Multiple Informants: Mortality Associated with Psychiatric Disorders in the Stirling County Study

Nicholas J. Horton,1,2 Nan M. Laird,3 Jane M. Murphy,4–6 Richard R. Monson,6 Arthur M. Sobol,4 and Alexander H. Leighton7

This paper applies new statistical procedures for analyzing multiple-source information about the relation of psychiatric diagnoses to mortality. The data come from the Stirling County Study, a longitudinal community investigation of adults, that collected multiple-source reports (self-report and physician-report) about psychiatric disorders. These reports are used as predictors of mortality risk over a 16-year follow-up period (1952–1968). Despite extensive efforts, one or both of these reports were sometimes missing. Missingness of self-report was related to demographic characteristics as well as to physician-reports of psychiatric diagnosis. The statistical procedures used here draw together into a single frame of reference both informant reports for the initial Stirling survey and relate these to mortality risk using weighted generalized estimating equation regression models for time to event data. This unified method has two advantages over traditional approaches: 1) the relative predictiveness of each informant can be assessed and 2) all subjects contribute to the analysis. The methods are applicable to other areas of epidemiology where multiple informant reports are used. The results for self-reports and physician-reports of disorders were comparable: Psychiatric diagnosis was associated with higher mortality, particularly among younger subjects. Am J Epidemiol 2001;154:649–56.

Many psychiatric epidemiologic studies of adults in the community utilize a single source of information, typically by interviewing a randomly selected member of a household about his or her psychiatric symptoms (1–7). An exception is the Stirling County Study, a 40-year study of a general population (8–10) that has made use of information collected from two informants (or sources). One is the typical household survey, but the other is an interview with general physicians about the same persons.

Received for publication September 8, 2000, and accepted for publication March 12, 2001.
Abbreviations: DPAX, self-reported data on depression and anxiety; GEE, generalized estimating equation; GP-DATA, general physician data.
1 Department of Epidemiology and Biostatistics, Boston University School of Public Health, Boston, MA.
2 Department of Medicine, Boston University School of Medicine, Boston, MA.
3 Department of Biostatistics, Harvard School of Public Health, Boston, MA.
4 Department of Psychiatry, Massachusetts General Hospital, Boston, MA.
5 Department of Psychiatry, Harvard Medical School, Boston, MA.
6 Department of Epidemiology, Harvard School of Public Health, Boston, MA.
7 Departments of Psychiatry and of Community Health and Epidemiology, Faculty of Medicine, Dalhousie University, Halifax, Nova Scotia, Canada.
Correspondence to Dr. Nicholas Horton, Department of Epidemiology and Biostatistics, Boston University School of Public Health, 715 Albany St. T3E, Boston, MA 02118 (e-mail: horton@bu.edu).
psychiatric disorders and mortality analyzed the informants separately. One paper was concerned with the relations between self-reported disorders and mortality (26). When such disorders were aggregated using the self-report information, the standardized mortality ratio was 1.5. Another paper addressed the association of physician-reported disorders and mortality (27). When disorders identified by physicians were similarly aggregated, the standardized mortality ratio was 1.6. Both investigations indicated that psychiatric disorder was significantly associated with mortality and that subjects who died before reaching the age of 50 years were especially likely to have had a psychiatric disorder.

A practical difficulty in analyzing multiple-source reports is that there is often a substantial amount of missing data (21, 28). Multiple-informant reports are commonly missing since, by definition, they are collected from sources other than the primary subject of the study. Missingness can induce bias as well as loss of inferential efficiency (29, 30). In psychiatric epidemiology, results can be seriously biased if subjects who are missing information suffer more psychopathology than those for whom it was possible to gather data. For example, the National Comorbidity Survey (6) showed that nonrespondents have significantly more psychopathology than respondents have, which emphasizes the importance of adjustment for differential nonresponse.

In the Stirling County Study, some subjects were missing one or both of their reports of psychiatric problems. Our hypothesis is that the absence of physician-reports is not related to whether the subject had a psychiatric disorder or not, and that missing reports were due to clerical omissions. On the other hand, we hypothesize that the absence of interview (self-report) data can be related to a subject’s psychiatric status. A person with a psychiatric disorder may be more likely to refuse the interview or to be repeatedly unavailable.

