ideation. We performed an additional analysis by using the Charlson score to adjust for medical comorbid conditions. We prepared survival curves by using the Kaplan–Meier method (36) to illustrate the mortality rate in each group.

To evaluate our prespecified study hypothesis, we used a test for effect modification of intervention assignment on the risk for death by baseline depression status. The formal test for effect modification introduced terms representing

<table>
<thead>
<tr>
<th>Medical Condition</th>
<th>Intervention Group</th>
<th>Usual Care Group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Deaths, n</td>
<td>Mortality Rate (95% CI), n/1000 person-years</td>
</tr>
<tr>
<td>All patients with depression (n = 599)</td>
<td>60</td>
<td>44.7 (34.1–57.6)</td>
</tr>
<tr>
<td>Cardiovascular disease</td>
<td>23</td>
<td>17.1 (10.9–25.7)</td>
</tr>
<tr>
<td>Stroke</td>
<td>3</td>
<td>2.2 (0.5–6.5)</td>
</tr>
<tr>
<td>Cancer</td>
<td>12</td>
<td>8.9 (4.6–15.6)</td>
</tr>
<tr>
<td>Respiratory disease</td>
<td>8</td>
<td>6.0 (2.6–11.7)</td>
</tr>
<tr>
<td>Accident</td>
<td>2</td>
<td>1.5 (0.2–5.4)</td>
</tr>
<tr>
<td>Suicide</td>
<td>1</td>
<td>0.7 (0.0–4.2)</td>
</tr>
<tr>
<td>Other diseases</td>
<td>11</td>
<td>8.2 (4.1–14.7)</td>
</tr>
<tr>
<td>Patients with major depression disorder (n = 396)</td>
<td>39</td>
<td>43.6 (31.0–59.6)</td>
</tr>
<tr>
<td>Cardiovascular disease</td>
<td>16</td>
<td>17.9 (10.2–29.1)</td>
</tr>
<tr>
<td>Stroke</td>
<td>2</td>
<td>2.2 (0.3–8.1)</td>
</tr>
<tr>
<td>Cancer</td>
<td>8</td>
<td>8.9 (3.9–17.6)</td>
</tr>
<tr>
<td>Respiratory disease</td>
<td>5</td>
<td>5.6 (1.8–13.0)</td>
</tr>
<tr>
<td>Accident</td>
<td>1</td>
<td>1.1 (0.0–6.2)</td>
</tr>
<tr>
<td>Suicide</td>
<td>0</td>
<td>0.0 (0.0–4.1)</td>
</tr>
<tr>
<td>Other diseases</td>
<td>7</td>
<td>7.8 (3.1–16.1)</td>
</tr>
<tr>
<td>Patients with clinically significant minor depression (n = 203)</td>
<td>21</td>
<td>46.9 (29.1–71.7)</td>
</tr>
<tr>
<td>Cardiovascular disease</td>
<td>7</td>
<td>15.6 (6.3–32.2)</td>
</tr>
<tr>
<td>Stroke</td>
<td>1</td>
<td>2.2 (0.1–12.5)</td>
</tr>
<tr>
<td>Cancer</td>
<td>4</td>
<td>8.9 (2.4–22.9)</td>
</tr>
<tr>
<td>Respiratory disease</td>
<td>3</td>
<td>6.7 (1.4–19.6)</td>
</tr>
<tr>
<td>Accident</td>
<td>1</td>
<td>2.2 (0.1–12.5)</td>
</tr>
<tr>
<td>Suicide</td>
<td>1</td>
<td>2.2 (0.1–12.5)</td>
</tr>
<tr>
<td>Other diseases</td>
<td>4</td>
<td>8.9 (2.4–22.9)</td>
</tr>
<tr>
<td>Patients without depression (n = 627)</td>
<td>52</td>
<td>41.7 (31.2–54.7)</td>
</tr>
<tr>
<td>Cardiovascular disease</td>
<td>20</td>
<td>16.1 (9.8–24.8)</td>
</tr>
<tr>
<td>Stroke</td>
<td>3</td>
<td>2.4 (0.5–7.0)</td>
</tr>
<tr>
<td>Cancer</td>
<td>12</td>
<td>9.6 (5.0–16.8)</td>
</tr>
<tr>
<td>Respiratory disease</td>
<td>4</td>
<td>3.2 (0.9–8.2)</td>
</tr>
<tr>
<td>Accident</td>
<td>3</td>
<td>2.4 (0.5–7.0)</td>
</tr>
<tr>
<td>Suicide</td>
<td>0</td>
<td>0.0 (0.0–3.0)</td>
</tr>
<tr>
<td>Other diseases</td>
<td>10</td>
<td>8.0 (3.8–14.8)</td>
</tr>
</tbody>
</table>

