Self-reported use of mental health services versus administrative records: care to recall?

ANNE E. RHODES,[1] KINWAH FUNG[2]

1 Department of Psychiatry and Public Health Sciences University of Toronto, Canada; Suicide Studies and Inner City Health Research Units, St Michael's Hospital, Toronto, Canada; Adjunct Scientist, Health Systems Research and Consulting Unit, The Centre for Addictions and Mental Health, Toronto, Canada; Adjunct Scientist, Institute for Clinical Evaluative Sciences, Sunnybrook and Women's College Health Sciences Centre, Toronto, Canada
2 Biostatistician, Institute for Clinical Evaluative Sciences, Sunnybrook and Women's College, Health Sciences Centre, Toronto, Canada

Abstract

Estimates of the level of unmet need for mental health treatment often rely on self-reported use of mental health services. However, depressed persons may over-report their use in relation to administrative records if they are highly distressed. This study seeks to replicate and explicate the finding that persons at a high level of distress report more mental health service use than recorded in their healthcare records. The study sample, N = 36,892, 12 years and older, was drawn from the 1996/97 Ontario portion of the Canadian National Population Health Survey. Respondents were individually linked to their administrative mental healthcare records 12 months backward in time. Of these, 96.5% agreed to the link and 23,063 (62.5%) were linked. Almost two-thirds of those who were depressed in the past year were currently at a high level of distress. Differential reporting of use for highly distressed persons in excess of 100% remained in the use of different types of physician providers after adjustments for other potential determinants of use. Telescoping was also not an explanation. The patterns of differential reporting between groups expected to diverge and converge in their recall ability were consistent with a recall bias. As this study was not able to rule out a recall bias, it further accentuates concerns about the impact of bias in the measurement of mental health-service use and inferences made concerning the determinants of use.

Key words: self-reports, mental health services, health services research, depression, mental recall

Introduction

A consistent finding in mental health surveys internationally is the sizeable gap between those who have a mental disorder and their use of mental health services. This body of work has been critical in documenting the treated prevalence of mental illness and level of unmet in the population (Bilj et al., 2003). A potential difficulty with these findings has been the reliance on self-reported mental health service use. There is no gold standard in the measurement of mental health service use. Persons with mental illness may systematically under or over report their use. There is a scant literature concerning the consistency of self-reported and administrative records of mental health service use and large representative samples have been rare (Evans and Crawford, 1999; Rhodes et al., 2002).

This study follows up on findings from a previous study to determine why those in the general population with higher levels of distress reported more mental health service use than was recorded in their administrative healthcare records (odds ratio of 9.21 versus 2.67 yielding a relative difference of 245%). Among those who were depressed in the past year most were currently distressed at a high level (Rhodes et al., 2002). A major concern for researchers and healthcare planners, therefore, is how to measure the use of mental health services for those who are depressed. The aims of this study are twofold: first, to replicate the finding of differential reporting and, second, to examine the evidence for a recall bias.

Background and rationale

Recall bias is a form of differential misclassification

(Rothman and Greenland, 1998). In epidemiology, recall bias is often described in the context of case-control studies where the cases have knowledge of their diagnosis and this knowledge affects their recall of potential putative exposures. Recall bias occurs when cases are 'better' at recalling exposures or events than the controls. Mathematically, the sensitivity and specificity of what is being recalled is higher in the cases than in the controls. Accordingly, when an association (for example, such as an odds ratio) between the supposed risk factor and disease under study is calculated, it is biased upwards.

In a similar manner, survey respondents who are currently at a high level of distress may recall their use of mental health services 'better' than those who are not. Mood states are known to influence memory (Koriat et al., 2000). When persons are depressed, they are better at recalling past depressed states or stressful life events (Aneshensel et al., 1987; Simon and VonKorff, 1995). This may well extend to their past use of mental health services. Note that recall bias is about a difference between two groups in recall ability. Those who are not currently distressed may have used mental health services as much as, or more than, those who are currently distressed. However, the former group may not recall their past use as well as the latter.

Mathematically, the denominator of the odds ratio, the odds of use among those who are not highly distressed is too low. Therefore, when the odds ratio is calculated, it is biased upwards.

It is also important to recognize that even if *most* persons who were *depressed* in the past year are *currently distressed* and, therefore, report their use 'better' than those who are not currently distressed (the numerator of the odds ratio), this does not mean that those who are depressed are 'good' at recalling their use. Without a gold standard, this latter question cannot be addressed. In relation to a gold standard, the sensitivity and specificity of depressed person's self-reported use might be quite low. While neither administrative records nor self-reported use are gold standards of use, the discrepancy between them may be caused by a recall bias. Accordingly, the odds ratios of use would be higher in the self-reported data than in the administrative records.

