Improving the Detection of Mood Disorders in Primary Care Settings of Resource Poor Countries

A thesis

submitted by

Paul A. Vöhringer M.D.

In partial fulfillment of the requirements
for the degree of

Master of Science

in

Clinical and Translational Research

TUFTS UNIVERSITY

Sackler School of Graduate Biomedical Sciences

Date

May, 2012

ADVISER:

David Kent M.D. M.S.

Table of Contents

Acknowledgements

This work could not have been completed without the help and support of many individuals. I would like to thank everyone who has helped me along the way. Particularly: Dr. S. Nassir Ghaemi for providing me with an opportunity to conduct clinical research in the mood disorders field under his knowledgeable advice as well as for his guidance and kind support over the course of it. I would also like to thank Dr. David Kent for serving as chair on my thesis committee and for his numerous valuable suggestions and Dr. Tanya Logvinenko for her invaluable help in guiding me down the long path of statistical analysis.

Lastly, I would like to thank my family, my wife Carolina and my sons Tomas, Diego and Samuel, without whose support none of this would have been possible.

General Introduction

"Mental Health Research in low-income and scarce settings: how to mind the gap"

There is growing international evidence about the importance of mental health as a development issue in countries with low income [1]. This evidence suggests that people with mental conditions constitute a vulnerable population in urgent need of targeting for development assistance. International organizations have subscribed that mental health problems have huge social and economic costs and therefore it is crucial to determine *what kind of interventions* are needed to break the cycle of poverty and mental illness in these countries [2]. Whether interventions to promote mental health in poor populations should begin with economic intervention to increase financial status or with efforts to improve primary health outcomes (i.e. targeting early detection of symptoms and disabilities associated with mental illness) remains an unsettled question. A recent study indicated that improving health outcomes (intervening in the social drift pathway) and thereby increasing the capabilities of mentally ill patients could be more productive towards bettering their economic outcomes than solely increasing access to financial resources [3]. Although there is robust research questioning whether interventions for early detection and treatment of mental disorders could be sufficiently effective, most of such evidence has been derived from high-income countries [4]. Because of sociocultural and health system differences, the generalizability of such findings are therefore limited.

The present monograph, in two papers, aimed to address ways in which economically scarce settings can provide good research evidence, focusing more specifically on mood disorders detection in low-income primary care settings. The first

paper describes how screening tests could outperform typical assessment of mood disorders administered by general practitioners. The second paper reports how an easily applicable clinical predictive score might help physicians at this level of care to detect people with high risk of having a mood disorder. With improved ability to detect the presence of mood disorders, treatment of these disorders could be improved considerably, leading ultimately to greater economic outcomes for these low-income mentally ill patients.

References

1. WHO. Mental health and development: targeting people with mental health conditions as a vulnerable group. Geneva: world Health Organization and Mental Health and Poverty Project, 2010) report
2. UN. UN General Assembly Resolution on Global Health. A/65/l.27. N York: United Nations, 2010.
3. Lund C, De Silva M, Plagerson S, Cooper S, Chisholm D, Das J, Knapp M, Patel V. Poverty and mental disorders: breaking the cycle in low-income and middle-income countries. Lancet. 2011 Oct 22;378(9801):1502-14.
4. Patel V, Araya R, Chatterjee S, Chisholm D, Cohen A, De Silva M, Hosman C, McGuire H, Rojas G, van Ommeren M.Treatment and prevention of mental disorders in low-income and middle-income countries. Lancet. 2007 Sep 15;370(9591):991-1005

Word
count: 2157

Detecting Mood Disorder in Resource-Limited Primary Care Settings: A comparison of a self-administered screening tool to general practitioner assessment

Vöhringer PA[1,3,5], Jimenez MI[1], Igor MA[1], Fores GA[6], Alvear K[10], Correa MO[7], Sullivan M[3], Holtzman NS[3], Whitham EA[3], Barroilhet SA[2,3], Logvinenko T[9], Kent DM[8], Ghaemi SN[3,4].

1. Hospital Clínico Universidad de Chile. Unidad de Trastornos del Ánimo, Clínica Psiquiátrica Universitaria. Facultad Medicina, Universidad de Chile, Santiago, Chile
2. Escuela de Psicología, Universidad de los Andes, Santiago, Chile
3. Mood Disorders Program, Tufts Medical Center, Boston, MA
4. Tufts University School of Medicine
5. Graduate Program, Clinical and Translational Science Institute (CTSI), Sackler School of Graduate Biomedical Sciences, Tufts University, Boston, MA, United States
6. Departamento Neurología. Universidad Mayor. Santiago, Chile.
7. Departamento Psiquiatría, Facultad Medicina, Universidad de los Andes, Santiago, Chile
8. Institute for Clinical Research and Health Policy Studies, Tufts Medical Center, Boston, MA, USA
9. Biostatistics Research Center at the Institute for Clinical Research and Health Policy Studies, Tufts Medical Center, Boston, MA
10. Universidad Diego Portales, Santiago, Chile

Correspondence to:

Paul Vöhringer MD
Mood Disorder Program
Tufts Medical Center
800 Washington Street #1007
Boston, MA 02111
Telephone: 617-636-5735
Facsimile: 617-636-7795
Email: PVohringer@tuftsmedicalcenter.org

Abstract: (Word Count 318)

Introduction: Mood disorders are under-diagnosed in the primary care setting. While efficacious treatments for mood disorders are available in this level of care, under-diagnosis and delayed detection is associated with under-treatment and worse outcomes. Therefore, new strategies to improve detection in primary care settings are much needed. This study seeks to compare accuracy of self-administered screening tests to routine general practitioner assessment for the detection of current mood disorders.

