

Assessment of mental health and illness by telephone survey: Experience with an Alberta mental health survey

Scott B Patten, Carol E Adair, Jeanne V A Williams, Rollin Brant, Jian Li Wang, Ann Casebeer and Pierre Beauséjour

Abstract

Mental health is an emerging priority for health surveillance. It has not been determined that the existing data sources can adequately meet surveillance needs. The objective of this project was to explore the use of telephone surveys as a means of collecting supplementary surveillance information. A computer-assisted telephone interview was administered to 5,400 subjects in Alberta. The interview included a set of brief, validated measures for evaluating mental disorder prevalence and related variables. The individual subject response rate was 78%, but a substantial number of refusals occurred at the initial household contact. The age and sex distribution of the study sample differed from that of the provincial population prior to weighting. Prevalence proportions did not vary substantially across administrative health regions. There is a potential role for telephone data collection in mental health surveillance, but these results highlight some associated methodological challenges. They also draw into question the importance of regional variation in mental disorder prevalence—which might otherwise have been a key advantage of telephone survey methodologies.

Key words: mental health, methods, surveillance, telephone surveys

Introduction

Several features distinguish surveillance from other forms of health investigation. First, data collection is driven by a need for evidence rather than by research hypotheses.¹ Second, surveillance data are collected routinely or in an ongoing way, and data collection is integrated with analysis and interpretation, usually leading to the production of a surveillance product.² Chronic illnesses,³ including mental illnesses (www.who.int/whr/2001/en/), now rank among the most important public health issues. A need for enhanced chronic disease surveillance has been identified nationally⁴ and a lack of progress towards this goal

has received criticism at the national level (www.oag-bvg.gc.ca/domino/reports.nsf/html/20020902ce.html).

Mental illnesses may pose some particular challenges for surveillance. One challenge is the relative paucity of available data. The Public Health Agency of Canada's on-line chronic disease surveillance utility (www.oag-bvg.gc.ca/domino/reports.nsf/html/20020902ce.html) focuses on mortality, which does not address the most important sequelae of mental illness: impaired functioning and quality of life. The national health survey capable of region-level inference, the general health iterations of the Canadian Community Health Survey (www.statcan.ca/english/concepts/hs/index.htm), have only addressed two mental disorders, substance dependence and major depression, the latter only as optional content (www.statcan.ca/english/concepts/health/cycle3_1/pdf/cchs3documentation.pdf).

Improved mental health care might lead to increased service utilization because of increased accessibility or it might lead to diminished utilization because of improved population mental health. For these reasons, utilization statistics, such as physician billing data and hospital separations, provide incomplete information. Despite this, and perhaps because of the availability of such data to key stakeholders in government, they have assumed a preeminent role in mental health surveillance. One recent project concerned with the identification of mental health indicators (http://secure.cihi.ca/cihiweb/DispPage.jsp?cw_page=indicators_mental_e) resulted in a prototype report that consisted largely of data on hospital separation, inpatient hospital-days and length of stay. Similar analyses comprised the bulk of a recent Manitoba report.⁵

Efforts have been made to extend the scope of data available for mental health surveillance in Canada. Most notable in this regard is the 2002 Canadian Community Health Survey, Mental Health and Wellbeing (CCHS 1.2), which has provided a wealth of data about mental health and illness. Unfortunately, it

Author References

Scott B Patten, Carol E Adair, Jian Li Wang, Departments of Community Health Sciences and Psychiatry, Faculty of Medicine, University of Calgary, Calgary, Alberta, Canada

Jeanne VA Williams, Rollin Brant, Ann Casebeer, Department of Community Health Sciences, University of Calgary, Calgary, Alberta, Canada

Pierre Beauséjour, Department of Psychiatry, University of Alberta, Edmonton, Alberta, Canada

Correspondence: Scott B Patten, Department of Community Health Sciences, 3330 Hospital Drive NW, Calgary, Alberta, Canada T2N 4N1; fax: (403) 270-7307; e-mail: patten@ucalgary.ca

is not clear whether this type of survey will ever be repeated in the future, so its role as a source of surveillance data is uncertain. Furthermore, surveillance systems should be responsive to stakeholder needs¹ and it is not clear that large, specialized national surveys can be sufficiently responsive, for example, to needs manifesting at the health region level. The answer to this question depends at least partially on the extent to which key variables differ across geographical regions. If there is a high level of consistency across the country and a high level of consistency in the priorities of stakeholders, large-scale national projects may suffice.

Telephone surveys represent a potential mechanism for filling gaps in mental health surveillance and these methods have been employed in Canadian studies. Fournier explored the application of telephone survey methods in Quebec.^{6,7} In Saskatchewan, depression data deriving from the Saskatchewan Health and Dynamics Survey has recently been reported by D'Arcy.⁸ The Winnipeg Area Survey has generated prevalence data for several anxiety disorders.⁹⁻¹² In Calgary, the Calgary Health Region is currently carrying out a baseline telephone survey of generalized anxiety disorder in the region, building upon earlier studies of depression in Calgary and adjacent rural areas.¹³ Wild et al. used a telephone survey to evaluate the prevalence of alcohol-related problems and interest in self-help materials in Alberta.¹⁴ Computer-assisted telephone interviews (CATI) may be the only feasible method of data collection in geographically dispersed areas (see review of strengths and weaknesses of CATI by Choi).¹⁵

A project exploring the application of telephone survey methods for mental health surveillance in Alberta was initiated by the Alberta Mental Health Board in association with researchers from the University of Calgary. To guide the project, two consultative committees were initially formed. One of these was the Project Advisory Committee, consisting of decision makers from within the mental health care system. The other was the Technical Advisory Committee, which included epidemiologists, biostatisticians and health services researchers. Under the direction of the

investigators and the advisory committees, three projects were undertaken: 1) a content priorities survey, 2) a consensus workshop, and, 3) a population survey. This paper's objectives are to summarize our experience with the project, present some key findings and draw attention to several key methodological issues.

The content priorities survey targeted 110 key informants from the mental health system in Alberta. The target population included representatives of regional health authorities, senior program or service managers, government representatives, academic researchers and the Alberta Mental Health Board. Fifty eight of these subjects (52.7%) responded to the survey. Among the respondents were representatives of 16 of the 17 health regions in Alberta in 2003. The number of health regions in the province was subsequently reduced to nine in April 2003. The survey identified four priority areas for regional data collection: prevalence, service use, impact of disorders on functioning and quality of life. Of these priorities, only service-use data was routinely available to regional stakeholders.

The second project was a workshop to determine a course of action. This was held in Calgary on October 18, 2002. It was attended by 11 experts from government, health regions and academia, the investigative team and a professional facilitator. A plan for action was formulated which emphasized the potential value of using primary data collection to enhance the availability of data in the areas identified by the content priorities survey. It was decided that an initial survey should be conducted and that the target population for this should be the general household population in Alberta.