Based on the availability of advanced statistical methods, this is the first report from the Stirling County Study in which the two sources are drawn together in such a way that the measured effects of each can be compared. The question of interest is to quantify the effect of psychiatric morbidity on subsequent mortality. In this paper, we apply new statistical procedures for drawing together into a single frame of reference the information from the subjects and physicians of the Stirling County Study as gathered in the first survey that was conducted in 1952, as well as information about mortality over the following 16 years. We describe unified regression models that jointly model the associations between the reports of psychiatric disorders and subsequent death. These models allow separate regressions to be analyzed together, tested for differences, and simplified if appropriate to determine a final overall regression. Partially observed multiple-source reports may be incorporated into these regression models to account for differential missingness.

MATERIALS AND METHODS

Study sample

The site of the Stirling County Study is a county of 20,000 residents in Atlantic Canada, and the fictitious name of “Stirling” is used to protect identity (31). The area has undergone many of the influences of urbanism and has experienced the major social trends exhibited throughout North America (10, 32). The study consists of repeated cross-sectional surveys, as well as cohort follow-up investigations in which subjects from earlier samples were reassessed or certified for death at each field phase when a new sample was drawn. For the 1952 survey, 1,098 subjects were selected by probability sampling. Among these, 1,003 provided a personal interview (91 percent completeness), and 1,029 were known to at least one physician who was able to provide a report (94 percent completeness). Although 953 subjects had both types of information available, 1,079 had at least one report available (50 with only a self-report and 76 with only a physician-report). These 1,079 subjects constitute the sample of interest in the paper.

Measures

The personal interview data were processed by a computer algorithm for self-reported data on depression and anxiety (DPAX), “DP” standing for depression and “AX” for anxiety (33). The algorithm has steps for determining that the syndromes of depression and/or anxiety are reasonably complete, that they persisted over a minimal duration of 1 month, and that they were associated with impairment in everyday functioning. Ten physicians contributed information for this study, each having been interviewed by a psychiatrist who asked about syndrome completeness, duration, and impairment as components of diagnosis. While the self-report interviews focused on disorders involving depression and anxiety, the physician interviews covered the full range of psychiatric disorders. The separate physicians are not distinguished in this study, and their collective information is described as general physician data (GPDATA). In this paper, our focus is on whether the subject was identified as having a psychiatric disorder rather than on what type of psychiatric diagnosis pertained. In addition to using the case assessments from the two sources, we used the demographic variables gender and age. Age in 1952 was divided into three categories: <50, 50–69, and ≥70 years, because overall mortality is associated with age.

Outcome

June 1968 was taken as the common closing date for the search of death certificates. For the large majority of subjects who died in the province where “Stirling County” is located, death certificates were found. For those who died elsewhere, information was drawn from relatives, neighbors, physicians, and family bibles (26, 27). Complete information on vital status at this time was achieved. Approximately one quarter (n = 270) of the sample died during this 16-year follow-up period.

Statistical methods

We use survival regression models as described by Laird and Olivier (34) to assess the magnitude and significance of the relation between independent variables and mortality.
These methods specify a piecewise exponential survival distribution and generate maximum likelihood estimates and associated tests. They approximate proportional hazards regression methods but have the advantage that we can incorporate multiple-informant models, using generalized estimating equations (35, 36) (GEEs) for parameter estimation.

More formally, we assume that time in the study can be partitioned into \( J \) mutually exclusive, exhaustive intervals \( \Omega_1, \ldots, \Omega_J \) with a constant hazard function within each interval. We observe \( (Y_{ij}, T_{ij}) \) for \( i = 1, \ldots, n \) and \( j = 1, \ldots, J \), where \( Y_{ij} \) denotes whether the \( i \)th subject died during the \( j \)th period, and \( T_{ij} \) denotes the time at risk for the \( i \)th subject during the \( j \)th period. Here \( Y_{ij} = 1 \) for some \( i \) and \( j \) implies that \( T_{j<k} = 0 \) for \( k = j + 1, \ldots, J \). We divided the follow-up period into \( J = 4 \) intervals, each having a duration of 4 years, and allowed the constant rate parameter to vary for each interval. The last interval (interval 4) was used as the reference group. Other predictors were assumed to be constant over time (i.e., no interaction with interval).