However, there are several other mechanisms that may have accounted for the discrepancy observed. The first and perhaps most obvious explanation is that the additional use reported by the currently distressed was real, due to use not captured in the administrative healthcare records. Under coverage could have come from two sources: providers not covered in the administrative records and from providers who were covered but the contacts were not coded as mental health claims. Within Ontario, only physicians are reimbursed through the Ontario Health Insurance Program (OHIP) for mental health services; therefore, the use of non-medical providers is not captured within administrative healthcare records. Under coverage from the first source would not appear to be a full explanation of differential reporting, if it remained after these users were removed. It is more complicated to examine under coverage from the second source due to the nature of the codes used for reimbursement purposes. Yet, the presence of under coverage does not preclude the presence of a recall bias. In fact, some undercoverage scenarios may be indicative of recall bias. One such scenario is the use of psychiatrists *only*.

By definition, all claims made by a psychiatrist were identified as mental health claims so the coverage of mental health contacts to psychiatrists would be better than the coverage of mental health contacts to other types of physicians in the administrative records. If differential reporting continued to be found in those who reported seeing only psychiatrists, this would possibly be due to recall bias. Given the way recall bias works, highly distressed persons recall their use 'better' than those who were not highly distressed. Accordingly, if differential reporting remained for those who reported seeing psychiatrists only, it may be because these clients really did have contacts with their psychiatrists. The contacts would not have been eligible for reimbursement if they were brief or over the phone. While it is possible that some clients misclassified their use of psychiatrists – for example, if they really saw a psychologist – it is suspected such an effect would be small given the way mental healthcare is publicly financed. Only a small proportion would have been able to access non-medical providers (CANMAT, 1999) and presumably an even smaller portion would have misclassified this use.

In addition, if recall bias were present, then one could anticipate certain patterns of differential reporting. Differential reporting would be most apparent in groups where one would expect divergence in recall ability and least apparent in groups where one would expect convergence in recall ability. An example of where divergence would be greater would be

comparisons between those who had used in the past year in contrast to those who had not (never used or used longer than a year ago). Convergence in recall and, therefore, little differential reporting would be expected between groups that had used mental health services in the past year – in other words, among users of primary care physicians and psychiatrists.

Lastly, rather than recall bias, differential reporting may arise from telescoping. Those who are currently distressed may not be 'better' at recalling their use in the past year than those who are not. They may simply be telescoping their use more than those who are not currently distressed. Telescoping occurs when events from the past are recalled as taking place more recently than they did (Bradburn, 1983). Accordingly, persons who are highly distressed and/or depressed may report mental health visits as taking place in the past year when really, they occurred several years ago. If telescoping rather than recall bias were the explanation, then one would expect the self-reported use of the currently distressed in the past year to correspond more closely to the use captured in the administrative records for periods of time extending beyond one year.

Methods

Study design and sample
This was a cross-sectional study that compared the self-reported use of mental health services within the Ontario portion of the 1996/1997 National Population Health Survey in Canada (OHS 96) individually linked with administrative mental healthcare records: physician reimbursement claims in the Ontario Health Insurance Plan (OHIP) and inpatient discharge abstract data from the Canadian Institute for Health Information (CIHI). The OHS 96 employed a complex household survey sampling design with computer-assisted telephone interviews. Persons living on reserves, armed forces bases, institutions and remote areas were excluded. Some general information was collected on all members of the household and then one member of the household 12 years of age or older was randomly selected for a more detailed health interview and asked consent for record linkage (Statistics Canada, 1998).

The overall response rate to the OHS 96 was calculated by multiplying the Ontario household response rate (78.8%) with the Ontario selected person response rate (94.4%). This yielded a full sample of 36,892. Weights were created by Statistics Canada to represent the population of Ontario: 96.5% consented to the linkage and 23,063 respondents were successfully linked (62.5%) and formed the *study sample*. Linkage was performed using custom-written computer programs and based on combinations of OHIP number, name, birth date and address fields. A separate set of weights was also created for the linked sample. In Canada, 1% to 2% of households do not have a telephone service. Differences between those with and without telephones were taken into account in the weighting scheme (Ciok, 1993; Y. Beland, personal communication, Household Survey Methods Division, Statistics Canada, 2002). As only about 46.5% of the original sample was in the linked sample, caution is required in making inferences from these data to the original sample and the population of Ontario. Table 1 indicates that the differences in the distributions of respondent characteristics and self-reported use were minimal between the full and linked samples.

Measures

Self-reported mental health service use (SR)
The questions in the 96/97 NPHS concerning mental health service use that were asked of the individual respondents were as follows:

> In the past 12 months, that is, from (date 12 months ago) to yesterday, have you seen or talked on the telephone to a health professional about your emotional or mental health?

> How many times (in the past 12 months)? Whom did you see or talk to? (The interviewer is instructed to read the list and mark all that apply).