Methods

Sampling procedures

Alberta has a population of 2.97 million, dispersed over an area of 661,190 km². The population is 80% urban and 20% rural. The health care system is currently divided into nine health regions, with populations (within the 18–64 year age range targeted by this survey) ranging from 757,741 in

the Calgary Health Region to 45,824 in the Northern Lights Health Region. A map showing the Alberta health regions is available at: www.statcan.ca/english/concepts/health/cycle3_1/maps/alta_alb.pdf. Because of the vast geographical areas involved, telephone survey methods (rather than “face to face” interviews) were considered necessary. Since a high priority was placed by the stakeholders on region-specific estimates, a stratified sample was chosen. Precision-based sample size estimates determined that 95% confidence intervals of plus or minus 2% for attributes with a frequency of five percent could be achieved with a sample size of N = 600 per region. The target sample size was therefore set to N = 5,400.

Data collection was carried out by the population survey unit associated with the Quality Improvement and Health Information (QIHI) portfolio within the Calgary Health Region. A listing of provincial residential telephone numbers is maintained and updated by the survey unit. A random sample of these numbers was initially selected. Since unlisted numbers were not included in this database, there was concern that bias might be introduced in the event that households with unlisted numbers differed from those with listed numbers. The strategy of changing the final digit in the telephone number was adopted as a means of ensuring that non-listed numbers could also be included in the sampled list.¹⁶ A value of 1 was added to each of the randomly selected numbers. This “plus one” approach avoids the need to identify working “blocks” of telephone numbers encountered other random-digit dialing protocols,¹⁷ and avoids the clustering inherent in such sampling procedures. When a household was reached, the “next birthday” method¹⁸ was used to randomly select a single subject from the household.

Measures

The telephone interview administered to selected subjects included the Mini Neuropsychiatric Interview (MINI),¹⁹ which is a brief diagnostic interview. The MINI was developed jointly at the University of South Florida and the

National Institute for Mental Health in Paris. The MINI is not a single instrument, but rather a “family” of instruments that have been modified over time in response to modifications occurring to DSM criteria with the release of *DSM-IV*.²⁰ The MINI was originally developed for use in primary care, where it was felt that a brief structured interview could lead to improved detection by allowing non-physician clinical staff to derive preliminary psychiatric diagnoses.¹⁹ In keeping with the original goal of the MINI as a case-finding tool for primary care, the development process emphasized sensitivity over specificity. In community surveys, due to a lower base-rate, even high specificity might lead to overestimation of prevalence since a small false positive rate could give rise to a considerable proportion of false positive results in a sample where the vast majority of respondents do not have the condition being evaluated. Validation data for the MINI was originally reported in two European papers,^{21,22} and subsequently was summarized in a paper published in the *Journal of Clinical Psychiatry*.¹⁹ When compared to the SCID-P (a gold standard semi-structured interview) the MINI was found to be 96% sensitive and 88% specific. In the validation sample, a positive predictive value of 87% was achieved. However, in a community sample (which would probably have a lower point prevalence than the clinical validation sample) a lower predictive value might be expected. In the Paris validation study, the CIDI (long form) was used as a validation standard and 94% sensitivity and 79% specificity were reported. Estimates of the test-retest reliability of the MINI have ranged from 1.0¹⁹ to 0.64.²³

The general concept of a mental disorder in *DSM-IV* involves the requirement for a clinical syndrome but also a requirement that the syndrome be associated with significant distress, or an impact on functioning.²⁴ The latter requirement is intended to help to distinguish mental disorders from non-pathological expressions of depressive symptoms (e.g., an adjustment disorder resulting from stressful life events). One brief instrument resembling the MINI, called the Patient Health Questionnaire (PHQ), incorporates this concept of altered functioning by including an “interference”

item, which asks whether symptoms interfered with life activities, whatever these may be.²⁵ This approach addresses a concern raised by Narrow et al.²⁶ that inconsistencies in diagnoses arising from lay-administered psychiatric interviews often relate to indicators of clinical significance, such as interference with activities. The Narrow et al. study, which was a re-assessment of prevalence data from the Epidemiological Catchment Area studies in the United States²⁷ and the National Comorbidity Survey,²⁸ also used help-seeking behaviors as an index of clinical significance (e.g., taking medications, seeing a health professional). These characteristics were not used as clinical significance indicators in the current project. The concern that precluded their use was an effort to avoid the creation of tautological definitions of diagnosis in relation to some of the study’s key objectives—in particular, the goal of developing a way of monitoring health care use and treatment receipt in relation to diagnosis.

The interference item from the current study had the following wording: “How much did any of the problems we just talked about interfere with your life or activities?” The response options were read to the respondent: a lot, some, a little and not at all. The analysis by Narrow et al. used interference items from the NCS (“How much did your [symptom(s)] ever interfere with your life or activities—a lot, some, a little, or not at all?”) and ECA (“Did [symptom(s)] interfere with your life or activities a lot?”). The PHQ item is worded as follows: “... how difficult have these problems made it for you to do your work, take care of things at home, or get along with other people?”

The interview also included a generic quality of life instrument, the EuroQoL EQ-5D (www.euroqol.org/index.htm) and a measure of disability, the World Health Organization’s Disability Assessment Schedule (WHO DAS II).²⁹ In addition, items evaluating demographic variables and health care utilization were included. Interviewers working on the project were experienced telephone interviewers with the QIHI unit, and the data collection was preceded by a series of training sessions

incorporating both didactic instruction and practice.

Analysis

There are a variety of determinants of selection probabilities in telephone surveys: the number of telephone lines reaching each household and the number of household residents. An initial sampling weight was calculated by dividing the number of household residents by the number of voice-telephone lines into the household. The demographic characteristics of the sample were then compared to the demographic distribution of the target population within each health region. Data for the 18–64 age group (reflecting the survey’s eligibility criteria) from the Alberta Health Care Insurance Plan Stakeholder Registry was used for this purpose. This is a database of registrants in the provincial health care insurance plan. The relevant tables were provided by the Health Surveillance Branch of Alberta Health and Wellness (www.health.gov.ab.ca). The telephone sample was found to overrepresent women, especially those over the age of 35, and to underrepresent men, especially those under the age of 50. For these reasons, an adjustment was made to the sampling weights so that the results would closely approximate the age and sex distribution of the underlying population. Another set of sampling weights were calculated on a region-by-region basis and these sets of weights were used for making region-specific estimates. To account for the regionally stratified sampling, the weights used in the provincial estimates were also multiplied by the inverse of the stratified sampling probability, with the sampling probability being defined as the regional sample size (N = 600) divided by the population of the region. All analyses used the survey (“svy”) commands in Stata, version 8.0.³⁰

Results

In total, 29,941 telephone numbers were dialed. Of these, 6,121 were not working numbers, 3,048 were businesses and 2,525 were fax machines. There were 143 blocked calls and in 1,453 cases it was not possible to get past answering machines or

TABLE 1
Prevalence proportion estimates and 95% confidence intervals of MINI*-defined major depression (current) in survey respondents, with clinical interference thresholds (N=422)

Interference threshold response	N	Unweighted estimates		Weighted estimates	
		%	95% CI	%	95% CI
No interference requirement	422	7.8	7.1-8.6	7.7	6.9-8.4
"A little" interference	384	7.1	6.4-7.8	7.0	6.3-7.7
"Some" interference	297	5.5	4.9-6.1	5.3	4.7-5.9
"A lot" of interference	182	3.4	2.9-3.9	3.2	2.7-3.7

* MINI – the MINI Neuropsychiatric Interview

voice mail. A small number (N = 101) were busy each time called and 1,701 numbers rang but were never answered. At the time that the study was terminated, there were 376 numbers in the process of follow-up, which included 148 households. In 2,644 households, there were no eligible residents and in 377 it was not possible to successfully communicate with a householder because of a language barrier. In total, 11,680 numbers contacted an eligible household.