* Data from the Prevention of Suicide in Primary Care Elderly: Collaborative Trial (1999 to 2004) (14).
interaction and main effects for depression status and intervention condition into the Cox model. We assessed the proportional hazards assumption by including time-dependent terms in the unadjusted model and measuring the global effect (37). We used SAS, version 9.1 (SAS Institute, Inc., Cary, North Carolina), for these analyses. We set an $\alpha$ level of 0.05 to denote statistical significance, recognizing that tests of statistical significance are approximations to aid interpretation and inference.

Role of the Funding Source
This study was funded by the National Institute of Mental Health. The funding source had no role in the design and conduct of the study; the collection, management, analysis, and interpretation of the data; or in the preparation, review, and approval of the manuscript.

RESULTS
Sample Characteristics
The Consolidated Standards of Reporting Trials (CONSORT) flow diagram for the PROSPect trial has been published elsewhere (14). Figure 1 shows a flow diagram for our investigation. After 5 years, 223 patients had died: 108 patients without depression (17%) and 115 patients with depression (19%). The median length of follow-up in ascertainment of vital status was 52.8 months (range, 0.8 to 68.4 months). Table 1 shows baseline characteristics for all patients with depression in the intervention and usual care groups. Of the baseline characteristics, suicidal ideation statistically significantly differed across groups, with a higher proportion of patients in the intervention group having 1 or more symptoms on the Scale for Suicidal Ideation.

Mortality Rates
Table 2 shows the number of deaths due to specific medical conditions and mortality rates (with 95% CIs) by group assignment stratified by baseline depression status. One documented suicide during the study was identified in data obtained from NDI Plus (Table 2). Among patients with major depression, fewer deaths due to cancer occurred among patients in the intervention group.

Risk for Death
Table 3 shows the final adjusted and unadjusted Cox proportional hazards model in the variable-specific effects on time to death. We derived all adjusted hazard ratio estimates from models that included terms for potentially influential covariates, namely age, sex, educational attainment, smoking status, baseline comorbid conditions (cardiovascular disease, stroke, diabetes, and cancer), baseline Mini-Mental State Examination score, and suicidal ideation. No statistically significant departure from the proportional hazards assumption occurred in the model (chi-square = 3.856; $P = 0.28$). Results were the same in models that adjusted for the Charlson score at baseline (not shown).

Effect of the Practice-Based Intervention on Death
Because of the statistically significant interaction between depression status and the intervention ($P = 0.031$), we used unadjusted and adjusted hazard ratios for intervention effects stratified by baseline depression status (major, minor, and nondepressed) (Table 4). For patients with major depression, the intervention was associated with a reduction in deaths (adjusted hazard ratio, 0.55 [95% CI, 0.36 to 0.84]). The intervention did not statistically significantly differ in deaths among patients with clinically significant minor depression (adjusted hazard ratio, 0.97 [CI, 0.49 to 1.92]) or among patients without depression (adjusted hazard ratio, 1.14 [CI, 0.84 to 1.53]). Figure 2 shows the Kaplan–Meier curves for randomization groups according to baseline depression status. Attrition due to dropout was reflected in the change in number of patients at risk over time (14).

