

Symptom profiles of depression among general medical service users compared with speciality mental health service users

TONGWOO SUH AND JOSEPH J. GALLO¹

From the Department of Mental Hygiene, The Johns Hopkins University School of Hygiene and Public Health, Baltimore, Maryland, USA; and Korea Institute of Health Services Management, Seoul, Korea

ABSTRACT

Background. Since depressive disorders are now eminently treatable and early detection and treatment could bring substantial benefits, it is critical to address alternative presentations of depression in the general medical setting. Concern regarding under-diagnosis of depression in general medical settings has given rise to the question of whether the clinical disorder of depression differs qualitatively or only quantitatively across care settings.

Methods. Symptom profiles of depression were compared across care sectors to investigate how the presentation of depression among general medical service users might differ qualitatively from speciality mental health service users. Data on depression symptoms within 6 months of interview gathered in three community surveys that were part of the NIMH Epidemiologic Catchment Area Program were analysed using methods developed to assess item bias. The subjects were 4931 and 363 persons who reported a visit to the general medical sector or to speciality mental health respectively, within 6 months of interview.

Results. Compared with speciality mental health service users, general medical service users were less likely to present dysphoria (adjusted Odds Ratio, aOR = 0.57; 95% Confidence Interval, CI = 0.38–0.84) and feeling worthless, sinful, or guilty (aOR = 0.63; 95% CI = 0.40–0.98), but were more likely to present fatigue (aOR = 1.71; 95% CI = 1.09–2.69), even after holding constant other characteristics that might influence reporting of symptoms as well as level of depression.

Conclusions. These results suggest that there are qualitative differences in depression presenting in general medical care compared with speciality mental health care and call for a re-conceptualization of depression in the general medical setting.

INTRODUCTION

Many epidemiological studies have reported that depressive disorder is among the most common mental disorder in the general medical setting as well as in the speciality mental health setting (Hoepfer *et al.* 1979; Katon, 1982; Kessler *et al.* 1985, 1987). From the community samples

of the Epidemiologic Catchment Area (ECA) study, for example, the 1-year prevalence rate of major depression was about 5% (Regier *et al.* 1993) and the 6-month prevalence was 2.4 to 3.8% (Myers *et al.* 1984). The prevalence of major depression measured through structured interview instruments in the primary-care setting ranges from 4.1 to 8.6%, higher than that found in community samples (Hoepfer *et al.* 1979; Schulberg *et al.* 1985; Kessler *et al.* 1987; Barrett *et al.* 1988; Blacker & Clare, 1988; Coulehan *et al.* 1990; Ormel *et al.* 1990; Zich *et al.* 1990). The estimated prevalence of major

¹ Address for correspondence: Dr Joseph J. Gallo, Department of Mental Hygiene, Johns Hopkins University School of Hygiene and Public Health, 624 North Broadway, Baltimore, Maryland, 21205, USA.

depression in the primary-care setting varies depending on whether the morbidity is defined by clinician examination, patient self-report, or lay-administered psychiatric structured interview (Schulberg *et al.* 1985).

It has been argued that general practitioners (GPs) are in an advantageous position to detect depression because GPs are likely to see a significant proportion of those who are suffering from depression (Hankin *et al.* 1982; Burvill & Knuiman, 1983; Shapiro *et al.* 1984; Regier *et al.* 1993; Marino *et al.* 1995). However, the concept of depression has been based on clinicians' experience of patients in psychiatric clinics and diagnostic criteria for depressive disorders have been developed and modified by studies in psychiatric clinic settings (Spitzer *et al.* 1977; Mirowsky & Ross, 1989). Therefore, clinicians and investigators have been concerned about misclassification of depression among general medical service users, especially under-diagnosis (Knights & Folstein, 1977; Goldberg & Huxley, 1980; Mann *et al.* 1981; Lebo *et al.* 1988; Sanson-Fisher & Hennrikus, 1988; Pérez-Stable *et al.* 1990). Since depressive disorders are now eminently treatable and early detection and treatment could bring substantial benefits, it is critical to address alternative presentations of depression in the general medical setting.

Depression in the general medical sector

Concern regarding under-diagnosis of depression in general medical settings has given rise to the question of whether the clinical disorder of depression differs qualitatively or only quantitatively when patients in the general medical sector are compared with patients in the speciality mental health sector. There are two different views regarding the nature of the depression seen in the general medical sector. One perspective holds that the majority of depression presented in general practice falls within a single broad diagnostic category. From this viewpoint, depression seen in general medical practice is essentially the same as depression seen in psychiatric clinics by psychiatrists, but at a less severe or early stage. If this is true, patients of the general physician can be identified and classified according to the same standard criteria that have been constructed based on the experience of psychiatrists, such as DSM-III (American Psychiatric Association, 1980). Some

American studies have supported this viewpoint. Limiting analysis to respondents who met the DSM-III criteria for major depression, Cooper-Patrick *et al.* (1994) compared symptom profiles (e.g. suicidal ideation/attempts, vegetative/non-vegetative symptoms, somatization syndrome) between depressed patients in the general medical and the speciality mental health sector. Similar symptom profiles across care sectors were found, although there were differences in some demographic factors and in psychiatric comorbidity. Coulehan *et al.* (1988) could not find substantial differences in the symptom profiles of patients presenting with a major depressive episode between primary-care clinics and community mental health centres.

Another perspective on differences in depression in the general medical setting takes the stance that depression is not the same as in the psychiatric clinic setting, in terms of the symptom profiles, natural course, prognosis, and the response to pharmacological treatment. Most British studies and some American studies on depression support this perspective. Investigators have found that depressive patients seen in the primary-care sector have briefer episodes, express changes in mood less frequently, and more often present in somatic terms than depressed patients in the psychiatric sector (Goldberg & Blackwell, 1970; Fahy, 1974; Pilowsky & Spence, 1978; Sireling *et al.* 1985; Emmons *et al.* 1987). Williamson & Yates (1989) reported that depressed medical patients' chief complaints were more somatic, obscure and less psychologically focused when compared with psychiatric patients. Once present at a diagnosable stage, however, the symptom profiles seemed to be very similar in both patient groups. This finding is consistent with previous findings by Coulehan *et al.* (1988) and Cooper-Patrick *et al.* (1994), studies which limited analysis to patients meeting standard diagnostic criteria, thus excluding patients with subthreshold depression and alternative presentations.