> 1. Family doctor or general practitioner
> 2. Psychiatrist
> 3. Psychologist
> 4. Nurse
> 5. Social worker or counselor
> 6. Other (specify)

Self-reported versus no self-reported use was defined in a hierarchical fashion based upon contact with physicians. Anyone reporting contact with a psychiatrist was classified as being seen by a psychiatrist; anyone reporting contact with a family doctor or general practitioner (but not to a psychiatrist) was classified as primary care user as in other studies (Wells et al., 1987; Leaf et al., 1988; Kessler et al., 1997). Use to each of

Table 1. Differences in the distributions of respondent characteristics and self-reported use

Baseline respondent characteristics	Full sample N = 36,892 (miss) unweighted %	Full sample weighted %	Linked study sample n = 23,063 (miss) unweighted %	Linked Study sample weighted %
Permission to link data	96.5	96.6	100	100
Telephone interview	99.6	99.6	99.4	99.4
Depression	4.3 (1268)	4.1	4.6 (638)	4.1
Distress	(1519)		(752)	
None	40.3	40.6	37.4	37.9
Low	26.2	26.0	27.6	27.2
Medium	21.1	21.3	22.1	22.1
High	12.3	12.2	13.0	12.8
Age				
12-29	24.8	29.1	23.7	29.0
30-41	26.1	25.6	24.9	25.5
42-59	25.5	26.5	25.7	26.3
60-102	23.6	18.8	25.7	19.2
Sex				
Male	46.4	49.0	46.7	49.0
Female	53.6	51.0	53.3	51.0
Marital status	(95)		(12)	
Partner	53.8	57.4	54.1	57.9
Single	28.3	31.1	26.7	30.5
Separated/divorced	8.9	6.2	9.4	6.3
Widowed	9.1	5.4	9.7	5.4
Residence	(43)		(17)	
Urban	84.8	87.2	84.3	86.9
Ethnicity	(189)		(86)	
Caucasian	92.1	86.6	93.4	88.0
Education	(445)		(175)	
Secondary or less	27.5	28.1	27.8	28.2
Secondary grad.	18.6	18.4	18.2	18.2
Some post sec.	14.4	13.4	15.0	13.8
Some university	6.5	7.0	6.4	7.0
Trade/diploma	17.5	16.4	17.8	16.9
University	15.6	16.7	14.8	15.8
Labour force activity	(641)		(242)	
Ill/disability	4.4	3.7	4.5	3.6
Responsible	5.3	5.2	5.2	5.3
School	5.6	7.4	4.9	7.0
Off work	4.3	4.7	4.3	4.9
Retired	15.1	12.1	16.2	12.2
Work	65.4	66.9	65.0	67.1
Household income	(9440)		(4179)	
None to 9,999	4.5	3.4	4.3	3.2
$10,000 to 14,999	8.1	5.6	8.3	5.6
$15,000 to 19,999	7.9	6.4	8.0	6.5
$20,000 to 29,999	14.0	12.2	14.2	12.0
$30,000 to 39,999	14.2	13.9	14.4	14.0
$40,000 to 49,999	13.0	13.9	12.7	13.7
$50,000 to 59,999	12.8	14.2	12.8	14.4
$60,000 to 79,999	12.4	14.1	12.3	14.2
$80,000 or more	13.0	16.4	13.1	16.6
Disability days	(34)		(18)	
None	90.7	91.6	89.7	91.1
1–7	6.9	6.6	7.6	6.8
8–14	2.4	1.9	2.8	2.1

Table 1. Cont.

Baseline respondent characteristics	Full sample N = 36,892 (miss) unweighted %	Full sample weighted %	Linked study sample n = 23,063 (miss) unweighted %	Linked Study sample weighted %
Perceived health status				
Excellent	24.5	25.4	24.0	25.5
Very good	38.6	39.2	38.9	39.4
Good	25.4	25.3	25.4	25.4
Fair	8.5	7.5	8.7	7.3
Poor	3.0	2.6	3.0	2.4
Chronic health problems	(298)		(171)	
None	38.4	42.5	36.4	41.9
One	26.4	26.7	26.7	27.2
Two or more	35.3	30.8	37.0	30.9
Alcohol dependence	(572)		(269)	
	0.7	0.7	0.8	0.8
Alcohol consumption – past month	(385)		(155)	
At least 1	53.9	52.4	54.9	53.3
Less than 1	21.2	20.6	21.2	20.7
None past year	14.1	13.0	14.3	12.8
Never	10.8	14.0	9.7	13.3
Self-reported mental health service use[1]	(922)		(454)	
No use	92.9	93.6	92.6	93.4
Psychiatrist	1.6	1.5	1.6	1.5
Family physician or general practitioner	3.3	2.9	3.4	3.0
Other providers	2.3	2.1	2.4	2.1

[1]Assigned in hierarchical fashion such that categories are mutually exclusive.

the remaining types of providers listed above was less common and therefore, a third category mutually exclusive of physician use, was created which grouped these non-medical providers together.