DISCUSSION
Depression has been linked to increased deaths, approximately doubling the risk for death in community samples across a wide range of depression assessment strategies (1–5). In our study, compared with older adults receiving usual care, older adults with depression in practices...
randomly assigned to an intervention consisting of a depression care manager working with primary care physicians to provide algorithm-based care were less likely to die. We found a statistically significant interaction between depression status and practice intervention assignment. Specifically, older adults in the intervention group who met standard criteria for major depression were less likely to die over the 5-year follow-up than were older adults who met criteria for major depression in the usual care group.

The reduction in death seemed to be almost entirely attributable to a reduction in deaths due to cancer. The mechanism for this effect is not apparent. Few investigators (we are aware of none from an intervention trial) have reported data pertaining to cause of death in association with depression assessed prospectively (see review [38]). By linking prospective New Haven Epidemiologic Catchment Area study interview data with the Connecticut Tumor Registry, Desai and colleagues (39) found that women with major depression were at increased risk for late-stage breast cancer diagnoses compared with women without major depression. Although this supports the plausibility that deaths due to cancer, for example, might be diminished if major depression were managed more appropriately, the potential for misclassification in cause of death derived from death certificates may be substantial and evidence of a potential association of practice intervention assignment and specific causes of death must be viewed as an opportunity for generation of hypotheses to be tested in future intervention research.

The PROSPECT intervention reduced hopelessness and depression (14, 40). Depression and hopelessness have been studied as independent predictors of increased deaths in observational studies, which have also drawn attention to the link among depression, cardiovascular disease, and death (1, 2). Intervention trials, such as the Enhancing Recovery in Coronary Heart Disease Patients (ENRICHD) trial (9, 10) and the Sertraline Anti-Depressant Heart Attack Randomized Trial (SADHART) (11), have focused on depression treatment for persons with advanced cardiovascular disease (7). The ENRICHD trial did not show any benefit on risk for death over a 40-month follow-up, and researchers did not report separate analyses for patients with major depression. Secondary analysis from ENRICHD showed that an antidepressant prescription (a selective serotonin reuptake inhibitor) was associated with reduction in all-cause mortality (12). The SADHART reported a statistically nonsignificant beneficial trend after 24 weeks for a combined outcome of death and nonfatal cardiovascular events for patients with a major depression disorder after myocardial infarction (11). Perhaps the target population represented by the ENRICHD and SADHART samples were further along in the pathway between depression and death than were participants in PROSPECT, whose sample was derived from community practices. The focus of depression treatment in high-risk patients may need to be counterbalanced by attention to primary health care and other community settings to achieve maximum reductions in depression-related deaths (40).

We cannot rule out the possibility that the mortality rate reduction we observed among patients with depression in intervention practices may be due to factors other than the specific effects of a depression management program, such as the nonspecific effects of having an additional person in the practice. We found no evidence for a substantial effect of the intervention on survival for patients with clinically significant minor depression or among patients who did not meet criteria for depression. Rather, the effect of this practice-level intervention on patient-level outcomes depended on the baseline depression status of the patient, that is, the effect of the PROSPECT intervention on death was modified by major depression status. Unützer and colleagues (41) found that an intervention designed to test the cost-effectiveness of a targeted package of preventive services did not attenuate the risk for death associated with high depression symptoms. The beneficial effect of placing a depression care manager in practices seems to be specific to depression care and is not because of the general influence of an additional clinician in a practice on the medical care of other conditions that might affect death. Nevertheless, the specific mediators between practice intervention assignment and patient outcomes deserve further study, including patient adherence, patterns of long-term care for mental health and physical conditions, and improvement in depression.