A dimensional specification of depression

Classification of depression can be based on either categorical or dimensional models in terms of summarizing individuals' depression symptoms (Goldberg & Huxley, 1992; Gallo, 1995). Goldberg & Huxley (1992) recognized the value of a dimensional specification of psychopath-

ology in understanding patient presentations in primary care. As Newmann (1989) describes, the dimensional model is consistent with a psychometric tradition, which uses symptom scales and assumes that depression varies on a continuum, while the categorical model is consistent with the clinical tradition, which emphasizes caseness and assumes that depression is a disease that is either 'present' or 'absent'. The categorical model for common mental disturbances seen in general medical settings, such as depression, has been criticized for a number of reasons. For example, depression symptoms that would not meet full standard criteria for major depression ('subthreshold depression') are nevertheless associated with poor functional outcomes (Crum *et al.* 1994; Horwath *et al.* 1992). Furthermore, subthreshold depression is associated with other psychiatric disturbances as well as increased utilization of health services, even in the absence of full diagnostic criteria (Olfson & Klerman, 1992; Wells *et al.* 1992; Spitzer *et al.* 1995). Regier *et al.* (1993) reported that 45% of persons who received care in the speciality mental health sector did not meet any diagnostic criteria for a mental disturbance. This evidence suggests that using a dimensional model that encompasses the whole range of depression might be advantageous in comparing depression across care settings.

To compare symptom profiles across care settings, we employed logistic regression models to implement item response theory in the framework of the symptom criteria of major depression in DSM-III. Item response theory was developed in the educational and psychometric field in the 1960s to investigate the assertion that the principal reason for the great disparity in test performance between minority and White students on tests of cognitive ability was that the tests contained items that were biased against persons from minority cultures (Angoff, 1993). 'Item bias' is defined as a systemic error in the measurement process such that equally able (or proficient) individuals from different groups do not have equal probabilities of responding to an item correctly (Osterlind, 1984; Angoff, 1993). A test item is said to be 'unbiased' when the probability of responding correctly on the item is the same for equally able examinees regardless of their membership in a particular subgroup (Osterlind, 1984).

In relation to the criteria for major depression, the concept of item bias can be used to assess whether persons in the general medical sector are as likely to report a given symptom compared to persons who report a visit in the speciality mental health sector. Since the criteria on which the *Diagnostic and Statistical Manual* are based have been drawn from clinicians' experience of patients in psychiatric clinics (Spitzer *et al.* 1977; Mirowsky & Ross, 1989), the pattern of symptoms among psychiatric patients is a reference or an expected stereotypical form of depression for other patients (namely, general medical service users). If a certain depression symptom functions significantly differently for general medical service users compared with speciality mental health service users, the symptom can be regarded as 'biased'. Given the centrality of the dysphoria symptom for the diagnosis of major depression, item functioning of the dysphoria symptom has a critical impact on the estimate of the prevalence of depression across different subgroups. Homogeneous (or unbiased) functioning of the dysphoria symptom implies that people in the general medical setting would present with dysphoria to the same extent as speciality mental health service users at the same level of depression. If the dysphoria item does not function homogeneously among general medical service users when compared with speciality mental health service users, the implication is that general medical service users are not categorized into 'case or normal' in the same way as speciality mental health service users. That is, if we find a significant underpresentation pattern for dysphoria among general medical service users, the prevalence of major depression in general medical settings might be underestimated.

Hypothesis

This study investigates depression in the general medical setting by comparing symptom profiles of depression in the general medical setting to the speciality mental health setting. Through comparison of symptom profiles, this study tests the hypothesis that, even adjusting for level of depression, compared to speciality mental health service users, general medical service users will be less likely to present dysphoria and other psychological symptoms (e.g. feeling worthless/sinful/guilty), but will be more likely to present

somatic (vegetative) symptoms (e.g. fatigue). In making these comparisons, we adjust for potentially influential characteristics, such as gender, that may be associated with reporting of depression symptoms.

METHOD

The Epidemiologic Catchment Area (ECA) Program

The ECA program was a series of epidemiological surveys conducted by collaborators between 1980 and 1984 at five sites in the United States: New Haven, Connecticut (Yale University), Baltimore, Maryland (Johns Hopkins University), St. Louis, Missouri (Washington University), Durham-Piedmont, North Carolina (Duke University), and Los Angeles, California (University of California, Los Angeles). ECA data include both community and institutional populations interviewed in person. At each site, ECA collaborators used multi-stage probability sampling to select 3000 to 5000 adult respondents aged 18 years and older (nearly 20000 adults). Other details of the ECA program, including sampling procedures, instrument design, and several studies of Diagnostic Interview Schedule reliability and validity, have been described elsewhere (Anthony *et al.* 1985; Eaton & Kessler, 1985; Helzer *et al.* 1985; Robins & Regier, 1991).

Study sample

Subjects in this study are household respondents who utilized medical services – general medical and/or speciality mental health services during the 6 months prior to the interview. Persons did not have to meet standard criteria for major or minor depression to be included in the study. Recency information, which establishes the time frame in which a given symptom last occurred, plays a central role in this research. Because data on the recency of individual depression symptoms were collected at only three ECA sites (Baltimore, Durham, and Los Angeles), the current analysis is restricted to respondents from three sites. The response rate for screening the households designated was over 90% at all the sites, and the effective response rate, combining nonresponse at both the household screener and respondent level, resulted in the following completion rates: Baltimore, 78%;

Durham, 79%; and Los Angeles, 68% (Von Korff *et al.* 1985; Kessler *et al.* 1987).

Measurement strategy

All variables under study were measured by standardized and generally pre-coded questions, as part of a highly structured interview administered by an agency lay-interviewer with Diagnostic Interview Schedule (DIS) training. Each interview was usually conducted in a private place, typically inside the respondent's home. The questions on medical service use preceded the DIS, and the DIS was designed to limit potential biases introduced by early disclosure of recent medical service use.

Utilization of medical services

The ECA interview schedule includes a range of questions on the utilization of ambulatory care during the 6-month period prior to the interview. The core question about general medical service use was 'Not counting any care you may have received while you were a bed patient in a hospital or nursing home, how many times altogether did you receive care or treatment from a health professional in an office, clinic or emergency room in the past 6 months, that is, since [date of 6 months ago]?' The core question about speciality mental health services was 'Did you go to a psychiatrist, psychologist, social worker or counsellor for help with problems with your emotions, nerves, drugs, alcohol, or your mental health in the past 6 months, that is, since [date of 6 months ago]?'