Administrative records of mental healthcare use (AD)

The OHIP is a computerized system in which physicians submit claims of medical service use encounters for payment of services. In Ontario, the majority of physicians' direct patient care (94 %) is captured within the administrative healthcare records (Chan, 1999; Chan and Austin, 2003) including psychiatrists (B Woodside, personal communication with the President of the Canadian Psychiatric Association, 2003). Mental healthcare was defined as in the prior study according to physician reimbursement *claims* made for a subset of a list of codes designated as core mental healthcare by the Ontario Ministry of Health

(Rhodes et al., 2002). All claims made by psychiatrists were identified as core mental health claims. For physicians who were not psychiatrists, core mental health claims were made up of codes that include procedures related to inpatient, emergency department and ambulatory contacts with physicians for individual, group or family therapy or other mental health procedures such as electroconvulsive therapy. Although diagnostic information was present in the claims data, it was not used due to uncertainty about reliability (Schwartz et al., 1980; Towery, Sharfstein and Goldberg, 1980) particularly within primary care settings (Pignone et al., 2002). Core mental health *claims* were specific to each respondent and corresponded to a 12 months time interval before the survey interview date.

A *mental health visit* was defined according to whether one or more core mental health claims was made to the same physician on the same day for a

respondent. A mental health visit was categorized in the same hierarchical manner as noted for self-reported use. An inpatient mental health stay was determined by specific OHIP claims pertaining to inpatient service and/or an inpatient stay with a most responsible diagnosis of an ICD-9 mental disorder (codes: 290–319) in the hospital abstract data collected by the Canadian Institute of Health Information (CIHI). Persons who had an inpatient mental health stay (n = 88), less than 1% of the sample, were excluded as this group has been found to under report their use in relation to administrative records (Clark et al., 1996; Rhodes et al., 2002; Hennesy and Reed, 1992).

Respondent characteristics

The presence of depression in the past year was based on the University of Michigan Composite International Diagnostic Interview (UM-CIDI) short form (CIDI-SF). This scale was developed in the National Comorbidity Survey (NCS) and piloted within the National Health Interview Survey in the United States. It is a shortened version of the original UM-CIDI that imparts a diagnosis of major depression based on DSM-III-R and ICD-10. Within the NCS, the sensitivity of the CIDI-SF depression for the UM-CIDI depression was 90% and the specificity was 94% (Kessler, Andrews, Mroczek, Ustun and Wittchen, 2001; Kessler et al., 2001). A cut point of a 90% probability of depression in the NCS was used in this study as in the prior study (Rhodes et al., 2002) and others (Beaudet, 1996; Diverty et al., 1997; Cairney et al., 1999) to define the presence of depression in the past year.

Distress was measured with the K-6 (Kessler et al., 2002). This scale was designed to supplement diagnostic measures of mental illness with a dimensional measure of severity within population-based surveys. Items were generated from a large pool of items from existing surveys and then reduced using item response theory. The scale was then piloted in two surveys and a clinical calibration study. The properties of the instrument were further investigated in the Australian National Survey of Mental Health and Well-Being and has been found to be a good indicator of current anxiety disorder and major depression (Andrews and Slade, 2001; Furukawa et al., 2003). The K-6 is made up of 6 items scored on a five-point scale (none of the time = 0; a little of the time = 1; some of the time = 2, most of the time = 3 and all of the time = 4). As in the prior study, this scale was divided into 4 levels: none

(score of zero); low (score of 1 to 2); medium (score of 3 to 5) and high (score of 6 to 24) (Rhodes et al., 2002).

Measures of health and function included one or more disability days in the past 14 days, perceived health status, chronic health problems, alcohol dependence in the past year and level of alcohol consumption in the past month. Respondents were also classified by their age, sex, urban/rural residence, marital status, ethnicity, highest level of education attained, labour force activity and household income.

Statistical analyses

Differential reporting was examined as in the previous study. An odds ratio of 1 implies that use is equally likely in those who are highly distressed and those who are not. An odds ratio greater than 1 implies that use is more likely in those who are highly distressed than in those who are not. Odds ratios and respective confidence intervals of self-reported use (SR) and of administrative record use (AD) were calculated. The percentage differences in these odds ratios, (SR OR - AD OR /AD OR) *100 and a lack of overlap in confidence intervals was determined (Austin and Hux, 2002).

There were two dependent variables: self-reported physician use (no use, primary care physician use only, psychiatrist use) and administrative recorded physician use (no use, primary care physician use only, psychiatrist use). To estimate associations between levels of distress and depression with the type of physician use, two separate multinomial logistic regression models were fit, one for self-reported use (SR) and the other for administrative recorded use (AD). Within each model, four equations were estimated. One of the four equations was the inverse of another, so three were of interest: primary care physician use only versus no primary care physician use only; psychiatrist use versus no psychiatrist use and psychiatrist use versus primary care use only. For the purposes of this study, the first two equations permitted an assessment of differential reporting in groups expected to be divergent in their recall ability and the third equation, in groups expected to be convergent in their recall ability. Odds ratios were adjusted for respondent characteristics. Multinomial logistic regression was employed to estimate these equations instead of logistic regression as the odds ratios estimated in multinomial logistic regression are mathematically equivalent to those in

logistic regression. Advantages of multinomial regression are fewer models need to be fit and there is some gain in precision (Hosmer and Lemeshow, 1989; Allison, 2000).