There were 4,512 household-level refusals (39% of households contacted), so that 7,168 subjects were selected for inclusion. There were 146 subjects who were never contacted because they were not at home during any of the call-back attempts and 82 who were not contacted because the study ended while they were in process. There were 6,940 individual subjects who were contacted in person by an interviewer and asked to provide assent for participation. Of these, 95 did not speak English, there were 1,314 refusals and 121 interrupted interviews, so that interviews were complete in 5,410 instances. Complete data were collected from 5,383 of these subjects. If the individual response rate is calculated as the proportion of eligible subjects who did not refuse or interrupt the interview, the response rate was 78%. If those who were selected but never contacted are counted as individual non-responders, the response rate was 75%.

The unweighted mean age of the sample was 40.8 years. The weighted mean age was 39.2 years. The unweighted sample included 2,087 men (38.8%) and 3,296 women (61.2%). The weighted proportions of men and women were 47.8% and 52.2%, respectively. The marital status

of 65.6% of subjects was married, 11.7% were divorced, widowed or separated and 22.6% had never married. Most (60.2%) of the subjects had more than secondary-level education (grade 12 graduation), 36.3% reported a secondary-level education and only 3.5% had less than a secondary-level education.

The direct application of the MINI diagnostic algorithms (without the interference item) tended to produce prevalence estimates higher than what would be predicted based on the literature. Addition of the interference item in the diagnostic algorithms had the expected effect of reducing the estimated prevalence for each disorder. Table 1 shows weighted and unweighted estimates for one disorder, major depres-

sion, with application of several interference thresholds. Estimated prevalence predictably diminished as the requirement for interference with activities was made more stringent. Since the MINI detects current major depressive episodes, the requirement for "a lot" of interference with current activities might be seen as a suitable requirement for identification of clinically significant cases. Consistent with this idea, the prevalence associated with this threshold was 3.2%. This estimate is slightly higher than that of the recent Canadian National Survey of Mental Health and Well-being (CCHS 1.2), which placed the thirty-day prevalence at 1.8%³¹ but lower than the 4.9% point prevalence reported from the National Comorbidity Survey in the United States.³² The ECA study reported a 2.2% thirty-day prevalence³³ and the Edmonton study reported a 2.3% estimate.³⁴ The weighted estimates tended to be slightly lower than the unweighted estimates for major depression, probably because of the female preponderance in the unweighted data. However, the differences were small.

A similar relationship was observed when the interference criterion was applied to other disorders. For example, the twelve-month prevalence of alcohol dependence

TABLE 2
Prevalence proportions and 95% confidence intervals of mood disorders in survey respondents†, by Alberta provincial health region

Health region	#	Prevalence proportions and 95% confidence intervals					
		Major depressive episode* (14 day)		Dysthymia** (2 year)		Any mood disorder*	
		%	CI	%	CI	%	CI
Chinook	1	2.7	(1.2-4.2)	0.9	(0.1-1.8)	4.8	(2.8-6.8)
Palliser	2	3.4	(1.8-5.0)	0.9	(0-1.7)	5.0	(3.1-6.9)
Calgary	3	3.7	(2.0-5.4)	0.9	(0.1-1.6)	5.7	(3.6-7.7)
David Thompson	4	2.3	(1.1-3.5)	1.1	(0.2-2.0)	3.9	(2.3-5.5)
East Central	5	4.2	(1.4-6.9)	1.6	(0.7-2.5)	6.3	(3.3-9.3)
Capital	6	3.1	(1.5-4.8)	1.4	(0.4-2.4)	6.7	(4.4-9.1)
Aspen	7	6.8	(3.4-10.2)	0.7	(0-1.5)	8.3	(4.8-11.7)
Peace Country	8	2.1	(1.0-3.1)	1.5	(0-2.9)	3.8	(2.3-5.3)
Northern Lights	9	3.0	(1.4-4.7)	1.1	(0.2-2.1)	4.9	(2.9-6.9)
Alberta (total)		3.2	(2.7-3.7)	1.2	(0.9-1.5)	5.4***	(4.8-6.1)

† Respondents were assessed using the Mini Neuropsychiatric Interview.

* Clinical significance component of the prevalence definition required "a lot" of impairment.

** Clinical significance component of the prevalence definitions required any reported impairment.

*** Includes bipolar disorders.

TABLE 3
Prevalence proportions and 95% confidence intervals of substance-use disorders in survey respondents*, by Alberta provincial health region

Health region	#	Prevalence proportions (95% confidence intervals)				
		Alcohol		Other substances		Any substance-use disorder
		Dependence	Abuse	Dependence	Abuse	
		(12 month)	(12 month)	(12 month)	(12 month)	(12 month)
Chinook	1	1.0%	2.1%	2.1%	1.6%	6.2%
		(0-2.0)	(0.8-3.7)	(0.5-3.5)	(0.1-3.0)	(3.6-8.7)
Palliser	2	3.6%	4.2%	1.2%	1.6%	9.2%
		(1.7-5.5)	(1.7-6.7)	(0-2.6)	(0-3.2)	(5.9-12.5)
Calgary	3	3.8%	1.9%	2.4%	1.6%	7.9%
		(2.1-5.5)	(0.1-3.0)	(1.0-3.9)	(0.6-2.6)	(5.5-10.4)
David Thompson	4	4.8%	4.6%	3.0%	2.5%	11.7%
		(0.2-9.5)	(0.3-8.8)	(0-7.6)	(0.8-4.2)	(5.6-17.8)
East Central	5	6.6%	1.2%	1.5%	1.9%	9.5%
		(2.5-10.7)	(0.4-2.1)	(0-3.1)	(0.2-3.6)	(5.3-13.8)
Capital	6	3.1%	1.9%	1.6%	1.6%	6.6%
		(1.7-4.5)	(0.4-3.4)	(0.7-2.6)	(0.2-3.0)	(4.4-8.9)
Aspen	7	4.6%	1.4%	1.4%	1.8%	7.7%
		(2.6-6.6)	(0.3-2.5)	(0.2-2.6)	(0-3.8)	(4.8-10.7)
Peace Country	8	3.9%	1.2%	1.9%	1.7%	6.7%
		(1.8-6.0)	(0-2.5%)	(0.4-3.5)	(0.1-3.2)	(3.9-9.6)
Northern Lights	9	4.6%	1.2%	1.4%	0.3%	6.4%
		(2.5-6.7)	(0.5-2.2)	(0.2-2.6)	(0-0.7)	(3.9-8.9)
Alberta (total)	–	3.6%	1.9%	1.6%	1.5%	7.1%
		(3.0-4.1)	(1.5-2.4)	(1.2-2.0)	(1.1-1.8)	(6.4-7.9)