The general medical and speciality mental health service users are defined for this study as mutually exclusive categories: 'the general medical service users' defined in this study generally represent those who reported visits to non-psychiatric clinicians, and 'the speciality mental health service users' represent those who reported visits to psychiatric clinicians in the 6 months prior to interview, whether or not a visit was made to the general medical setting.

Depression criteria

The DIS contains 17 questions related to the depressive syndrome. Positive responses to the questions were followed by further questioning to determine whether a threshold for severity had been met and whether the symptom was plausible as a psychiatric symptom; that is, the

symptom could not be explained by physical illness, medications, alcohol, or drug use (Robins *et al.* 1981; Robins & Regier, 1991). The items were grouped into the dysphoria symptom and eight symptom clusters referable to the DSM-III Criterion B for Major Depressive Episode (Von Korff & Anthony 1982). Thus, item parcels were created from the symptom-level data (West *et al.* 1995). Depression symptoms corresponding to the DSM-III criteria are defined as present if any of the constituent DIS items are reported as present during the preceding 6 months. Specifically, item parcels were created for the symptom criteria; appetite disturbance (3 items), sleep disturbance (2 items), psychomotor symptoms (2 items), thought disturbance (2 items), and suicidal ideation/behaviour (4 items).

Analytical strategy

Item response theory assumes that test items are intended to measure only a single attribute, which means that test items should have a single definable dimension or a single latent trait (Osterlind, 1984; Camilli & Shepard, 1993). For this reason, our initial step was to test the unidimensionality of symptoms in DSM-III criteria for major depression employing exploratory factor analysis. Depression symptoms in DSM-III criteria are dichotomous measures: a respondent either has a symptom or not. Correlation of variables that are measured dichotomously cannot be estimated with Pearson product-moment correlations. Therefore, the standard factor analysis technique, based on Pearson product-moment correlations, was not used (Muthén, 1989, 1995). Instead, dichotomous variable factor analysis based on tetrachoric correlations, as provided in LISCOMP (Linear Structural Equations with a Comprehensive Measurements; Muthén, 1988), was used in this study. If we find a single factor, a single dimension explains the variance in depression symptoms of the DSM-III criteria. The logistic regression model was then employed to detect item bias (Swaminathan & Rogers, 1990; Anthony & Aboraya, 1992). In the logistic regression model, each depressive symptom (a dichotomous variable) is the dependent variable. The probability function of each depressive symptom accounting for level of depression (measured by the number of depressive symptoms) was fit into a logistic curve to estimate a

'threshold' and an 'estimated probability ratio'. The 'threshold' is defined as the number of depression symptoms present when half of the respondents report the particular depressive symptom that is the dependent variable. The 'estimated probability ratio' is measured by the odds ratio for reporting each depressive symptom among general medical service users compared to speciality mental health service users, holding constant the level of depression. Multiple logistic regression models were then used to detect a differential presentation pattern for each depression symptom across care settings, and after holding constant other covariates, such as age, and the level of depression. Finally, to investigate if the item bias across care settings varied by other sociodemographic characteristics ('non-uniform item bias'), interaction terms were included in the multiple logistic regression models.

RESULTS

Study sample

Among 5880 medical service users (402 speciality mental health service users and 5478 general medical service users) 576 subjects (9.8%) were excluded because of missing data, so that responses from 5294 subjects (363 speciality mental health service users and 4931 general medical service users) were available for analysis. Sociodemographic characteristics by care setting are shown in Table 1. Except for gender, there were significant differences between the two groups according to the sociodemographic factors ($P < 0.001$). The subjects who were interviewed in Los Angeles, aged less than 65 years, White, educated 12 years or more, married/living with spouse, employed, had other psychiatric disorders, or lived alone reported higher utilization of speciality mental health services. Table 2 shows the 6-month prevalence of major depression symptoms in the DIS/DSM-III criteria according to care setting as well as the constituents of the item parcels forming the criteria groups. The results in Table 2 demonstrate that speciality mental health service users were more likely to report all the depression symptoms. This tendency was accounted for in subsequent analyses by including a term for the level of depression in multiple logistic regression models.

Table 1. Comparison of characteristics of the study sample across care settings. Data from Baltimore, Durham and Los Angeles ECA study samples (1981–1984)*

Characteristics	General medical service users (<i>N</i> = 4931) No. (%)	Speciality mental health service users (<i>N</i> = 363) No. (%)	Total medical service users (<i>N</i> = 5294) No. (%)
Site			
Baltimore	1792 (36.3)	111 (30.6)	1903 (35.9)
Durham	1895 (38.4)	97 (26.7)	1992 (37.6)
Los Angeles	1244 (25.2)	155 (42.7)	1399 (26.4)
Age (years)			
18–64	3609 (73.2)	348 (95.9)	3957 (74.7)
≥ 65	1322 (26.8)	15 (4.1)	1337 (25.3)
Gender			
Male	1722 (35.9)	129 (35.5)	1901 (35.9)
Female	3159 (64.1)	234 (64.5)	3393 (64.1)
Race			
White	3012 (61.1)	247 (68.0)	3259 (61.6)
Non-white	1919 (38.9)	116 (32.0)	2035 (38.4)
Education (years)			
12 or more	2679 (54.3)	257 (70.8)	2936 (55.5)
< 12	2252 (45.7)	106 (29.2)	2358 (44.5)
Marital status			
Married/living with spouse	2453 (49.7)	230 (63.4)	2611 (49.3)
Living without spouse	2478 (50.3)	133 (36.6)	2683 (50.7)
Employment status			
Employed	2559 (51.9)	221 (60.9)	2780 (52.5)
Unemployed	2372 (48.1)	142 (39.1)	2514 (47.5)
Other psychiatric disorders†			
Absent	3930 (79.7)	211 (58.1)	4141 (78.2)
Present	1001 (20.3)	152 (41.9)	1153 (21.8)
Number in household			
1	1205 (24.4)	123 (33.9)	1328 (25.1)
2	1712 (34.7)	100 (27.5)	1812 (34.2)
3	829 (16.8)	53 (14.6)	882 (16.7)
≥ 4	1185 (24.0)	87 (24.0)	1272 (24.0)

* All the differences across care settings are statistically significant ($P < 0.001$), except for gender ($P = 0.879$).