Laird and Olivier show that treating each \( Y_{ij} \) as a Poisson random variable: \( Y_{ij} \sim \text{Poisson}(T_{ij}a_{ij}) \), where the model for the rate is specified by
\[
\begin{pmatrix}
a_{i1} \\
a_{i2} \\
\vdots \\
 a_{ij}
\end{pmatrix} = a_i = \exp(X_i'\beta),
\]
yields the correct piecewise exponential likelihood. Here \( X_i \) is a vector of covariates for the \( i \)th subject and \( \beta \) denotes a set of regression parameters for the rate.

We denote the expected number of events:
\[
m_i = E[Y_i|T_i, X_i, \beta] = T_i a_i = T_i \exp(X_i'\beta),
\]
and we model
\[
\log(m_i|T_i, X_i, \beta) = \log(T_i) + X_i'\beta,
\]
where \( \log(T_i) \) is a known offset. Maximum likelihood estimation of the parameters of \( \log(m_i) \) is straightforward to calculate in existing Poisson regression software packages (e.g., PROC GENMOD in SAS) (37) for a single informant and fully observed predictors.

We extend the model, by considering whether use of a particular informant will yield different results for the overall regression model, using techniques proposed independently by Pepe et al. (38) and Horton et al. (17). Let \( X_1 \) and \( X_2 \) denote two multiple-informant reports where, for our example, \( X_1 \) is the GP-DATA (physician-report) and \( X_2 \) is the DPAX (self-report). We denote by \( Z \) a vector of other covariates and by \( (Y, T) \) indicators of survival status and follow-up time. We model the hazard for mortality as a function of the reports of psychiatric disorders and other covariates.

Fitting separate regression models is equivalent to specifying
\[
\log(m_i|T_i, X_{1i}, Z_i, \alpha, \beta) = \log(T_i) + \beta_0 + \beta_1X_{1i} + \beta_2Z_i
\]
and
\[
\log(m_i|T_i, X_{2i}, Z_i, \alpha, \beta) = \log(T_i) + (\alpha_0 + \beta_0) + (\alpha_1 + \beta_1)X_{2i} + (\alpha_2 + \beta_2)Z_i
\]
where other variables and interactions may be included. A limitation of this approach has been the lack of a method for deciding if regression coefficients are different in the two models and how to combine results if they are not. Pepe et al. (38) and Horton et al. (17) showed how the GEE regression model can be used with existing general purpose statistical software (39), including PROC GENMOD in SAS, to fit models that allow the same parameters to appear in both equations.

For example, it might be assumed that there is no effect of informant on the associations between the covariates \( Z_i \) and the outcome. Under this assumption, \( \alpha_2 \) would be presumed to be 0. In this case, the regression model is simplified:
\[
\log(m_i|T_i, X_{1i}, Z_i, \beta) = \log(T_i) + \beta_0 + \beta_1X_{1i} + \beta_2Z_i
\]
The \( \alpha_1 \) parameter assesses how the association between diagnosis and outcome varies by informant. If this parameter is significantly different from 0, this would imply that there is a different mortality risk associated with one of the informant reports. This might be due, in our example, to the physician-report’s being associated with a more serious type of disorder. The \( \alpha_0 \) parameter in this model denotes the difference in baseline log rate using the second informant. This may differ from 0 if the informant effects are estimated from different subsets of the data, and if missingness is related to the outcome.

While the same outcome is used in both regression equations, the distribution of \( Y \) in the two regression models is different as long as the predictors are different. Including separate parameters for models 1 and 2 implies that the use of different sources leads to different inferences about the model, while assuming that \( \alpha = 0 \) suggests that the data from different informants lead to the same inferences concerning the association of diagnosis and other predictors of mortality.

A limitation of the use of separate regression models is that, if some of the informant reports \( X_1 \) and \( X_2 \) are missing, then models 1 and 2 will be fit using different subsets of the sample. The GEE permits a subject to contribute to one equation and not to the other, but use of these “available case” methods may be inefficient, as well as biased if missingness is not completely at random (30).