* Respondents were assessed using the Mini Neuropsychiatric Interview.

according to the MINI was 5.2%, which is consistent with some published estimates (see discussion below). When any reported interference was incorporated into the definition, the prevalence became 3.6% and when “a lot” of interference was required, the prevalence was 0.9%. For dysthymia, the study’s prevalence estimates were most consistent with the literature when “some” interference was required. Adoption of this threshold seems justifiable in the sense that a lower level of interference is expected from dysthymia than from current major depression, yet some amount of interference with functioning seems reasonable as a means of ensuring that the cases identified were of clinical significance.

Prevalence proportion data from the MINI interview for mood disorders, substance-use disorders and anxiety disorders are presented in Tables 2–4. These specific estimates reflect the application of the interference criterion as described above: a lot of

interference for major depression, at least some interference for dysthymia and no requirements for substance-use disorders. For reasons similar to those for dysthymia, anxiety disorders were required to be associated with at least some interference with functioning in order to be considered clinically significant. The MINI produces period prevalence estimates covering variable periods of time. For major depression, past fourteen-day prevalence was assessed; for dysthymia the prevalence period was the past two years. For anxiety and substance-use disorders, the MINI produces twelve-month period prevalence estimates. None of these tabulations provided evidence of regional variations in mood disorder prevalence. The final row of the Tables 2–4, therefore, presents the provincial prevalence proportions estimates, which seem preferable for reasons of precision.

The MINI evaluates current major depressive episodes and current (past 30 days) manic or hypomanic episodes. However,

not unexpectedly (as the sample size was calculated based on a five percent attribute frequency), the prevalence proportions of current hypomanic episodes (0.8%; 95% CI: 0.5-1.0) and manic episodes (0.5%; 95% CI: 0.3-0.7) were too low for the data to support region-specific estimation.

Generalized anxiety disorder is not included as a separate column in Table 4 because the prevalence proportions resulting from the MINI syndromal definition and this level of clinical significance resulted in an unrealistically high prevalence proportion estimate, 9.6%. When “a lot” of interference with functioning was required, the prevalence dropped to 4.2% (95% CI: 3.6-4.7).

The MINI includes a diagnostic module for eating disorders. In this survey no subjects met diagnostic criteria for anorexia nervosa. The prevalence proportion of bulimia nervosa was 1.2%. With the addition of the clinical significance criteria requiring interference, the estimated prevalence of bulimia nervosa was 0.9% (95% CI: 0.7-1.2). The odds ratio for the female sex in bulimia was 3.2 (95% CI: 1.6-6.3).

While regional variations in prevalence were not identified in this analysis, sizable differences in relation to age and sex were observed. Figure 1 shows the prevalence proportions of major disorder categories by sex. As expected mood and anxiety disorders were more common in women and substance-use disorders were more common in men. Figure 2 presents the same prevalence proportions estimates stratified by age. The highest prevalence proportions of mood and anxiety disorders occurred in the 26–44 year age group, whereas substance-use disorders were most frequent in the 18–25 age group. No significant associations were observed between prevalence of these disorders and level of education.

Mood disorders were more common in divorced, widowed or separated subjects (9.7%, 95% CI: 7.4-12.0) than in married (including common-law) subjects (4.5%, 95% CI: 3.8-5.2) or never-married subjects (5.7%, 95% CI: 4.3-7.1). A similar pattern was observed for anxiety disorders. For

TABLE 4
Prevalence proportions and 95% confidence intervals of anxiety disorders in survey respondents†, by Alberta provincial health regions

Health region	#	Prevalence proportions (95% confidence intervals)			
		Panic disorder (Lifetime)	Agoraphobia* (Current)	Social phobia (30 day)	Any anxiety disorder**
Chinook	1	1.9% (0.2-1.6)	1.6% (0.4-2.8)	1.8% (0.5-3.2)	4.3% (2.4-6.1)
Palliser	2	1.3% (0.3-2.3)	1.5% (0.4-2.5)	2.3% (0.9-3.7)	4.4% (2.6-6.2)
Calgary	3	1.0% (0.1-1.8)	2.1% (0.7-3.5)	2.6% (1.1-4.0)	5.1% (3.2-6.9)
David Thompson	4	1.1% (0.4-1.9)	1.2% (0.4-2.0)	1.5% (0.4-2.5)	4.7% (2.9-6.4)
East Central	5	1.3% (0.4-2.2)	1.3% (0.4-2.1)	2.8% (0.1-5.4)	5.9% (3.0-8.8)
Capital	6	0.9% (0.2-1.6)	2.2% (0.8-3.5)	2.6% (1.4-3.9)	6.0% (3.9-8.1)
Aspen	7	2.9% (0.4-5.3)	1.0% (0.3-1.7)	2.5% (0.4-4.6)	7.5% (4.1-10.9)
Peace Country	8	1.4% (0.5-2.3)	1.1% (0.2-2.0)	2.1% (0.9-3.3)	3.7% (2.2-5.1)
Northern Lights	9	0.7% (0-1.3)	1.7% (0.7-2.8)	1.9% (0.8-2.9)	4.4% (2.5-6.3)
Alberta (total)	–	1.4% (1.0-1.7)	1.5% (1.2-1.9)	2.2% (1.8-2.7)	5.1% (4.5-5.6)

† Respondents were assessed using the Mini Neuropsychiatric Interview.

* Includes two MINI categories: panic disorder with agoraphobia and agoraphobia with panic symptoms but no panic disorder, associated with at least “some” impairment in functioning.

** Category includes generalized anxiety disorder. Clinical significance criteria required “a lot” in interference with function.

substance-use disorders, a gradient was observed, with 4.3% (95% CI: 3.6-5.0) of married subjects, 8.2% (95% CI: 6.1-10.4) of divorced, widowed or separated subjects and 14.7% (95% CI: 12.4-17.1) of never-married subjects having one of these disorders. The possibility of confounding by age, however, is suggested by Figure 2. The high prevalence of substance-use disorders in never-married subjects could be accounted for by the younger age of these subjects. With married subjects comprising the baseline group, the odds ratio for never-married subjects was 3.9 and for divorced, widowed or separated subjects, it was 2.0. After adjustment for age using logistic regression, both categories continued to have an elevated prevalence of substance-use disorder, but the difference between them disappeared. The age adjusted odds ratio was 2.3 (95% CI: 1.7-3.0) for never-married subjects and 2.4

(95% CI: 1.7-3.4) for divorced, widowed and separated subjects.