† Other psychiatric disorders include schizophrenic disorder, mania, panic disorder, obsessive-compulsive disorder, phobia, somatization disorder, anorexia, substance use disorder, cognitive impairment.

Factor analysis

Latent roots for the sample covariance matrix of the nine symptoms of major depression in DSM-III criteria were obtained from exploratory factor analysis in LISCOMP. This statistical procedure is equivalent to principal components analysis for dichotomous data. The scree plot of latent roots from the sample covariance matrix was consistent with a single dimension constituted by nine depression criteria (Kim & Mueller, 1978; Gibbons *et al.* 1985). The first eigenvalue accounted for 61 % of the variance in the responses.

Item characteristic analysis

Table 3 describes the item characteristics of nine depression symptoms in terms of threshold and probability ratio for each symptom presented

among general medical service users compared with speciality mental health setting estimated from nine logistic regression models, adjusted for level of depression. Since the logistic regression model cannot identify the distribution of threshold statistics, statistical testing of difference in thresholds across care settings was not possible. However, compared with speciality mental health service users, thresholds for dysphoria and worthless/sinful/guilty among general medical service users seem much higher, while thresholds for lost sexual interest and fatigue seem much lower. The difference of thresholds across care settings was consistent with the estimated probability ratio for reporting those symptoms measured by odds ratios in the logistic regression model; compared with speciality mental health service users, general medical service users were less likely to report dysphoria

Table 2. Six-month prevalence of major depression symptoms by care setting. Data from Baltimore, Durham and Los Angeles ECA study samples (1981–1984)

	Speciality mental health (<i>N</i> = 363) %	General medical health (<i>N</i> = 4931) %
Diagnostic Interview Schedule Items based on DSM-III Criteria		
Dysphoria	25.7*	6.9*
72 Have you ever had two weeks or more during which you felt sad, blue, depressed, or when you lost all interest and pleasure in things that you usually cared about or enjoyed?		
Appetite disturbance	21.5	9.2
74 Has there ever been a period of two weeks or longer when you lost your appetite?	9.9	3.0
75 Have you ever lost weight without trying to – as much as two pounds a week for several weeks (or as much as 10 pounds altogether)?	6.1	2.4
76 Have you ever had a period when your eating increased so much that you gained as much as two pounds a week for several weeks (or 10 pounds altogether)?	11.8	5.6
Sleep disturbance	30.9	14.4
77 Have you ever had a period of two weeks when you had trouble falling asleep, staying asleep, or waking up too early?	26.2	12.0
78 Have you ever had a period of two weeks or longer when you were sleeping too much?	9.4	3.4
Fatigue	19.6	9.0
79 Has there ever been a period lasting two weeks or more when you felt tired out all the time?		
Psychomotor symptoms	17.1	5.7
80 Has there ever been a period of two weeks or more when you talked or moved more slowly than is normal for you?	10.2	3.1
81 Has there ever been a period of two weeks or more when you had to be moving all the time – that is, you couldn't sit still and paced up and down?	10.5	3.4
Lost sexual interest	7.2	3.0
82 Was there ever a period of several weeks when your interest in sex was a lot less than usual?		
Feeling of worthless/sinful/guilty	16.5	3.8
83 Has there ever been a period of two weeks or more when you felt worthless, sinful, or guilty?		
Thought disturbance	19.8	6.5
84 Has there ever been a period of two weeks or more when you had a lot more trouble concentrating than is normal for you?	15.7	4.6
85 Have you ever had a period of two weeks or more when your thoughts came much slower than usual or seemed mixed up?	13.2	3.7
Suicidal ideation/behaviour	25.3	10.6
86 Has there ever been a period of two weeks or more when you thought a lot about death – either your own, someone else's, or death in general?	20.1	9.5
87 Has there ever been a period of two weeks or more when you felt like you wanted to die?	9.4	2.1
88 Have you ever felt so low you thought of committing suicide?	9.6	1.6
89 Have you ever attempted suicide?	1.4	0.1
At least one depression symptom	59.8	33.0
Number of reported depression symptoms		
Mean	1.83	0.69
(95% Confidence Interval)	(1.60–2.07)	(0.65–0.73)

* Numbers shown in italics represent the prevalence of symptom reporting for any of the symptoms in that symptom group.

(OR = 0.49; 95% CI = 0.33–0.72) and worthless/sinful/guilty (OR = 0.55; 95% CI = 0.35–0.86) after holding constant the level of depression, but were more likely to report fatigue (OR = 1.82; 95% CI = 1.17–1.83).

Multivariate analysis

Since this item bias for depression symptoms (differential reporting pattern among general medical service users compared with speciality mental health service users) might be con-

founded by differences in characteristics across care settings as shown in Table 1, terms were introduced into the logistic regression models that represent other factors that might be influential. Table 4 shows estimated differential item functioning across care setting for reporting each depression symptom from nine multiple logistic regression models that include terms for possible confounders. As hypothesized, general medical service users were less likely to report (or present) dysphoria compared to speciality

Table 3. *Item characteristics of depression symptoms of major depressive disorder in DSM-III Criteria across care settings. Data from Baltimore, Durham and Los Angeles ECA study samples (1981–1984)*

Depression symptom	Threshold*	OR (95% CI, <i>P</i> value)†
Dysphoria		
Speciality mental health service users	3.29	Reference
General medical service users	3.99	0.49 (0.33–0.72; < 0.001)
Appetite disturbance		
Speciality mental health service users	4.07	Reference
General medical service users	3.92	1.14 (0.78–1.68; 0.500)
Sleep disturbance		
Speciality mental health service users	2.70	Reference
General medical service users	2.58	1.19 (0.81–1.73; 0.373)
Psychomotor symptom		
Speciality mental health service users	4.80	Reference
General medical service users	4.59	1.18 (0.74–1.87; 0.485)
Lost sexual interest		
Speciality mental health service users	8.39	Reference
General medical service users	6.71	1.54 (0.89–2.65; 0.121)
Fatigue		
Speciality mental health service users	4.30	Reference
General medical service users	3.72	1.82 (1.17–2.83; 0.008)
Worthless/sinful/guilty		
Speciality mental health service users	5.19	Reference
General medical service users	5.61	0.55 (0.35–0.86; 0.008)
Thought disturbance		
Speciality mental health service users	4.29	Reference
General medical service users	4.27	1.00 (0.64–1.56; 0.997)
Suicidal ideation/behaviour		
Speciality mental health service users	3.49	Reference
General medical service users	3.48	1.00 (0.69–1.45; 0.980)

* Number of depression symptoms when 50% of respondents reported each depression symptom.