To address further the potential for a selection bias arising from the less than full linkage (see above and Table 1), the multinomial logistic regression model was re-run in the SR *unlinked* sample to determine whether differences between the SR unlinked (original sample) and SR linked sample would yield different results. As there was a sizeable amount of missing data in household income (18%) (see Table 1) a sensitivity analyses was conducted to determine what impact, if any, the missing data would have had on the differential reporting observed. Each missing value was substituted with each income level in successive imputations (Vach and Blettner, 1991; Greenland and Finkle, 1995). Adjusted odds ratios that changed by 10% or more and were no longer statistically significant were considered sensitive to assumptions in the missing data.

The persistence of differential reporting was examined again after removing persons who reported use to non-medical providers. People who *only* reported seeing non-medical providers, n = 515 or 2% of self-reported use, were excluded and the SR, AD associations re-estimated. Persons who reported they saw a physician and another type of provider, n = 155 or 0.49% of self-reported use, were not excluded as it would not be possible to determine whether coverage differences were arising from mental health contacts (but not full visits) with primary care physicians or non-medical providers.

In addition, the presence of differential reporting was examined under the scenario of psychiatrist use *only*. The sample was restricted to persons who reported they had not seen any provider in the past year (n = 20,899); those who reported they had *only* seen a primary care physician in the past year (n = 663) and those who reported they had *only* seen a psychiatrist in the past year (n = 243).

Finally, the potential for telescoping was assessed by the magnitude of the differential reporting in the past year (described above) to three analogous models in which AD use was measured for time periods extending beyond one year, i.e. over the past two years, three years and four years.

All analyses were initially conducted using unweighted data with the SAS statistical software package. Analyses were then repeated using the weights and a bootstrapping technique to provide estimates and their respective variances that represent the population of Ontario (Rao, Wu and Yue, 1992; Rust and Rao, 1996). Unless stated otherwise, all estimates reported were weighted.

Results

In the linked sample, 4.1% were depressed in the past year and 12.8% were at a high level of distress in the passed month (Table 1). Almost two-thirds (62%) of those who were depressed in the past year were *currently* at a high level of distress, i.e., in the past month. Among those who were depressed in the past year and *currently* at a high level of distress, 75.9% of their self-reported use was to physicians and the remaining 24.1% to other types of providers.

The SR and AD odds ratios of physician use for levels of distress and for depression in the past year are shown in Tables 2 (crude ORs) and 3 (adjusted ORs). The percentage differences between the SR and AD odds ratios demonstrate that persons at higher levels of distress or depression were more likely to report use than was recorded in the AD records. The differences exceed 100% for primary care physician use and for psychiatrist use (crude and adjusted ORs). There were fluctuations in the 95% confidence intervals with respect to overlap that may represent a loss of statistical power. For comparisons among physician users – psychiatrist versus primary care physician use, the percentage differences in the odds ratios were all under 100% and all the corresponding SR and AD 95% confidence intervals overlapped.

The direction of the results from the SR unlinked original sample and SR linked sample stayed the same with some shifts in statistical significance for persons at the lowest level of distress. The results were sensitive to income imputations in two places. In the SR model, the association between low distress and psychiatrist use no longer was statistically significant (regardless of income level imputed). The adjusted odds ratio was attenuated to about 1.2 for each imputation, a difference of about 41%. In both AD and SR models, the association between depression and use of psychiatrists as opposed to primary care physicians also became non-significant. The adjusted odds ratio was attenuated to about 1.4 (regardless of income level imputed). This difference was much smaller, just under 15%.

Table 2. A comparison between self-reported use of physicians for mental health reasons and administrative records – distress, depression crude associations

	AD OR[1] (95% CI)	SR [2] OR (95% CI)	% Difference in SR and AD ORs
Primary care physician only versus no primary care physician only			
No distress	Reference	Reference	
Low distress	1.24 (1.01;1.52)	1.96 (1.39;2.76)	57.8
Medium distress	1.97 (1.60;2.41)	4.08 (2.99;5.55)	107.3*
High distress	3.96 (3.24;4.85)	11.15 (8.36;14.87)	181.5*
Depression	6.13 (4.94;7.61)	13.46 (10.72;16.91)	119.6*
Psychiatrist versus no psychiatrist use			
No distress	Reference	Reference	
Low distress	1.28 (0.87;1.87)	1.61 (0.77;3.36)	26.1
Medium distress	2.37 (1.62;3.33)	4.53 (2.18;9.44)	95.3
High distress	7.67 (5.46;10.82)	18.95 (9.60;37.4)	146.5
Depression	11.71 (8.99;15.27)	23.86 (17.38;32.76)	103.7 *
Psychiatrist versus primary care physician only			
No distress	Reference	Reference	
Low distress	1.03 (0.66;1.60)	0.82 (0.36;1.87)	−20.1
Medium distress	1.18 (0.79;1.77)	1.11 (0.50;2.47)	−5.8
High distress	1.94 (1.32;2.86)	1.70 (0.81;3.57)	−12.4
Depression	1.91 (1.41;2.60)	1.77 (1.23;2.55)	−7.3