The disability and quality of life results were most striking in relation to comor-

bidity. The WHO DAS II produces a variety of scaled outputs that are consistent with the WHO ICF classification system (www3.who.int/icf/icftemplate.cfm), which is based on a division of body structures, body functions, participation and environment. Figure 3 presents the proportion reporting any problems in the function “understanding and communicating”, and a participation scaling, “participation in society”. As the number of MINI conditions increased, the proportions reporting no deficits decreased. One output of the EuroQol is an analogue scale of globally perceived health, called the “health thermometer”. With increasing comorbidity, these global perceptions of health declined substantially (see Figure 4).

Discussion

Any attempt to use telephone surveys for mental health surveillance will encounter certain challenges, most prominently involving measurement and subject selection. Within the domain of measurement, one problem is that although several brief instruments have been developed for various mental disorders,^{25,35,36} these have rarely been validated in general population samples. A partial exception is the CIDI Short Form, which was developed using data collected during the National Comorbidity Survey.³⁷ However, concerns have been raised about the validity of this instrument.^{38,39} With all of these instruments, the specificities reported by their validation studies raise concern about over-

FIGURE 1
Sex distribution of prevalence proportions (%) of major disorder diagnostic categories* in survey respondents

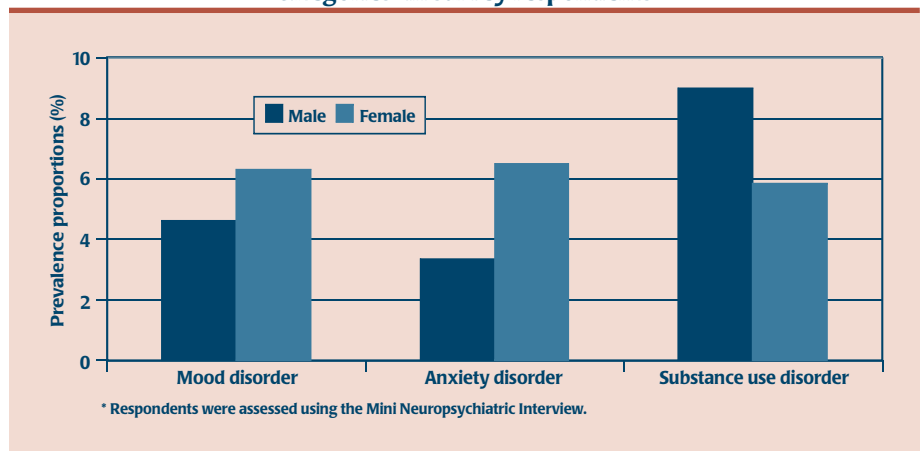
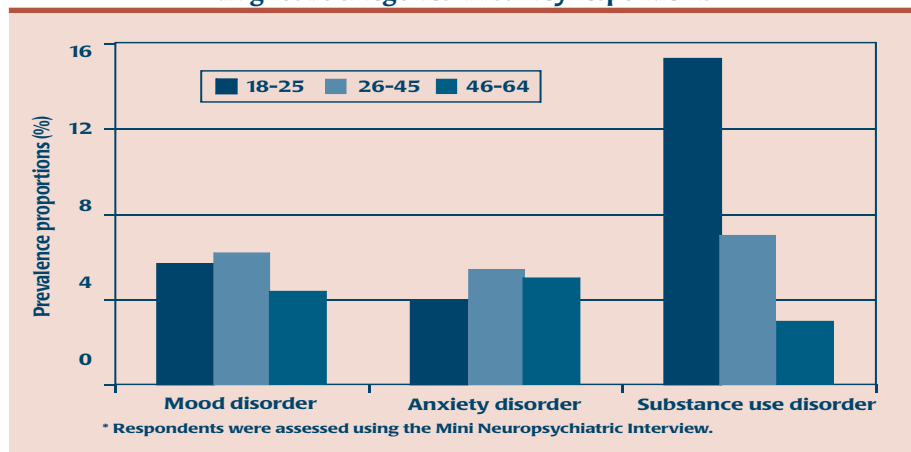


FIGURE 2
Age distribution of prevalence proportions (%) of major disorder diagnostic categories* in survey respondents



estimation of prevalence in general population surveys. With the CIDI Short Form for major depression, this has seemed to be the case.⁴⁰ There is only one population-based study that used the Patient Health Questionnaire (major depression section) and this found a prevalence proportion of 3.8% for current major depression.⁴¹

The procedure employed in this study to identify potential false positive diagnostic results was the application of an interference item similar to that used by the Patient Health Questionnaire. Imposing a requirement for higher levels of interference resulted in lower prevalence estimates, many of which were consistent with the existing literature. In the following paragraphs, the estimates deriving from this study are compared to published estimates and reviews in the literature. However, it should be noted that the augmentation of the MINI diagnostic algorithms by the addition of an interference criterion was an *ad hoc* procedure in the sense that it was suspected, but not known, before the data were analyzed that such a procedure would be necessary. As such, it cannot be claimed that the procedure for adjusting the prevalence estimates is replicable. However, the general approach makes sense both in terms of its consistency with the *DSM-IV* approach to diagnosis and the analysis of inconsistencies in survey output reported by Narrow et al.²⁶

In the case of depressive disorders, a review of prevalence studies by Wittchen et al.⁴²

reported a median and range of point prevalence proportion estimates from published studies as 3.1% (range: 1.5-4.5), which is consistent with the estimate reported here. The Australian National Mental Health Survey, using ICD-10 criteria, reported a thirty-day prevalence proportion nearly identical to that reported here: 3.3%.⁴³ A recent systematic review by Waraich et al.⁴⁴ arrived at a “best estimate” of annual dysthymia prevalence of 2.0% (95% CI: 1.3-2.8), which is slightly higher but not inconsistent with the 1.2% (95% CI: 0.9-1.5) identified in this study. The Australian National Mental Health Survey reported a 1.1% one-month and 1.3% twelve-month prevalence proportion of ICD-10 dysthymia. Narrow et al.²⁶ produced revised twelve-month prevalence proportion estimates of 1.6% from two large American epidemio-

logical surveys.²⁶ The Narrow et al. estimate for the twelve-month prevalence proportion of any substance-use disorder was 6.0% (compared to 7.1% in this survey). Narrow et al. reported that alcohol use disorders (5.2%) were much more common than other substance-use disorders (1.7%), which is the same pattern seen in this study. However, the Canadian Mental Health and Well-being Survey reported a lower (national) twelve-month prevalence proportion of alcohol dependence and drug dependence, 2.6% and 0.7%, respectively.⁴⁵ Finally, the overall twelve-month prevalence proportion of alcohol or substance-use disorders observed in this survey resembles a “best estimate” reported in a recent structured review of prevalence studies.⁴⁶ A puzzling feature of the substance-use disorder prevalence estimates is that the prevalence proportion of dependence exceeded that of abuse. This probably results from the structure of the MINI, which skips the abuse questions when criteria for dependence are met. A variable prevalence of panic disorder has been reported in previous studies, with the most notable discrepancy being a higher prevalence proportion (3.5%) in the National Comorbidity Survey²⁸ than in earlier studies. The lifetime prevalence proportion of panic disorder reported here is consistent with international estimates deriving from studies using methods comparable to those of the Epidemiologic Catchment Area studies, 1.5%⁴⁷ and the Alberta estimate from the Canadian Mental Health and Well-