† Estimated probability ratio (95% confidence interval, *P* value) for reporting each depression symptom among general medical service users compared to speciality mental health service users, holding constant the level of depression.

mental health service users, even after holding constant age, sex, self-reported ethnicity, education, marital status, employment status, other psychiatric disorders, study site, number in household as well as level of depression. The estimated probability ratio for reporting dysphoria across care settings was 0.57 (95% CI = 0.38–0.84). This means general medical service users were 43% less likely to report dysphoria compared with speciality mental health service users after holding constant other covariates as well as level of depression. In addition to dysphoria, general medical service users were less likely to report worthless/sinful/guilty (aOR = 0.63; 95% CI = 0.40–0.98) compared with speciality mental health service users, but were more likely to report fatigue (aOR = 1.71; 95% CI = 1.09–2.69). Lost sexual interest was marginally insignificant (aOR = 1.65; 95% CI = 0.94–2.90) according to traditional criteria of

statistical significance. The estimated probability ratios in adjusted multiple logistic regression models did not change markedly from unadjusted estimates in Table 3.

To investigate if the item bias across care settings varied by other sociodemographic characteristics, we introduced interaction terms. Only one significant interaction (indicating non-uniform bias) was found; namely, item functioning for dysphoria varied by the level of education across care settings (aOR = 0.43, 95% CI = 0.27–0.68, for subjects with 12 or more years of schooling compared to aOR = 1.04, 95% CI = 0.51–2.12, for subjects with less than 12 years of schooling). The interpretation is that, compared with speciality mental health service users with higher education, general medical service users with higher education were less likely to report dysphoria, while persons with lower education were no less likely to report dysphoria in the

Table 4. Summary of estimated item bias across care settings for depression symptoms of major depression in DSM-III Criteria in multiple logistic regression models. Data from Baltimore, Durham and Los Angeles ECA study samples (1981–1984)

Depression symptom	OR*	95% CI*	P
Dysphoria	0.57	0.38–0.84	0.005
Education			
≥ 12 years	0.43	0.27–0.68	< 0.001
< 12 years	1.04	0.51–2.12	0.913
Appetite disturbance	1.15	0.77–1.70	0.491
Sleep disturbance	1.21	0.82–1.78	0.330
Psychomotor symptoms	0.90	0.56–1.47	0.686
Lost sexual interest	1.65	0.94–2.90	0.080
Fatigue	1.71	1.09–2.69	0.021
Worthless/sinful/guilty	0.63	0.40–0.98	0.042
Thought disturbance	1.02	0.64–1.60	0.947
Suicidal ideation/behaviour	0.96	0.66–1.41	0.849

* Estimated probability ratio and 95% confidence interval for reporting each depression symptom of major depressive disorder in DSM-III criteria among general medical service users compared to speciality mental health service users, holding constant age, gender, self-reported ethnicity, education, marital status, employment status, other psychiatric disorders, number in household, study site as well as the level of depression.

general medical setting compared with the speciality mental health setting.

DISCUSSION

This study employed methods borrowed from testing theory to examine differences in symptom patterns. An under-presentation pattern of psychological symptoms (dysphoria and worthless/sinful/guilty) and over-presentation pattern of somatic symptoms (fatigue and lost sexual interest) was found among general medical service users compared with speciality mental health service users. The finding of significant item bias across care settings was consistent with previous research that found a somatization pattern among primary care or general medical service users (e.g. Katon *et al.* 1986). Therefore, our findings support the view that depression in the general medical setting is not the same as that in the psychiatric clinic setting. At first glance, our findings with regard to symptom presentation pattern seem inconsistent with two studies conducted in the United States that found no significant difference of symptom profiles of depression across care settings

(Coulehan *et al.* 1988; Cooper-Patrick *et al.* 1994). These studies, however, examined the difference among patients already diagnosed as a case of major depressive disorder. For this reason, it is not unexpected that the investigators found no difference in symptom profiles across care settings. In contrast, we used a study design that permitted inclusion of the full range of symptoms.

Somewhat unexpectedly, there was no significant item bias across care settings for reporting some other psychological symptoms, like thought disturbance and suicidal ideation/behaviour, or for some somatic symptoms, like appetite disturbance and sleep disturbance. If we look at this study's findings from a different perspective, these negative findings make sense. Suppose item bias across care settings for depression symptoms is related to the location of each symptom on a 'psycho-somatic' spectrum. Dysphoria and worthless/sinful/guilty are on the psychological side, while fatigue and lost sexual interest are on the somatic side of spectrum. In the same way, thought disturbance, suicidal ideation/behaviour, appetite disturbance, and sleep disturbance could be regarded as symptoms with equal psychological and somatic components (along the middle of the psycho-somatic spectrum), resulting in no significant differential reporting across care settings.

The differential reporting pattern of depression symptoms across care settings can be understood from two perspectives. First, during the utilization of medical services, patients might have been 'trained' to somatize their depression in the general medical setting and 'trained' to psychologize their depression in the speciality mental health setting by physicians and by the clinic *milieu* (an 'iatrogenic effect'). The second perspective concerns help-seeking behaviour related to the filter model of health care services (Goldberg & Huxley, 1980, 1992). Patients who had a 'tendency to somatize' their depression were more likely to keep utilizing the general medical sector, while those who had a 'tendency to psychologize' their depression were more likely to pass the filter from the general medical sector to the speciality mental health sector. Data presented in this study are cross-sectional and cannot disentangle the extent to which the differential reporting across care settings reflects these two perspectives.