[1] N=22,342 (633 missing) in the multinomial logistic regression analyses for depression
  N=22,228 (747 missing) in the multinomial logistic regression analyses for distress
[2] N=22,332 (643 missing) in the multinomial logistic regression analyses for depression
  N=22,222 (753 missing) in the multinomial logistic regression analyses for distress
* 95% CIs do not overlap
AD = administrative records SR = self-reported OR = odds ratio CI = confidence interval

The pattern of the percentage differences and non-overlap in 95% confidence intervals for persons at different levels of distress or depression when persons who *only* used non-medical providers were excluded was similar to the adjusted results observed above. When 90% confidence intervals were calculated, the adjusted odds ratio of use of primary care physicians continued to not overlap for persons at the highest level of distress. As well, the 90% confidence intervals for the adjusted odds ratio of psychiatrists did not overlap for persons at the medium level of distress. Comparisons of the SR model (linked sample) in the past year with AD models extending further back in time also showed a similar pattern. Consequently, telescoping was not apparent as the percent difference between the SR and AD adjusted odds ratios did not diminish over time.

Discussion

This population-based study in Ontario, Canada replicated the previous finding that persons at higher levels of distress are more likely to report their mental health use than those who are not. Differential reporting in excess of 100% for use of primary care physicians and for use of psychiatrists persisted after adjustments for other potential determinants of use and was not explained by the removal of users of non-medical providers or by telescoping. Moreover, the patterns of differential reporting in groups expected to be divergent and convergent in their recall ability were consistent with what would be expected given a recall bias. This was also true for the scenario of psychiatrist use *only*.

It is important to recognize that the study sample was a household one that excluded groups that may

Table 3. A comparison between self-reported use of physicians for mental health reasons and administrative records – distress, depression adjusted associations

	AD OR[1,3] (95% CI)	SR OR[2,3] (95% CI)	% Difference in SR and AD ORs
Primary care physician only versus no primary care physician only			
No distress	Reference	Reference	
Low distress	1.02 (0.81;1.27)	1.37 (0.96;1.95)	34.4
Medium distress	1.50 (1.19;1.89)	2.35 (1.65;3.33)	56.6
High distress	2.18 (1.68;2.83)	4.50 (3.13;6.48)	106.2 *
Depression	3.29 (2.54;4.26)	5.06 (3.80;6.74)	53.7
Psychiatrist versus psychiatrist use			
No distress	Reference	Reference	
Low distress	1.11 (0.76;1.63)	2.03 (1.01;4.07)a,b,c,d	82.2
Medium distress	1.65 (1.12;2.44)	3.94 (1.99;7.79)	138.5
High distress	3.25 (2.15;4.94)	7.42 (3.60;15.29)	128.0
Depression	5.41 (3.64;8.05)	8.63 (5.41;13.76)	59.5
Psychiatrist versus primary care physician only			
No distress	Reference	Reference	
Low distress	1.10 (0.71;1.69)	1.49 (0.68;3.24)	3.8
Medium distress	1.10 (0.71;1.71)	1.68 (0.79;3.58)	35.6
High distress	1.49 (0.91;2.44)	1.65 (0.74;3.67)	52.4
Depression	1.64 (1.06;2.55) a,b,c,d	1.71 (1.02;2.85) a,b,c,d	3.8

[1]N = 17,710 (5265 missing) in the multinomial logistic regression analyses
[2]N = 17,708 (5267 missing) in the multinomial logistic regression analyses
[3]Adjusted for: age, sex, marital status, urban residence, ethnicity, education, labour force activity, household income, disability days, perceived health status, chronic health problems, alcohol dependence and alcohol consumption
[a,b,c,d] Sensitive to imputations when missing values in income are set to a) none to less than or equal to $9,999 or to b) greater than or equal to $10,000 but less than or equal to $29,999 or to c) greater than or equal to $30,000 but less than or equal to $80,000 or d) greater than or equal to $80,000 in previous year
* 95% CIs do not overlap.
AD = administrative records SR = self-report OR = odds ratio CI = confidence interval

well suffer from more severe depression and other less common, more severe forms of mental illness. A central issue for policymakers is how much money should be allocated to treat less severe forms of mental illness. Given the greater prevalence of subthreshold conditions in the population, they probably pose a greater burden to society than major depression (Horwath et al., 1992; Judd et al., 1996; Pincus et al., 1999). It is therefore, crucial to document their use of mental health services accurately.