FIGURE 3
Proportion of survey respondents reporting no impairment on two WHO DAS II scalings, “Understanding and communicating” and “Participation in society”, by MINI comorbidity status

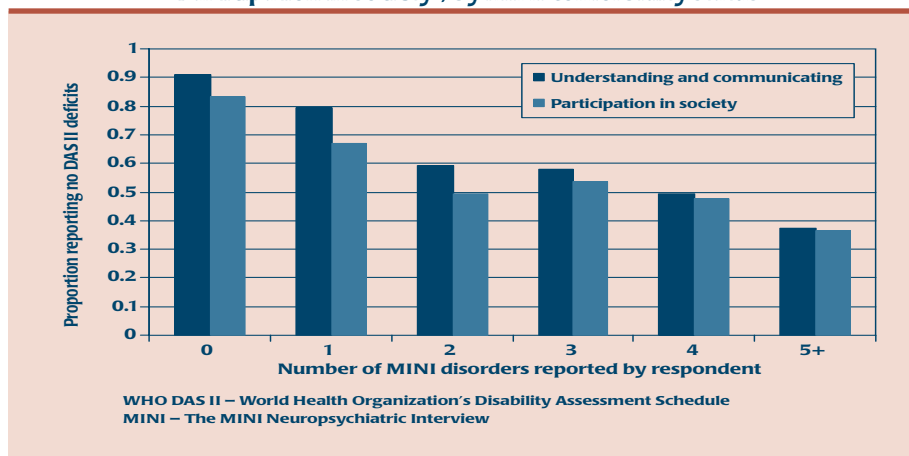
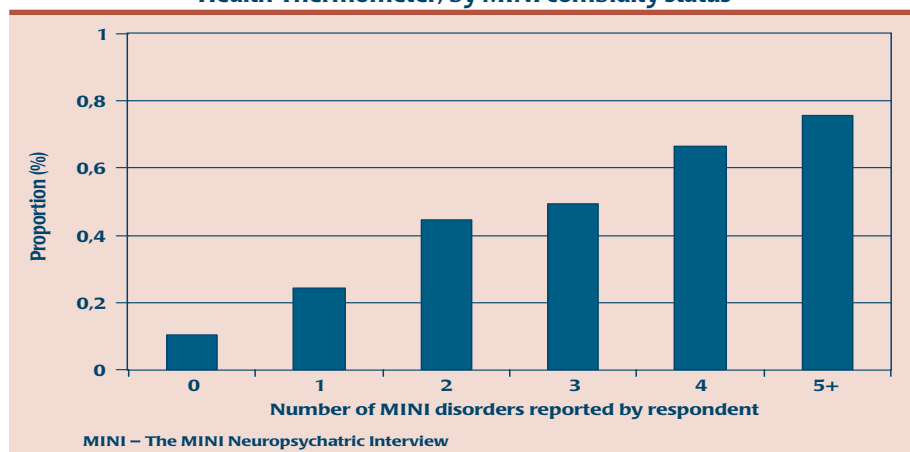


FIGURE 4
Proportion of survey respondents scoring < 70/100 on the EuroQoL EQ-5D Health Thermometer, by MINI comorbidity status



being Survey (1.7%),⁴⁵ noting that the latter estimate was for twelve-month prevalence. The Narrow et al. revised estimate for the twelve-month prevalence of panic disorder was 1.4%.²⁶ A review by Kessler and Wittchen⁴⁸ placed the twelve-month prevalence of generalized anxiety disorder at three to five percent. Few studies have reported the thirty-day prevalence of social phobia, which is the estimate that the MINI produces, most reporting lifetime or twelve-month period prevalence.⁴⁹ For example, the Canadian Mental Health and Well-being survey reported a twelve-month Alberta prevalence proportion of social phobia of 3.1%,⁴⁵ which seems consistent with the 2.2% reported here. A point estimate was reported by the Australian National Mental Health Survey, using ICD-10 criteria: 1.4%.⁴³

Another set of challenges to the use of telephone survey methods in mental health surveillance involves subject selection. The target population for this study consisted of household residents and can be expected to underrepresent institutionalized and homeless populations. In order to provide comprehensive coverage of the population, parallel sampling strategies, or alternative surveillance procedures for these populations would need to be developed. The National Population Health Survey conducted by Statistics Canada, for example, initiated an institutional survey that runs in parallel to the household residents' survey⁵⁰ in order to deal with this issue. Broader concerns related to the

validity of random digit dialing as a sampling strategy for mental health surveys cannot be fully addressed by this project. A large proportion of telephone numbers sampled could not be reached and there was an appreciable extent of household non-response. The exact proportions in these various categories could not be determined because the availability of call-screening technology may have resulted in disinterested household refusals not being reached or not answering their phones. If factors related to willingness to participate in the survey, either at the household or residential level, are also related to mental health status, then selection bias may have been introduced.

The approach taken in this project to the evaluation of clinical significance differs from that usually adopted in psychiatric epidemiology. More lengthy structured diagnostic interviews, such as the Composite International Diagnostic Interview (CIDI),⁵¹ typically assess clinical significance at the level of specific syndromes or even symptoms, incorporating probes (very similar to those used in this survey enquiring about help-seeking, medication taking and interference) to evaluate the clinical significance of syndromes or symptoms. However, in the current study, clinical significance probes were administered in an omnibus fashion, enquiring about interference using a single set of items covering all reported syndromes and symptoms. This allowed the interview to be quite brief (the total interview took, on average,

approximately 20 minutes, which is considered a reasonable limit for maintenance of response rates in telephone surveys), yet preserved the ability to evaluate probable clinical significance of the syndromes identified. This approach, however, does not allow confirmation of clinical significance for each disorder separately when comorbid disorders are present.