Our study found non-uniform bias in dysphoria across care settings according to education level. Differential reporting for dysphoria was present only among subjects with higher education (12 years or more of schooling). In other words, compared with speciality mental health service users with higher education, general medical service users with higher education were less likely to report dysphoria, while those with lower education were no more likely to report dysphoria in the speciality mental health setting. Non-uniform item bias perhaps can be explained by considering what effect education level might have on presentation of depression. First, education level can be regarded as representing the ability to comprehend the surrounding situation and to adapt to situational needs. Subjects with higher education reported dysphoria that they considered appropriate to the care setting they utilized. Persons with higher education could adapt themselves into the situation by taking 'culturally acceptable roles' better than persons with lower education who could not differentiate the situational needs in terms of presentation pattern of depression according to care settings. In other words, those with higher education obtained more 'iatrogenic effects' than those with lower education. The second perspective is related to help-seeking behaviour. Education level represents the difference in knowledge about the medical sector and ability to match sector to symptoms. In other words, those with higher education could appropriately select the sector (general medical or speciality mental health sector) according to their types of symptoms, while those with lower education were less adept at matching their symptoms to care sector. Consequently, the reporting pattern for dysphoria was not different across care settings among subjects with lower education, but differed among those with higher education.

We note several limitations of this study. First, although the ECA had reasonably high response rates across sites (Von Korff *et al.* 1985; Kessler *et al.*, 1987), non-response could result in selection bias. Prior analyses of the data on non-response in the ECA, however, did not indicate strong biases among sociodemographic or psychopathological variables according to response status, differences in sample composition, or response rate (Von Korff *et al.* 1985).

Secondly, in this study, entire observations were deleted if a missing value was found on any of the variables in the model since logistic regression models do not allow missing values. We checked the potential for selection bias by comparing the total sample and the study sample (sample after excluding subjects with missing values). There were no differences after excluding those with missing values except for age (data not shown) and age was controlled in the multivariate models to decrease the selection bias as well as any confounding effect. Thirdly, this study used unweighted data. We were interested in the relative effect of care sectors on reporting each depression symptom, not on national estimates of number of depression symptoms in each care sector. However, to some degree, this study accounted for the complex sampling design by including the number of individuals in the household in the models. This allowed control for the most important sampling scheme, over-sampling individuals living alone. Fourthly, since this study used self-reported retrospective data, there are several limitations with regard to measurement shared by similar studies; for example, due to socially desirable responding, recall bias, and other biases in retrospective interview data (Anthony *et al.* 1985; Eaton & Kessler, 1985). Fifthly, there are limitations in statistical analysis. The multiple logistic regression model does not take into account measurement error in relation to each depression symptom and the level of depression. We simply measured level of depression from nine dichotomous depression symptoms in DIS/DSM-III criteria. Better measurement of 'true' level of depression might result from a psychiatrist's severity assessment measurement after semi-structured detailed interview, but this would have been prohibitively expensive. Finally, this study could only detect relative discrepancies of DSM-III criteria of depression across care settings, not any constant bias of the DSM-III criteria (Camilli & Shepard, 1993).

Notwithstanding the limitations, this study has a number of strengths that set the research apart from other work in this area. First, this study used a community-based sample drawn to be representative of the adult population living in households in three locations in the United States, which is expected to have less selection bias (or Berksonian's bias) than studies based on

hospital or speciality clinic samples. This large sample size increases statistical power and allows control for many possible confounders, providing an opportunity to assess interaction. Secondly, the data were gathered through a highly standardized interview instrument. Since data collection was not affected by the environment of clinical settings, we could focus on patient factors without controlling variation from physician and systemic factors. Thirdly, data on symptoms are not wasted since the dimensional model is employed. In other words, we used samples with a full range of severity of depression, and were not restricted to persons who met diagnostic criteria for major depression. Finally, unlike other studies, which compared depression across care settings, we adjusted our estimates of item bias for level of depression as well as a number of possible confounders.

The results and methodology of this study help elucidate the nature of major depression in the general medical setting. Depression in the general medical setting might have a different nature compared with the speciality mental health setting in terms of an under-presenting pattern of psychological depression symptoms and an over-presenting pattern of somatic symptoms. The results call for re-conceptualization of depression in the general medical setting; namely, reformulation of a classification system of depression for the general medical setting. The results of this study also have implications for screening guidelines for depression in primary care. For example, dysphoria might not be appropriate for a gateway (or pathognomonic) question when exploring or screening depression among patients in the general medical setting. Another implication of this study's findings might be on the treatment of depression in the general medical setting. Most data on treatment strategies originate from psychiatric clinic settings, and many of these have not been evaluated in general medical settings (Burvill, 1988). Since this study found substantial heterogeneity of symptom profiles in depression across care settings, some of the current recommendations for treatment in general medical settings might need to be modified. For example, dysphoria might not be the best indicator of change of depression during the treatment for patients in general medical settings. In conclusion, this collaborative study

found heterogeneity in symptom profiles of depression between the general medical and the speciality mental health setting. This heterogeneity in the presentation of depression should be considered in estimating level of depression in primary care, and in designing intervention strategies.

This work was supported by National Institute of Drug Abuse grant DA03992 and National Institute of Mental Health grant MH47447. Dr Gallo is a Brookdale Foundation National Fellow in Geriatrics. Data gathering was supported by the Epidemiologic Catchment Area (ECA) Program of the National Institute of Mental Health Division of Biometry and Epidemiology. The principal investigators and grant award numbers during data gathering were Jerome K. Myers at Yale University (MH34224), Morton Kramer at the Johns Hopkins University (MH33870), Lee N. Robins at Washington University (MH33883), Dan Blazer and Linda George at Duke University (MH35386), Richard Hough and Marvin Karno at the University of California, Los Angeles (MH35865). At the National Institute of Mental Health, principal collaborators during data gathering were Darrel A. Regier, Ben Z. Locke, William W. Eaton, Carl A. Taube, and Jack D. Burke, Jr.