Xie and Paik (40) proposed a simple approach to handling missing covariates in GEE models when the probability of missingness depends on the outcomes and/or observed covariates. This method assumes that the missingness mechanism is missing at random (29): Miss-
ingness of the unobserved covariate does not depend on the value of the unobserved covariate. In certain settings, the model reduces to a weighted GEE (41, 42).

We define the missingness indicators

\[
R_{i1} = \begin{cases} 
0 & \text{if } X_{i1} \text{ is observed} \\
1 & \text{if } X_{i1} \text{ is missing}
\end{cases}
\]

and

\[
R_{i2} = \begin{cases} 
0 & \text{if } X_{i2} \text{ is observed} \\
1 & \text{if } X_{i2} \text{ is missing}
\end{cases}
\]

Under the missing at random assumption, we have that

\[
p_1 = \Pr(R_1 = 1 \mid Y, T, X_1, X_2, Z) = \Pr(R_1 = 1 \mid Y, T, X_2, \gamma_1),
\]

and

\[
p_2 = \Pr(R_2 = 1 \mid Y, T, X_1, X_2, Z) = \Pr(R_2 = 1 \mid Y, T, X_1, Z, \gamma_2).
\]

Following the approach of Xie and Paik, we use logistic regression models, governed by \( \gamma \), for fitting models 4 and 5 and estimating \( p_1 \) and \( p_2 \). The inverses of \( \hat{p}_1 \) and \( \hat{p}_2 \) are used as weights in the GEE analysis; that is, \( \hat{p}_1^{-1} \) weights the equation with \( X_1 \) as a covariate and \( \hat{p}_2^{-1} \) for the equation with \( X_2 \). We assume a working independence correlation structure, which reduces to standard Poisson regression with estimated weights. Subjects with two informants contribute two observations; those missing one contribute one observation. We use the bootstrap (43) to estimate the standard errors of \( \alpha \) and \( \beta \). We create a data set of size 1,079 by randomly sampling persons with replacement from the original data set and refit the same model using this new sample. This procedure is repeated 1,000 times, with the estimated standard error calculated from the sample variance from these 1,000 data sets.

With time to event data, it is not possible to fit a saturated logistic regression model for models 4 and 5. Instead, we categorized the survival times into \( J \) intervals plus 1 for survivors) and treat the distribution as a multinomial. We also dropped higher order interactions from the logistic regression model. With simplification of the missingness law in these ways, there is a potential increase in bias due to misspecification of the model for weights. The overall goal is to provide a parsimonious yet flexible representation of the missingness mechanism.

A more serious complication is that missingness is not monotone: Some subjects are missing DPAX while others are missing GP-DATA. Estimation of model 4, the probability of observing \( X_1 \), is complicated because \( X_2 \) is not fully observed. Similar complications arise in the estimation of the probability of observing \( X_2 \) (\( \hat{p}_2 \)). In our example, the DPAX report (\( X_2 \)) was not predictive of missingness for GP-DATA, so we dropped \( X_2 \) from model 4 and estimated \( \hat{p}_1 \) from the full sample. This was not the case for model 5, because GP-DATA (\( X_1 \)) were predictive of missingness for DPAX (\( X_2 \)). We gave weight 1 to observations with missing predictors in the missingness model for \( \hat{p}_2 \).

### RESULTS

The majority of the sample in 1952 is female (54 percent), and the mean initial age (in 1952) was 48.5 (standard deviation, 16.2; range, 17–92) years. Table 1 displays the contingency table representing the GP-DATA and DPAX reports. Among the 1,079 subjects with data from one or both sources, the physicians reported a prevalence rate of psychiatric disorders of 14 percent, while the prevalence using self-report was 12 percent. The agreement between GP-DATA and DPAX was low (44) (\( \kappa = 0.18 \)).

We note that 12 percent (6 of 50) of the subjects with missing GP-DATA reports had positive DPAX reports and that 24 percent (18 of 76) of the subjects with missing DPAX reports had positive GP-DATA reports.