A survey of this size requires the investment of considerable resources. However, the use of telephone survey methods kept total costs down to approximately \$40 per subject. This total cost included not only direct interviewing costs, but also associated infrastructural and start-up costs. Geographical sampling procedures, even those involving sampling clusters in multiple stages, would have been considerably more expensive. On the other hand, the increasing use of call-screening and voice messaging may lead to higher costs for telephone surveys in the future by requiring more frequent follow-up calls to identify and interview selected households and subjects. These same factors may lead, in the future, to unacceptably low response rates or bias. The use of telephone survey methods, as opposed to the "face to face" interviewing that is generally employed in epidemiological surveys, may make it more feasible to repeat surveys of this type periodically, an essential element of surveillance.

Periodic monitoring of mental health across time in the same population could provide a useful set of population mental health indicators. This would enhance the capacity of decision makers to attach measurable goals to their policy decisions and to better target services towards recognizable needs within the population. However, it cannot be convincingly argued that telephone surveys are the preferred method for accomplishing these goals. The lack of regional differences found in this study might suggest that provincial (or, by extension, even national) samples may be sufficient to achieve the surveillance goals identified by the stakeholders. With the various unresolved measurement issues, future work will need to explore procedures for gaining valid estimates with brief measures. The challenges involved

should not be underestimated since the extent of agreement even between more lengthy structured interviews and clinical interviews is not generally impressive.⁵²⁻⁵⁴ Furthermore, if a suite of brief, valid measures can be identified, it may increase the scope of what can be feasibly accomplished by large national efforts such as the CCHS. Similarly, the challenges associated with telephone sampling are likely to intensify with advancing technology and heightened concerns about privacy. Despite these uncertainties, as Canada's national approach to chronic disease surveillance is taking shape, it seems reasonable that telephone survey methods should remain "on the table" as potential contributors to mental health surveillance.

Acknowledgements

The Calgary portion of this survey was supported by the Calgary Health Region through its biannual mental health research funding competition. Data collection in the remainder of the province of Alberta was supported by the Alberta Mental Health Board. Developmental and pilot studies were funded by the Health Research Fund, which is administered by the Alberta Heritage Foundation for Medical Research. A consensus workshop that contributed to the planning of the survey was funded by the Institute of Health Economics. The authors would like to thank the members of Project Advisory and Technical Advisory Committees, and additional participants in the Consensus Workshop: Don Addington, Yvonne Collinson, Elliot Goldner, Rob James, Gerald McDougall, Marlene Reimer, Don Schopflocher, Doug Watson, Cam Wild, Di Vosburgh, Sujaya Parthasarathy and Tom Noseworthy.

References

1. Choi BC. Perspectives on epidemiological surveillance in the 21st century. *Chron Dis Canada* 1998;19:145-51.
2. Mowat D. Health Surveillance in Canada. *Chron Dis Can* 1998;19:143-4.
3. Murray CJL, Lopez AC. *Global Burden of Disease and Injury*. Boston: Harvard School of Public Health; 1996.

4. Chronic Non-communicable Disease Infrastructure Sub-group for the Health Surveillance Working Group. National surveillance networks for chronic disease in Canada. Charting a path forward. 2001. Ref Type: Unpublished Work
5. Martens P, Fransoo R, McKeen N, Burland E, Jebamani L, Burchill C, et al. Patterns of regional mental illness disorder diagnoses and service use in Manitoba: a population-based study. 2004.
6. Fournier L, Kovess V. A comparison of mail and telephone interview strategies for mental health surveys. *Can J Psychiatry* 1993;38:525-33.
7. Fournier L, Lesage AD, Toupin J, Cyr M. Telephone surveys as an alternative for estimating prevalence of mental disorders and service utilization: a Montreal catchment area study. *Can J Psychiatry* 1997;42:737-43.
8. D'Arcy C, Kosteniuk J, Smith P, Nilson R, Cholowsky M, Bowen R, et al. Depression in Saskatchewan: An analysis of the Saskatchewan Population Health and Dynamics Survey 1999-2000. *Applied Research/Psychiatry*; 2004.
9. Stein MB, Torgrud LJ, Walker JR. Social phobia symptoms, subtypes, and severity: findings from a community survey. *Arch Gen Psychiatry* 2000;57:1046-52.
10. Stein MB, Forde DR, Anderson G, Walker JR. Obsessive-compulsive disorder in the community: an epidemiological survey with clinical reappraisal. *Am J Psychiatry* 1997;154(8):1120-6.
11. Stein MB, Walker JR, Forde DR. Gender differences in susceptibility to posttraumatic stress disorder. *Behav Res Ther* 2000;38:619-28.
12. Stein MB, Walker JR, Forde DR. Public-speaking fears in a community sample. Prevalence, impact on functioning, and diagnostic classification. *Arch Gen Psychiatry* 1996;53:169-74.
13. Patten SB, Stuart HL, Russell ML, Maxwell CJ, Arboleda-Florez J. Epidemiology of depression in a predominantly rural health region. *Soc Psychiatry Psychiatr Epidemiol* 2003;38:360-5.
14. Wild TC, Roberts AB, Cunningham J, Schopflocher D, Pazderka-Robinson H. Alcohol problems and interest in self-help: A population study of Alberta adults. *Can J Pub Hlth* 2004;95:127-32.
15. Choi BCK. Computer assisted telephone interviewing (CATI) for health surveys in public health surveillance: methodological issues and challenges ahead. *Chronic Diseases in Canada* 2004;25:21-7.
16. Mullet GM. The efficacy of plus-one dialing: Self reported status. *American Statistical Association: Proceedings of the Section on Survey Research Methods* 1982;575-6.
17. Potthoff RF. Telephone sampling in epidemiologic research: to reap the benefits, avoid the pitfalls. *American Journal of Epidemiology* 139(10):967-78, 1994 May 15.
18. Watson EK, Firman DW, Heywood A, Hauquitz AC, Ring I. Conducting regional health surveys using a computer-assisted telephone interviewing method. *Australian Journal of Public Health* 1995;19:508-11.
19. Sheehan DV, LeCubier Y, Sheehan H, Amorim P, Janavs J, Weiller E, et al. The Mini-International Neuropsychiatric Interview (M.I.N.I.): The development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *J Clin Psychiatry* 1998;59(Suppl 20):22-33.
20. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders (DSM-IV-TR)*. 4th ed. Washington: American Psychiatric Association; 2002.
21. Lecubier Y, Sheehan DV, Weiller E, Amorim P, Bonora I, Sheehan KH, et al. The MINI International Neuropsychiatric Interview (MINI). A short diagnostic structured interview: reliability and validity according to the CID-I. *European Psychiatry* 1997;12:224-31.
22. Sheehan DV, Lecubier Y, Sheehan KH, Janavs J, Weiller E, Keskiner A, et al. The validity of the Mini International Neuropsychiatric Interview (MINI) according to the SCID-P and its validity. *Eur Psychiatry* 1997;12:232-41.