REFERENCES

- American Psychiatric Association (1980). *Diagnostic and Statistical Manual of Mental Disorders (3rd ed.)*. APA: Washington, DC.
- Angoff, W. H. (1993). Perspectives on differential item functioning methodology. In *Differential Item Functioning* (ed. P. W. Holland and H. Wainer), pp. 3–24. Lawrence Erlbaum Associates: Hillsdale, NJ.
- Anthony, J. C. & Aboraya, A. (1992). The epidemiology of selected mental disorders in later life. In *Handbook of Aging and Mental Health* (ed. J. E. Birren, R. B. Sloane and G. D. Cohen), pp. 27–73. Academic Press: New York.
- Anthony, J. C., Foldstein, M., Romanoski, A. J., Von Korff, M., Nestadt, G. R., Chahal, R., Merchant, A., Brown, C. H., Shapiro, S., Kramer, M. & Gruenberg, E. M. (1985). Comparison of the lay Diagnostic Interview Scheduled and a Standardized Psychiatric Diagnosis. *Archives of General Psychiatry* **42**, 667–675.
- Barrett, J. E., Barrett, J. A., Oxman, T. E. & Gerber, P. D. (1988). The prevalence of psychiatric disorders in a primary care practice. *Archives of General Psychiatry* **45**, 1000–1006.
- Blacker, C. V. R. & Clare, W. (1987). Depressive disorder in primary care. *British Journal of Psychiatry* **150**, 735–751.
- Burvill, P. W. (1988). Psychiatric disorders in primary care. In *Handbook of Social Psychiatry*. (ed. A. S. Handerson and G. D. Burrows), pp. 221–232. Elsevier: Amsterdam.
- Burvill, P. W. & Knuiman, M. W. (1983). The influence of minor psychiatric morbidity on consulting rates to general practitioners. *Psychological Medicine* **13**, 635–643.
- Camilli, G. & Shepard, L. A. (1993). *Methods for Identifying Biased Test Items*. Sage Publication: Newbury Park, CA.
- Cooper-Patrick, L., Crum, R. M., & Ford, D. E. (1994). Characteristics of patients with major depression who received care in general medical and speciality mental health settings. *Medical Care* **32**, 15–24.

- Coulehan, J. L., Schulberg, H. C., Block, M. R. & Zettler-Segal, M. (1988). Symptom patterns of depression in ambulatory medical and psychiatric patients. *Journal of Nervous and Mental Disorder* **176**, 284–288.
- Coulehan, J. L., Schulberg, H. C., Block, M. R., Janosky, J. E. & Arena, V. C. (1990). Medical comorbidity of major disorder in a primary medical practice. *Archives of Internal Medicine* **150**, 2363–2367.
- Crum, R. M., Cooper-Patrick, L. & Ford, D. E. (1994). Depressive symptoms among general medical patients: prevalence and one-year outcome. *Psychosomatic Medicine* **56**, 109–117.
- Eaton, W. W. & Kessler, L. G. (1985). *Epidemiologic Field Methods in Psychiatry. The NIMH Epidemiologic Catchment Area Program*. Academic Press: Orlando, FL.
- Emmons, C., Fetting, J. H. & Zonderman, A. B. (1987). A comparison of the symptoms of medical and psychiatric patients matched on the Beck Depression Inventory. *General Hospital Psychiatry* **9**, 398–403.
- Fahy T. (1974). Pathways of specialist referral of depressed patients from general practice. *British Journal of Psychiatry* **124**, 231–234.
- Gallo, J. J. (1995). Epidemiology of mental disorders in middle age and late life: conceptual issues. *Epidemiologic Review* **17**, 83–94.
- Gibbons, R. D., Clark, D. C., Cavanaugh, S. V. & Davis, J. M. (1985). Application of modern psychometric theory in psychiatric research. *Journal of Psychiatric Research* **19**, 43–55.
- Goldberg, D. P. & Blackwell, B. (1970). Psychiatric illness in general practice: a detailed study using a new method of care identification. *British Medical Journal* **ii**, 439–443.
- Goldberg, D. P. & Huxley, P. (1980). *Mental Illness In The Community*. Tavistock: London.
- Goldberg, D. P. & Huxley, P. (1992). *Common Mental Disorders: A Bio-social Model*. Tavistock Press: London.
- Hankin, J. R., Steinwachs, D. M., Rieger, D. A., Burns, B. J., Goldberg, I. D. & Hoepfer, E. W. (1982). Use of general medical services by mental disorders. *Archives of General Psychiatry* **39**, 225–231.
- Helzer, J. E., Robins, L. N., McEvoy, L. T., Spitznagel, E. L., Stoltzman, R. K., Farmer, A. & Brockington, I. F. (1985). A comparison of clinical and diagnostic interview schedule diagnosis. *Archives of General Psychiatry* **42**, 657–666.
- Hoepfer, E., Nycz, G. R., Cleary, P., Regier, D. & Goldberg, I. (1979). Estimated prevalence of RDC mental disorder in primary medical care. *International Journal of Mental Health* **8**, 6–15.
- Horwath, E., Johnson, J., Klerman, G. L. & Weissman, M. M. (1992). Depressive symptoms as relative and attributable risk factors for first-onset major depression. *Archives of General Psychiatry* **49**, 817–823.
- Katon, W. (1982). Depression: somatic symptoms and medical disorders in primary care. *Comprehensive Psychiatry* **23**, 274–287.
- Katon, W., Vitaliano, P. P., Russo, J., Cormier, L., Anderson, K. & Jones, M. (1986). Panic disorder: epidemiology in primary care. *Journal of Family Practice* **23**, 233–239.
- Kessler, L. G., Cleary, P. D. & Burke, J. D. (1985). Psychiatric disorders in primary care: results of a follow-up study. *Archives of General Psychiatry*, **42**, 583–587.
- Kessler, L. G., Burns, B., Shapiro, S., Tischler, G. L., George, L. K., Hough, R. L., Bodison, D. B. & Miller, R. H. (1987). Psychiatric diagnoses of medical service users: evidence from the epidemiologic catchment area program. *American Journal of Public Health* **77**, 18–24.
- Kim, J.-O. & Mueller, C. W. (1978). *Introduction to Factor Analysis*. Sage Publication: Newbury Park, NY.
- Knights, E. B. & Folstein, M. F. (1977). Undiscovered emotional and cognitive disturbance in medical patients. *Annals of Internal Medicine* **87**, 723–724.
- Lebo, A., Perez-Echeverria, M.-J., Artal, J., Rubio, L., Escobar, M.-V., Gonzalez-Torreccillas, J.-L., Morera, B., Dia, J.-L. & Miranda, M. (1988). Psychiatric morbidity among medical outpatients in Spain: A case for new methods of classification. *Journal of Psychosomatic Research* **32**, 355–364.
- Mann, A. H., Jenkins, R. & Belsey, E. (1981). The twelve month outcome of patients with neurotic illness in general practice. *Psychological Medicine* **11**, 535–550.
- Marino, S., Gallo, J. J., Ford, D. & Anthony, J. C. (1995). Filters on the pathway to mental health care: I. Incident mental disorders. *Psychological Medicine* **25**, 1135–1148.
- Mirowsky, J. & Ross, C. E. (1989). Psychiatric diagnosis as reified measurement. *Journal of Health and Social Behaviour* **30**, 11–25.
- Muthén, B. O. (1988). *LISCOMP: Analysis Of Linear Structural Equations With A Comprehensive Measurement Model (2nd edn)*. User's Guide. Scientific Software: Mooresville, IN.
- Muthén, B. O. (1989). Dichotomous factor analysis of symptom data. *Sociological Methods and Research* **18**, 19–65.
- Muthén, B. O. (1995). Factor analysis of alcohol abuse and dependence symptom items in the 1988 National Health Interview Survey. *Addiction* **90**, 637–645.
- Myers, J. K., Weissman, M. M., Tischler, G. L., Holzer III, C. E., Leaf, P. J., Orvaschel, H., Anthony, J. C., Boyd, J. H., Burke, J. D., Kramer, M. & Stoltzman, R. (1984). Six-month prevalence of psychiatric disorders in three communities, 1980 to 1982. *Archives of General Psychiatry* **41**, 959–967.
- Newmann, J. P. (1989). Aging and depression. *Psychology and Aging* **4**, 150–165.
- Olsson, M. & Klerman, G. L. (1992). Depressive symptoms and mental health service utilization in a community sample. *Social Psychiatry and Psychiatric Epidemiology* **27**, 161–167.
- Ormel, J., Van Den Brink, W., Koeter, M. W. J., Giel, R., Van Den Meer, K., Van De Willige G. & Wilmsink, F. W. (1990). Recognition, management, and outcome of psychological disorders in primary care: a naturalistic follow-up study. *Psychological Medicine* **20**, 909–923.
- Osterlind, S. J. (1984). *Test Item Bias*. Newbury Park: Sage Publication.
- Peréz-Stable, E., Miranda, J., Muñoz, R. & Ying, Y. (1990). Depression in medical outpatients: Underrecognition and misdiagnosis. *Archives of Internal Medicine* **150**, 1083–1088.
- Pilowsky, I. & Spence, N. (1978). Depression: inside and outside the hospital setting. *British Journal of Psychiatry* **132**, 265–270.
- Regier, D. A., Narrow, W. E., Rae, D. S., Manderscheid, R. W., Locke, B. Z. & Goodwin, F. K. (1993). The de facto US mental and addictive disorders service system: epidemiologic catchment area prospective 1-year prevalence rates of disorders and services. *Archives of General Psychiatry* **50**, 85–94.
- Robins, L. N., Helzer, J. E., Croughan, J. & Ratcliff, K. S. (1981). National Institute of Mental Health Diagnostic Interview Schedule: its history, characteristics, and validity. *Archives of General Psychiatry* **38**, 381–389.
- Robins, L. N. & Regier, D. A. (1991). *Psychiatric Disorders in America*. Free Press: New York.
- Sanson-Fisher, R. W. & Hennrikus, D. J. (1988). Why do primary care physicians often fail to detect psychological disturbance in their patients? In *Handbook of Social Psychiatry* (ed. A. S. Handerson and G. D. Burrows), pp. 245–256. Elsevier: Amsterdam.
- Schulberg, H. C., Saul, M., McClelland, M., Ganguli, M., Christy, W. & Frank, R. (1985). Assessing depression in primary medical and psychiatric practices. *Archives of General Psychiatry* **42**, 1164–1170.
- Shapiro, S., Skinner, E. A., Kessler, L. G., Von Korff, M., German, P. S., Tischler, G. L., Leaf, P. L., Bemham, L., Cottler, L. & Regier, D. A. (1984). Utilization of health services: three Epidemiologic Catchment Area sites. *Archives of General Psychiatry* **41**, 971–978.
- Sireling, L. I., Freeling, P., Paykel, E. & Rao, B. (1985). Depression in general practice: clinical features and comparison with outpatients. *British Journal of Psychiatry* **147**, 119–125.
- Spitzer, R. L., Sheehy, M. & Endicott, J. (1977). DSM-III: guiding principles. In *Psychiatric Diagnosis* (ed. V. Rakoff, H. Stancer and H. Kedward), pp. 1–24. Macmillan: London.
- Spitzer, R. L., Kroenke, K., Linzer, M., Hahn, S. R., Williams, J. B. W., deGruy III, F. V., Brody, D. & Davies, M. (1995). Health-