Although our findings deserve attention, we recognize that our methods have limitations. Misclassification of depression status can result in misleading inference. Depression and other mental health problems may be underestimated in the elderly because stigma leads many elderly persons to minimize reports of sadness or anhedonia and to attribute other symptoms of depression to physical health causes (42, 43). On the other hand, the prevalence of psy-

### Table 4. Hazard Ratios for Intervention Effects (Intention-to-Treat Analyses) Stratified by Depression Status Assessed at Baseline

<table>
<thead>
<tr>
<th>Baseline Depression Status</th>
<th>Hazard Ratio for Intervention Effects (95% CI)†</th>
<th>Unadjusted</th>
<th>Adjusted†</th>
</tr>
</thead>
<tbody>
<tr>
<td>All patients with depression</td>
<td>0.89 (0.59–1.34)</td>
<td>0.67 (0.44–1.00)</td>
<td></td>
</tr>
<tr>
<td>Major depression disorder</td>
<td>0.83 (0.53–1.29)</td>
<td>0.55 (0.36–0.84)</td>
<td></td>
</tr>
<tr>
<td>Clinically significant minor depression</td>
<td>1.03 (0.51–2.07)</td>
<td>0.97 (0.49–1.92)</td>
<td></td>
</tr>
<tr>
<td>Patients without depression</td>
<td>1.10 (0.81–1.50)</td>
<td>1.14 (0.84–1.53)</td>
<td></td>
</tr>
</tbody>
</table>

* Data from the Prevention of Suicide in Primary Care Elderly: Collaborative Trial (1999 to 2004) (14).
† Intervention versus usual care.
‡ Includes terms for baseline age, sex, education, smoking, cardiovascular disease, stroke, diabetes, cancer, cognition, and suicidal ideation.
chopathology can be inflated by misattributing symptoms of medical illness, medication side effects, or treatment sequelae to depression. An advantage of PROSPECT was the use of sensitive instruments (that is, the SCID-I and HDRS, which are accepted semistructured clinical interviews) by trained research associates in conducting a thorough evaluation of depression diagnosis and severity. Research associates were expected to respond to visual cues (tearfulness, appearance, and demeanor), to challenge inconsistencies (patients who report no activities yet deny anhedonia), and to request clarification (for example, “Is sleep disturbance due to frequent awakening to use the bathroom or ruminating that disturbs the ability to fall asleep after using the bathroom?”). This approach contrasts sharply with other studies that used survey instruments, retrospective recall, or nonstandardized clinical assessments.

Misclassification of vital status was also a potential limitation of our study findings. However, our experience and that of others suggest a minimal likelihood of misclassification given our methods. For example, in the 9-year follow-up of the New Haven Epidemiologic Catchment Area study (4), 1194 deaths were identified among a cohort of 3560 persons through a variety of sources, including the NDI. Survivors were confirmed through self-report, relatives, and medical records. National Death Index data confirmed 1187 of these deaths (99.4% correct classification of decedents) and correctly classified all survivors (that is, there were no false-negative results). The investigators obtained this high rate of coverage before implementation of the current computerized reporting system (NDI Plus) that includes cause of death and, unlike our study, did not benefit from the use of Social Security numbers in searching the NDI. Similarly, the overall sensitivity of the NDI for ascertainment of vital status was 98% in the Nurses’ Health Study (44) and has been much greater than 90% in most studies (see review [34]).

We believe that our investigation has implications for public health, primary care organization, and mental health services research methods. Mental health services research differs in design, implementation, and inference from other types of intervention research. The design of PROSPECT and our analytic strategy reflect the contrasts among an observational study of treatment received, a randomized trial with strictly prescribed treatment, and an effectiveness trial like PROSPECT that has features of both. In PROSPECT, adequacy of treatment of individuals can be thought of as independent from intervention assignment. Because PROSPECT was an effectiveness trial focused on how services were delivered, the intervention versus usual care assignment represented different probabilities that a patient with depression recruited into the study would receive adequate treatment. Inference from trials of practice-level interventions may require reasoning that differs from inference in observational studies or in patient-randomized clinical trials.

Depressed older adults are much more likely to present and be managed in primary care settings (23), and our study underscores the public health effect that could accrue by providing resources to help primary care clinicians bet-

---

**Figure 2.** Survival curves and 95% CIs for patients with major depression (top), minor depression (middle), or no depression (bottom) in practices randomly assigned to the intervention or usual care group.