23. Balázs J, LeCrubier Y, Csiszér N, Koszták J, Bitter I. Prevalence and comorbidity of affective disorders in persons making suicide attempts in Hungary: importance of the first depressive episodes and of bipolar II diagnoses. *J Affect Disord* 2003;76:113–9.
24. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders (DSM-IV-TR)*. Washington: American Psychiatric Association; 2000.
25. Spitzer RL, Kroenke K, Williams JBW, The Patient Health Questionnaire Primary Care Study Group. Validation and utility of a self-report version of the PRIME-MD. The PHQ Primary Care Study. *JAMA* 1999;282:1737–44.
26. Narrow WE, Rae DS, Robins LN, Regier DA. Revised prevalence estimates of mental disorders in the United States: using a clinical significance criterion to reconcile 2 surveys' estimates. *Arch Gen Psychiatry* 2002;59:115–23.
27. Regier DA, Myers JK, Kramer M, Robins LN, Blazer DG, Hough RL, et al. The NIMH Epidemiologic Catchment Area program. Historical context, major objectives, and study population characteristics. *Arch Gen Psychiatry* 1984;41:934–41.
28. Kessler RC, McGonagle KA, Zhao S, Nelson CB, Hughes M, Eshleman S, et al. Lifetime and 12-month prevalence of DSM-III-R psychiatric disorders in the United States. Results from the National Comorbidity Survey. *Arch Gen Psychiatry* 1994;51:8–19.
29. Epping-Jordan J. *The WHODAS II*. Geneva, Switzerland: The World Health Organization; 2000.
30. Stata, version 8.0 [computer program]. Version 8.0. College Station, TX: Stata Corporation; 2003.
31. Patten SB, Wang JL, Williams JVA, Currie SR, Beck CA, Maxwell CJ, et al. Descriptive epidemiology of major depression in Canada. *Can J Psychiatry* 2006;51:84–90.
32. Blazer DG, Kessler RC, McGonagle KA, Swartz MS. The prevalence and distribution of Major Depression in a national community sample: the National Comorbidity Survey. *Am J Psychiatry* 1994;151:979–86.
33. Regier DA, Boyd JH, Burke JD, Myers JK, Kramer M, Robins LN, et al. One-month prevalence of mental disorders in the United States. Based on five epidemiological catchment area sites. *Arch Gen Psychiatry* 1988;45:977–86.
34. Bland RC, Newman SC, Orn H. Period prevalence of psychiatric disorders in Edmonton. *Acta Psychiatr Scand* 1988;Suppl 338:33–42.
35. Diez-Quavedo C, Rangil T, Sanchez-Planell L, Kroenke K, Spitzer RL. Validation and utility of the Patient Health Questionnaire in Diagnosing Mental Disorders in 1003 general hospital Spanish inpatients. *Psychosom Med* 2001;63:679–86.
36. Spitzer RL, Williams JBW, Kroenke K, Hornyak R, McMurray J, Patient Health Questionnaire Obstetric-Gynecology Study Group. Validity and utility of the PRIME-MD Patient Health Questionnaire in assessment of 3000 obstetric-gynecologic patients: The PRIME-MD Patient Health Questionnaire Obstetric-Gynecology Study. *Am J Obstet Gynecol* 2000;183:759–69.
37. Kessler RC, Andrews G, Mroczek D, Ustun B, Wittchen HU. The World Health Organization Composite International Diagnostic Interview Short-Form (CIDI-SF). *Int J Methods Psychiatr Res* 1998;7:171–85.
38. Patten SB, Brandon-Christie J, Devji J, Sedmak B. Performance of the Composite International Diagnostic Interview Short Form for major depression in a community sample. *Chron Dis Can* 2000;21:68–72.
39. Aalto-Setälä T, Haarasilta L, Marttunen M, Tuulio-Henriksson A, Poikolainen K, Aro H, et al. Major depressive episode among young adults: CIDI-SF versus SCAN consensus diagnoses. *Psychological Medicine* 2002;32:1309–14.
40. Patten SB, Beck CA, Wang JL, Maxwell CJ. Measurement issues related to the evaluation and monitoring of major depression prevalence in Canada. *Chron Dis Can* 2005;27:100–106.
41. Rief W, Nanke A, Klaiberg A, Braehler E. Base rates for panic and depression according to the Brief Patient Health Questionnaire: a population-based study. *J Affect Disord* 2004;82:271–6.
42. Wittchen H-U, Knäuper B, Kessler RC. Lifetime risk of depression. *Br J Psychiatry* 1994;165(suppl. 26):16–22.
43. Andrews G, Henderson S, Hall W. Prevalence, comorbidity, disability and service utilisation. Overview of the Australian National Mental Health Survey. *Br J Psychiatry* 2001;178:145–53.
44. Waraich PS, Goldner EM, Somers JM, Hsu L. Prevalence and incidence studies of mood disorders: a systematic review of the literature. *Can J Psychiatry* 2004;49:124–38.
45. Statistics Canada. *Canadian Community Health Survey Mental Health and Well-being*. 2004 Mar 9. Available at: www.statcan.ca/bsolc/english/bsolc?catno=82M0021G.
46. Somers JM, Goldner EM, Waraich P, Hsu L. Prevalence studies of substance-related disorders: a systematic review of the literature. *Can J Psychiatry* 2004;49:373–84.
47. Weissman MM, Bland RC, Canino GJ, Faravelli C, Greenwald S, Hwu HG, et al. The cross-national epidemiology of panic disorder. *Arch Gen Psychiatry* 1997;54:305–9.
48. Kessler RC, Wittchen H-U. Pattern and correlates of generalized anxiety disorder in community samples. *J Clin Psychiatry* 2002;63(Suppl 8):4–10.
49. Wittchen H-U, Fehm L. Epidemiology and natural course of social fears and social phobia. *Acta Psychiatr Scand* 2003;108(Suppl. 417):4–18.
50. Statistics Canada Health Statistics Division. *National Population Health Survey Overview 1994–95*. Ottawa: Minister of Industry; 1995. Report No.: 82–587.
51. Robins LN, Wing J, Wittchen HU, Helzer JE, Babor TF, Burke J, et al. The Composite International Diagnostic Interview. An epidemiological instrument suitable for use in conjunction with different diagnostic systems and in different cultures. *Arch Gen Psychiatry* 1988;45:1069–77.

-
52. Brugha TS, Bebbington PE, Jenkins R. A difference that matters: comparisons of structured and semi-structured psychiatric diagnostic interviews in the general populations. *Psychol Med* 1999;29:1013–20.
 53. Brugha TS, Jenkins R, Taub N, Meltzer H, Bebbington PE. A general population comparison of the Composite International Diagnostic Interview (CIDI) and the Schedules for Clinical Assessment in Neuropsychiatry (SCAN). *Psychol Med* 2001;31:1001–13.
 54. Kessler RC, Berglund P, Demler O, Jin R, Koretz D, Merikangas KR, et al. The epidemiology of major depressive disorder: results from the National Comorbidity Survey Replication (NCS-R). *JAMA* 2003;289:3095–105.