- related quality of life in primary care patients with mental disorders. Result from the PRIME-MD 1000 Study. *Journal of American Medical Association* **274**, 1511–1517.
- Swaminathan, H. & Rogers, H. (1990). Detecting differential item functioning using logistic regression procedures. *Journal of Educational Measurement* **27**, 361–370.
- Von Korff, M. & Anthony, J. C. (1982). The NIMH Diagnostic Interview Schedule modified to record current mental status. *Journal of Affective Disorder* **4**, 365–371.
- Von Korff, M., Cottler, L., George, L. K., Eaton, W. W., Leaf, P. J. & Burnam, A. (1985). Nonresponse and nonresponse bias in the ECA surveys. In *Epidemiologic Field Methods in Psychiatry. The NIMH Epidemiologic Catchment Area Program* (ed. W. W. Eaton and L. G. Kessler), pp. 85–98. Academic Press: Orlando, FL.
- Wells, K. B., Burnam, M. A., Rogers, W., Hays, R. & Camp, P. (1992). The course of depression in adult outpatients: results from the Medical Outcomes Study. *Archives of General Psychiatry* **49**, 788–794.
- West, S. G., Finch, J. F. & Curran, P. J. (1995). Structural equation models with nonnormal variables: problems and remedies. In *Structural Equation Modeling: Concepts, Issues, and Applications* (ed. R. H. Hoyle), pp. 56–75. Sage: Oaks, California.
- Williamson, P. S. & Yates, W. R. (1989). The initial presentation of depression in family practice and psychiatric outpatients. *General Hospital Psychiatry* **11**, 188–193.
- Zich, J. M., Attkisson, C. C. & Greenfield, T. K. (1990). Screening for depression in primary care clinics: the CES-D and the BDI. *International Journal of Psychiatry in Medicine* **20**, 259–277.