We tested whether age, gender, DPAX report, or mortality during the follow-up period was predictive of observing the physician-report for those subjects with a complete DPAX report. None of these factors was significantly associated with observing the physician-report using a likelihood ratio test: \( \chi^2 = 4.16, p = 0.53 \). All \( p \) values for main effects were \( p > 0.50 \). These results suggest that missing physician-reports are not inconsistent with random processes and may be missing completely at random.

We also tested whether the probability of observing the DPAX diagnosis was associated with age, gender, their interaction, positive GP-DATA report, or mortality during the follow-up period using subjects with a complete GP-DATA report. We rejected the null hypothesis that the missingness was missing completely at random (\( \chi^2 = 19.88, p = 0.006 \)). There was a significant interaction between age and gender (\( \chi^2 = 6.90, p = 0.03 \)), as well as a significant positive association between the GP-DATA report and DPAX missingness (\( \chi^2 = 5.56, p = 0.02 \)). Mortality during the follow-up period was not significantly associated with missing the DPAX report (\( \chi^2 = 0.55, p = 0.46 \)).

Table 2 displays the estimated probabilities of missing the DPAX report, by age group, gender, and GP-DATA report for subjects that did not die during the follow-up period. Younger women have lower estimated missingness probability than younger men have, but older women have a higher missingness fraction than older men have, and

<table>
<thead>
<tr>
<th>TABLE 1. Cross-classification of GP-DATA* and DPAX* reports of psychiatric disorders (including partially observed subjects), Stirling County Study, 1952–1968</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>GP-DATA</strong></td>
</tr>
<tr>
<td><strong>Report</strong></td>
</tr>
<tr>
<td>Missing report</td>
</tr>
<tr>
<td>No disorder</td>
</tr>
<tr>
<td>Report of disorder</td>
</tr>
<tr>
<td><strong>Total</strong></td>
</tr>
</tbody>
</table>

* GP-DATA, general physician data; DPAX, self-reported data on depression and anxiety.
missingness is higher for subjects with a positive GP-DATA report. These results indicate that missingness of the DPAX report is related to factors of interest in the study: age, gender, and physician diagnosis of psychiatric disorder.

**Regression models for overall mortality**

We begin by considering models for the probability of dying during the 16-year follow-up period, given age at baseline, gender, and psychiatric disorder status. The youngest subjects and males were used as the reference groups in the regression models. We first consider separate regression models for GP-DATA and DPAX.

To account for the partially observed subjects, we weighted observations using the inverse of the probability of observing that value. We estimated these weights using a series of logistic regression models. For the probability of observing GP-DATA reports, this model included main effects for age (2 df), survival status (4 df), and gender (1 df). For the probability of observing DPAX reports, this model included main effects of age, gender survival status, and GP-DATA report, as well as interactions between gender and age and GP-DATA report and age. The goal is to model the missingness in a relatively rich fashion to avoid introducing bias; hence, some nonsignificant predictors were retained in these regressions.

Results for GP-DATA are provided in table 3, where the parameter estimates are log annual mortality rate ratios. To aid in interpreting these results, table 4 displays the estimated annual mortality rate for a younger male (<50 years) and an older male (≥70 years) with and without a psychiatric diagnosis using GP-DATA reports during the first time interval. Here the estimated log rate for a younger male subject with no diagnosis is

\[ \log(\text{rate}) = -5.70 - 0.95 = -6.65, \]

and the predicted annual mortality rate is \( \exp(-6.65) = 0.00129 \).

As expected, the force of mortality tends to increase with age. Older subjects and those with a psychiatric diagnosis have a significantly higher rate of mortality, but the association of a psychiatric diagnosis report with mortality is significantly larger for younger subjects.

Results for the DPAX separate regression (not reported here) are similar, but there is also a small effect of gender in the model: The difference in log rates for women is −0.224 (95 percent confidence interval: −0.489, 0.041). This may be due in part to the use of different subsamples used to estimate this model, since older females with positive GP-DATA reports were more likely to be missing their DPAX reports.

We next considered models using all subjects where the same outcome (indicator of survival and follow-up time) is repeated, but each subject with complete data has two observations (one with the GP-DATA report and one with the DPAX report), and those with a missing report contribute only one. We fit a GEE model using a working independence correlation structure, as displayed in table 5. This model allows the rate to vary by informant. Since the GP-DATA report is the baseline group, the intercept in the GEE model reported in table 5 (−5.70) is the same as the intercept for the GP-DATA model reported in table 3. Similarly, the estimated intercept for DPAX reports in table 5 (−5.70 + 0.22 = −5.48) is the same as the estimate from the separate DPAX model.