Data from Prevention of Suicide in Primary Care Elderly: Collaborative Trial (1999 to 2004) (14).
ter manage psychological distress and psychiatric disturbances. We need to avoid the temptation to apply simple fixes because patients and providers differ widely in expectations, beliefs, preferences, and experiences about mental health (45). The backdrop of varying patient and provider perspectives needs to be studied and considered in any redesign that seeks to mitigate system-level factors that currently discourage integration of mental health treatment into primary care settings—integrated care that is preferred by patients and providers (46). If we are to prepare for the increasing need for mental health services among older persons and to ease the burden of disability associated with depression, we must engage primary care practices as partners in developing services that interrupt the pathway from depression to death.

From the University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania; University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania; and Weill Medical College of Cornell University, White Plains, New York.

Grant Support: The mortality follow-up of PROSPECT participants was funded by the National Institute of Mental Health (principal investigator, Joseph J. Gallo, MD, MPH [R01 MH065539]). The PROSPECT was a collaborative research study funded by the National Institute of Mental Health. The 3 groups included in the funded study were the Advanced Centers for Intervention and Services Research of Cornell University (coordinating center; principal investigator, George S. Alexopoulos, MD, and co-principal investigators, Martha L. Bruce, PhD, MPH, and Herbert C. Schulberg, PhD [R01 MH59366, P30 MH68638]), University of Pennsylvania (principal investigator, Ira Katz, MD, PhD, and co-principal investigators, Thomas Ten Have, PhD, and Gregory K. Brown, PhD [R01 MH59380, P30 MH52129]), and University of Pittsburgh (principal investigator, Charles F. Reynolds III, MD, and co-principal investigator, Benoit H. Mulsant, MD [R01 MH59381, P30 MH52247]). Additional small grants came from Forest Laboratories and the John D. Hartford Foundation. Participation of Drs. Gallo, Bogner, Post, and Bruce was also supported by National Institute of Mental Health awards (K24 MH070407, K23 MH67671, K23 MH01879, and K02 MH01634). Dr. Bogner is a Robert Wood Johnson Foundation Generalist Physician Faculty Scholar (2004 to 2008).

Potential Financial Conflicts of Interest: None disclosed.

Requests for Single Reprints: Joseph J. Gallo MD, MPH, Department of Family Medicine and Community Health, School of Medicine, University of Pennsylvania, 3400 Spruce Street, 2 Gates Building, Philadelphia, PA 19104; e-mail, galloj@uphs.upenn.edu.

Current author addresses and author contributions are available at www.annals.org.

References
27. Beck AT, Brown GK, Steer RA. Psychometric characteristics of the Scale for...


Current Author Addresses: Drs. Gallo and Bogner: Department of Family Medicine and Community Health, School of Medicine, University of Pennsylvania, 3400 Spruce Street, 2 Gates Building, Philadelphia, PA 19104.
Dr. Morales: Center for Clinical Epidemiology and Biostatistics, School of Medicine, University of Pennsylvania, 423 Guardian Drive, 626 Blockley Hall, Philadelphia, PA 19104.
Dr. Post: University of Michigan and Ann Arbor Veterans Affairs Healthcare System, Health Services Research & Development (11H), 2215 Fuller Road, Ann Arbor, MI 48105.
Dr. Lin: Cambridge Health Alliance, 120 Beacon Street, Somerville, MA 02143.
Dr. Bruce: Cornell University, 21 Bloomingdale Road, White Plains, NY 10605.

Author Contributions: Conception and design: J.J. Gallo, H.R. Bogner, M.L. Bruce.
Drafting of the article: J.J. Gallo, H.R. Bogner, K.H. Morales, M.L. Bruce.
Critical revision of the article for important intellectual content: J.J. Gallo, H.R. Bogner, K.H. Morales, E.P. Post, M.L. Bruce.
Final approval of the article: J.J. Gallo, H.R. Bogner, K.H. Morales, E.P. Post, M.L. Bruce.
Obtaining of funding: J.J. Gallo, E.P. Post, M.L. Bruce.
Administrative, technical, or logistic support: E.P. Post.
Collection and assembly of data: J.J. Gallo, E.P. Post, M.L. Bruce.