It was difficult to attain precision of estimates of use of different provider types even with a linked study sample size of over 23,000 people for more common mental illness. For this reason, persons with an inpatient stay were excluded from the sample. As noted above, these persons may under report their use of mental health services in relation to administrative records. If the original sample had been fully linked, the precision of the adjusted estimates would likely have been improved. Fortuitously, there was little evidence of a selection bias between the original sample and the linked sample that would have affected the differential reporting observed. Another difficulty often inherent in household surveys is the sizeable amount of missing data in household income. In a setting of universal medical insurance coverage, though, income does not appear to be a determinant of mental health service use (Katz et al., 1997; Kessler, Frank, Edlund, Katz, Lin and Leaf, 1997; Alegria et al., 2000). In this study, differential reporting for persons at high levels of distress was not sensitive to income imputations.

If recall bias is present, then associations between distress and use may be too high in self-reported data.

This would extend to associations between depression and self-reported use as most persons who were depressed in the past year were also highly distressed. Depression may not be as strong a determinant of use as depicted in previous studies. Other characteristics, not necessarily equated with need, may be as or more strongly related to use than depression. From the standpoint of allocating resources to need, the magnitude of these associations with use compared to depression may not be acceptable.

Ideally all jurisdictions would be able to link survey data with administrative records to compare associations between SR and AD estimates of use in populations. If recall bias was a concern then associations could be examined in administrative records to determine whether inferences about associations with use were altered. However, capturing use beyond that which is insured becomes difficult. Administrative records may be available but difficult to access without relying on respondent recall. Privacy policies may also restrict access to these records. In these circumstances, differential recall of mental health services by highly distressed persons or of those who are currently suffering from ongoing depression may be minimized through shorter recall time intervals and carefully designed in person interviews. Among those who are all ready receiving mental health treatment from physicians, recall bias may be less of an issue due to similar recall ability. In general, recall bias may be minimized by selecting comparison groups with similar recall ability. However, recall bias is best eliminated through prospective study designs that employ administrative records.

Acknowledgements

References

Alegria M, Bijl R, Lin E, Walters E, Kessler R. Income differences in persons seeking outpatient treatment for mental disorders. A comparison of the United States with Ontario and the Netherlands. Archives of General Psychiatry 2000; 57, 383–91.

Allison P. Multinomial logit analysis. In Logistic Regression Using the SAS System: Theory and Application. 2 edn. Cary NC: SAS Institute, 2000, pp. 111–31.

Andrews G, Slade T. Interpreting scores on the Kessler Psychological Distress Scale (K10). Australian and New Zealand Journal of Psychiatry 25: 494–7.

Aneshensel C, Estrada A, Hansell M, Clark V. Social psychological aspects of reporting behavior: lifetime depressive episode reports. J Health Soc Behav 1987; 28: 232–46.

Austin P, Hux J (2002). A brief note on overlapping confidence intervals. Journal of Vascular Surgery 36: 194–5.

Beaudet M. Depression. Health Reports 1996; 7: 11–24.

Bijl R, De Graff R, Hiripi E, Kessler R, Kohn R, Offord D, Ustun T, Vicente B, Volleburgh W, Walters E, Wittchen H-U. The prevalence of treated and untreated mental disorders in five countries. Health Affairs 2003; 22, 122–33.

Bradburn N. Response Effects. In P Rossi, J Wright, A Anderson (eds) Handbook of Survey Research. New York: Academic Press, 1983, pp. 289–328.

Cairney J, Thorpe C, Rietschlin J, Avison W. 12-month prevalence of depression among single and married mothers in the 1994 National Population Health Survey. Canadian Journal of Public Health 1999; 90: 320–4.

CANMAT. Canadian Network for Mood and Anxiety Treatments: Guidelines for Diagnosis and Pharmacologic Treatment of Depression. Toronto, Ontario, 1999.

Chan B. Supply of Physicians' Services in Ontario (Rep. No. 1). Toronto, Ontario: Institute for Clinical Evaluative Sciences, 1999.

Chan BTB, Austin P. Patient, physician and community factors affecting referrals to specialists in Ontario, Canada. A Population-based, multi-level modelling approach. Medical Care 2003; 41: 500–11.

Ciok R. The Non-Telephone Population. Social Survey Methods Division, Statistics Canada; Ottawa, Canada: Statistics Canada, 1993.

Clark R, Ricketts S, McHugo G. Measuring hospital use without claims: a comparison of patient and provider reports. Health Services Research 1996; 31: 153–69.

Diverty B, Beaudet M. Depression. An undertreated disorder? Health Reports 1997; 8: 9–18.

Evans C, Crawford B. Patient self-reports in pharmacoeconomic studies. Their use and impact on study validity. Pharmacoeconomics 1999; 15: 241–56.

Furukawa T, Kessler R, Slade T, Andrews G. The performance of the K6 and K10 screening scales for psychological distress in the Australian National Survey of Mental Health and Well-Being. Psychological Medicine 2003; 33: 357–62.

Greenland S, Finkle WD. A critical look at methods for handling missing covariates in epidemiologic regression analyses. American Journal of Epidemiology 1995; 142: 1255–64.