The estimates of \( \alpha \) (for the main effects and interactions) were close to 0, so these nonsignificant terms were dropped from the model. The results from the final GEE model are displayed in table 6. We note that the parameter estimates for this model are roughly comparable to an average of the other models using GP-DATA only or DPAX only but that the standard error estimates of the GEE model are generally smaller. Overall, the GEE model provides a parsimonious summary of the association between diagnosis and overall mortality, while controlling for age and gender. This model also incorporates subjects that are partially observed.

To assess the effects of weighting the observations to

---

**TABLE 2. Predicted percentage missing DPAX* report, by age, gender, and GP-DATA report (for subjects that survived the follow-up period), Stirling County Study, 1952–1968**

<table>
<thead>
<tr>
<th>GP-DATA</th>
<th>Female</th>
<th>Male</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;50 years</td>
<td>1.3</td>
<td>5.1</td>
</tr>
<tr>
<td>50–69 years</td>
<td>3.1</td>
<td>11.6</td>
</tr>
<tr>
<td>≥70 years</td>
<td>4.6</td>
<td>11.6</td>
</tr>
<tr>
<td>No disorder</td>
<td>2.1</td>
<td>8.0</td>
</tr>
<tr>
<td>Report of disorder</td>
<td>3.0</td>
<td>11.1</td>
</tr>
</tbody>
</table>

* DPAX, self-reported data on depression and anxiety; GP-DATA, general physician data.

**TABLE 3. Piecewise exponential survival regression results using the GP-DATA* report (parameter estimates represent log annual mortality rate ratios, n = 1,029), Stirling County Study, 1952–1968**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate (SE*)</th>
<th>95% CI*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>−5.70 (0.27)</td>
<td>−6.24, −5.16</td>
</tr>
<tr>
<td>Interval</td>
<td>1</td>
<td>−0.95 (0.18)</td>
</tr>
<tr>
<td>2</td>
<td>−0.53 (0.17)</td>
<td>−0.86, −0.20</td>
</tr>
<tr>
<td>3</td>
<td>−0.38 (0.16)</td>
<td>−0.70, −0.07</td>
</tr>
<tr>
<td>4</td>
<td>0.00</td>
<td>0.00, 0.00</td>
</tr>
<tr>
<td>Gender</td>
<td>Female</td>
<td>−0.12 (0.13)</td>
</tr>
<tr>
<td>Male</td>
<td>0.00</td>
<td>0.00, 0.00</td>
</tr>
<tr>
<td>Age (years)</td>
<td>&lt;50</td>
<td>0.00</td>
</tr>
<tr>
<td>50–69</td>
<td>2.56 (0.27)</td>
<td>2.03, 3.09</td>
</tr>
<tr>
<td>≥70</td>
<td>3.75 (0.27)</td>
<td>3.23, 4.28</td>
</tr>
<tr>
<td>Diagnosis</td>
<td>1.69 (0.39)</td>
<td>0.93, 2.46</td>
</tr>
<tr>
<td>Diagnosis × age (years)</td>
<td>&lt;50</td>
<td>0.00</td>
</tr>
<tr>
<td>50–69</td>
<td>−1.15 (0.45)</td>
<td>−2.03, −0.26</td>
</tr>
<tr>
<td>≥70</td>
<td>−1.17 (0.48)</td>
<td>−2.12, −0.22</td>
</tr>
</tbody>
</table>

* GP-DATA, general physician data; SE, standard error; CI, confidence interval.
account for the missingness, we compared the results from the weighted models with those from the unweighted models (not reported here). The maximum absolute deviation was 0.082, and the maximum relative deviation was 27.1 percent (for one of the estimates of interval). In this setting, because missingness was not related to the outcome, the regression inferences were similar for the weighted and unweighted estimates.

**DISCUSSION**

It has long been suspected that nonresponse patterns are different for those with and without psychiatric problems.

**TABLE 5. Piecewise exponential survival GEE* models with repeated outcomes and no informant effects (parameter estimates represent log annual mortality rate ratios, n = 1,079), Stirling County Study, 1952–1968**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate (SE*)</th>
<th>95% CI*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>−5.58 (0.25)</td>
<td>−6.07, −5.09</td>
</tr>
<tr>
<td>Interval</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>−0.94 (0.17)</td>
<td>−1.28, −0.60</td>
</tr>
<tr>
<td>2</td>
<td>−0.52 (0.17)</td>
<td>−0.85, −0.19</td>
</tr>
<tr>
<td>3</td>
<td>−0.38 (0.16)</td>
<td>−0.70, −0.06</td>
</tr>
<tr>
<td>4</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>−0.18 (0.13)</td>
<td>−0.43, 0.07</td>
</tr>
<tr>
<td>Male</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>Age (years)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;50</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>50–69</td>
<td>2.50 (0.24)</td>
<td>2.03, 2.98</td>
</tr>
<tr>
<td>≥70</td>
<td>3.65 (0.24)</td>
<td>3.18, 4.12</td>
</tr>
<tr>
<td>Diagnosis</td>
<td>1.55 (0.31)</td>
<td>0.95, 2.15</td>
</tr>
<tr>
<td>Diagnosis × age (years)</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>&lt;50</td>
<td>−1.03 (0.38)</td>
<td>−1.77, −0.30</td>
</tr>
<tr>
<td>50–69</td>
<td>−1.19 (0.34)</td>
<td>−1.87, −0.52</td>
</tr>
<tr>
<td>≥70</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* GEE, generalized estimating equation; SE, standard error; CI, confidence interval.

The National Comorbidity Survey provided evidence for this effect, with people who were initially missing tending to be more ill (6). If not accounted for, biased estimates of prevalence and associations may result. In the Stirling County Study, we showed that missingness was related to demographic factors that are commonly collected (such as age and gender). In addition, missingness of self-reports was predicted by physician-reports of psychiatric disorders: The people you cannot interview in a household survey are not a random subset according to what you know from physicians.

Our results were not qualitatively different when accounting for missingness, partly because missingness was not related to overall mortality. However, missingness of self-reports was predicted by physician-reports of psychiatric disorder. According to the physician’s report, the subjects who could not be interviewed were more likely to suffer a psychiatric disorder than were those who were successfully interviewed. This bias appeared despite the high completion rates achieved in the Stirling County Study. Further, it appeared in the unusual situation of having as knowledgeable an informant as a physician give psychiatric information about the randomly chosen subjects. These factors emphasize that missingness and incomplete data are issues that need to be addressed in studies of the epidemiology of psychiatric disorders.

In psychiatric epidemiology, researchers are often interested in the results of separate regression models based on different reports of psychiatric states. These models differ
from approaches whereby both informants might be included in a regression model. Interpretability of the latter models can be problematic: How should one interpret the “effect” of a positive self-report while holding the physician-report constant? Separate regression models for each informant are straightforward to fit but difficult to interpret if they provide differing results. We have proposed methods using a single model that allows for statistical comparison of differences. These methods are broadly applicable to other epidemiologic investigations that utilize multiple-source reports, such as child psychiatric epidemiology, nutritional epidemiology, and geriatric studies.

One product of this analysis is the finding that there are similar associations between mortality and either self-report or physician-report of psychiatric disorders. The separate analyses reported earlier pointed in the same direction, but it might have been assumed from them that the similarity was due to the fact that the two sources identified the same subjects (though the agreement between the sources was low). The similarity may seem surprising because the self-report information is limited to what has sometimes been thought of as “minor psychiatric morbidity” involving depression and anxiety, while the physician-report data cover the full range of psychiatric disorders (45). In recent years, however, considerable evidence has accumulated that points to the fact that disorders identified through self-report are more serious than once realized. The seriousness is indicated not only in mortality risk, as here, but also in chronicity, disability, and use of services (46–49).

ACKNOWLEDGMENTS

The authors are grateful for the support provided by National Institute of Mental Health grants R01-MH54693, R01-MH59576, and T32-MH17119.

REFERENCES