Hennesy K, Reed S. Validating self-reports of mental health service use in a chronic population. Journal of Nervous and Mental Disorders 1992; 180: 399–400.

Horwath E, Johnson J, Klerman G, Weissman M. Depressive symptoms as relative and attributable risk factors for first-onset major depression. Arch Gen Psychiatry 1992; 49: 817–23.

Hosmer D, Lemeshow S. Applied Logistic Regression. New York: Wiley; 1989.

Judd L, Paulus M, Wells K, Rapaport M. Socieconomic burden of subsyndromal depressive symptoms and major depression in a sample of the general population. Am J Psychiatry 1996; 153: 1411–17.

Katz SJ, Kessler RC, Frank RG, Leaf P, Lin E. Mental health care use, morbidity, and socioeconomic status in the United States and Ontario. Inquiry 1997; 34: 38–49.

Kessler R, Andrews G, Colpe L, Hiripi E, Mroczek D, Normand S, Walters E, Zaslavsky A. Short screening scales to monitor population prevalences and trends in non-specific psychological distress. Psychological Medicine 2002; 32, 959–76.

Kessler R, Andrews G, Mroczek D, Ustun B, Wittchen H-U. The World Health Organization Composite International Diagnostic Interview – Short Form (CIDI–SF). International Journal of Methods in Psychiatric Research 2001; 7: 171–85.

Kessler R, Frank RG, Edlund M, Katz SJ, Lin E, Leaf P. Differences in the use of psychiatric outpatient services between the United States and Ontario. New England Journal of Medicine 1997; 336: 551–7.

Kessler R, Wittchen H-U, Abelson J, McGonagle K, Schwarz N, Kendler K, Knäuper B, Zhao S. Methodological studies of the Composite International Diagnostic Interview (CIDI) in the US National Comorbidity Survey (NCS). International Journal of Methods in Psychiatric Research 2001; 7: 33–55.

Koriat A, Goldsmith M, Plansky A. Towards a psychology of memory accuracy. Annu Rev Psychol 2000; 51; 481–537.

Leaf PJ, Bruce ML, Tischler GL, Freeman DHJ, Weissman MM, Myers JK. Factors affecting the utilization of specialty and general medical mental health services. Medical Care 1988; 26: 9–26.

Pignone M, Gaynes B, Rushton J, Burchell C, Orleans C, Multrow C, Lohr K. Screening for depression in adults: a summary of the evidence for the U.S. Preventive Services Task Force. Annals of Internal Medicine 2002; 136: 765–76.

Pincus H, Davis W, McQueen L. 'Subthreshold' mental disorders. A review and synthesis of studies on minor depression and other 'brand names'. Br J Psychiatry 1999; 288–96.

Rao J, Wu C, Yue K. Some recent work on resampling methods for complex surveys. Survey Methodology, Statistics Canada, Catalogue 1992; 18: 209–17.

Rhodes A, Lin E, Mustard C. Self-reported use of mental health services versus administrative records: should we care? International Journal of Methods in Psychiatric Research 2002; 11: 125–33.

Rothman K, Greenland S. Modern Epidemiology. 2 edn. Philadelphia: Lippincott-Raven; 1998.

Rust K, Rao J. Variance estimation for complex surveys using replication techniques. Stat Methods Med Res 1996; 5: 283–310.

Schwartz A, Perlman B, Paris M, Schmidt B, Thornton J. Psychiatric diagnosis as reported to Medicaid and as recorded in patient charts. American Journal of Public Health 1980; 70: 406–8.

Simon GE, VonKorff M. Recall of psychiatric history in cross-sectional surveys: implications for epidemiologic research. Epidemiol Rev 1995; 17: 221–7.

Statistics Canada. 1996/97 National Population Health Survey. Public Use Microdata File Documentation. Ottawa, Ontario: Minister of Supply and Services; 1998.

Towery OB, Sharfstein SS, Goldberg ID. The Mental and Nervous Disorder Utilization and Cost Survey: an analysis of insurance for mental disorders. American Journal of Psychiatry 1980; 137: 1065–70.

Vach W, Blettner M (1991). Biased estimation of the odds ratio in case-control studies due to the use of ad hoc methods of correcting for missing values for confounding variables. American Journal of Epidemiology 1991; 134: 895–907. Published erratum appears in Am J Epidemiol 1994; 140(1): 79.

Wells KB, Manning Jr, W, Duan N, Newhouse J, Ware Jr, J. Cost-sharing and the use of general medical physicians for outpatient mental health care. Health Services Research 1987; 22: 1–17.

Correspondence: Anne E Rhodes, Suicide Studies Unit, St Michael's Hospital, 30 Bond Street, 2 Shuter Wing, Suite 2010f, Toronto, Ontario, M5B 1W8.
Telephone (+1) 416 864-6060 x 2693.
Fax (+1) 416 864-5996.
Email rhodesa@smh.toronto.on.ca.