Methods: This study used a cross-sectional design. One hundred and ninety-seven consecutive patients attending primary care centers in Santiago, Chile filled out the Patients Health Questionnaire 9 (PHQ 9) for depression and the Mood disorder questionnaire (MDQ) for bipolar disorder . The diagnostic accuracy of these self-administered tools were compared to routine GP assessment, with the gold standard diagnosis established by a structured diagnostic interview applied by trained clinicians (SCID I).

Results: 75% of the sample was female; the mean age was 48.5 years (SD 16.8). 37% of the sample had a current mood disorder (positive results in SCID I for depression or SCID I for bipolar disorder). The sensitivity of the screening instruments (SI) was substantially higher than GP assessment (SI: 0.8, [95% CI 0.71, 0.81], versus GP: 0.2, [95% CI 0.12, 0.25]: p-value<0.0001, [95% CI 0.42, 0.59]), without any sacrifice in specificity (SI: 0.9, [95% CI 0.86, 0.96], versus GP: 0.9, [95% CI 0.88, 0.97]: p-value=0.7, [95% CI -0.06, 0.04]). This led to improvement in both the positive predictive value (SI: 0.8, [95% CI 0.82, 0.90], versus GP: 0.6, [95% CI 0.50, 0.64]: p-

value<0.001 [95%CI 0.1, 0.27]) and negative predictive value (SI: 0.9, [95% CI 0.78, 0.91] versus GP: 0.7, [95% CI 0.56, 0.72]: p-value<0.01, [95%CI 0.12, 0.27]) . Conclusions: A self-administered screening test may be useful for increasing the detection of current mood disorder in primary care settings. Further research is needed to test whether broad implementation of screening in primary care is feasible and would improve clinical outcomes.

Introduction

Mood disorders, including major depressive disorder and bipolar disorder, are among the most common mental illnesses in the U.S. and developing countries [1]. Initial evaluation often occurs in a primary care setting, where mood disorders often go unrecognized [2-4]. Some studies point to time constraints in general practice and the competing demands put upon general practitioners (GPs) as the potential cause of such oversight [5, 6]. In underserved populations, psychosocial problems such as occupational or marital difficulties may monopolize already scarce clinical time, masking symptoms or making thorough assessment difficult [5].

Many people suffering from mood disorders have repeatedly visited primary care facilities without receiving a correct diagnosis; in low-income settings, the average patient with a mood disorder has been evaluated by 8 physicians before the correct mood disorder diagnosis is made [7] . Moreover, the average patient has had symptomatic onset at least 10 years before any mood disorder diagnosis [8] .

Major depressive disorder (MDD) is the most common mental disorder seen in primary care patients in developing countries, with prevalence rates ranging from 23%-35% [9] . Some data indicate that about 80% of MDD patients are treated solely in primary care facilities [10]. MDD is usually described as a condition tending to recur over time and therefore become chronic. Epidemiological data show MDD is a major public health problem in Chile [8]. In one study examining prevalence rates of MDD in the primary care setting across 15 cities in 5 continents, Santiago had the highest MDD prevalence (27%) [11].

Bipolar disorder (BD) is a chronic and recurrent disease that also often goes undiagnosed in primary care settings [12-14]. It is the third leading cause of psychosocial disability adjusted life years (DALY, WHO) lost, surpassed only by depression and schizophrenia [15]. Epidemiological studies indicate that the prevalence of bipolar I disorder in general population is 1-2% [13], which is also the case in Chile in the same kind of non-consulting population (2%) according to the Chilean Psychiatric Prevalence Study [16].

Late diagnosis of these disorders may worsen their prognosis [17]. Currently, treatment options are available for mood disorders in primary care facilities, but due to late diagnosis, they are infrequently employed in a timely manner. Thus, new strategies to improve mood disorder detection are badly needed in primary care settings.

The use of self-administered screening tools may be a useful approach, improving detection of current mood disorders, lightening the general practitioner (GP) work load and allowing GPs to attend other demands. The aim of this study is to compare current mood disorder detection accuracy between screening tools and GP in a low income population in Santiago, Chile.

Methods

Participants

Subjects were recruited consecutively from the general medicine programs in ten different low income population primary care centers in Santiago, Chile, between 2009 and 2011. Patients from 18 to 75 yrs old seeking primary care medical evaluation for common illnesses were included. In each primary care center, potential participants were

approached in the waiting area and asked if they were interested in a study on "mood disorders." To participate they had to have a cognitive status compatible with the assessment and give informed consent. Regular clinical assessment of the patients was performed by approximately five general practitioners in each primary care center, although because of turnover, this list of GP was not stable during the enrollment period in each center. The only additional exclusion criterion was the presence of any mood disorder in the 6 months preceding the last month in which the patient enrollment was performed.

Procedure

A cross sectional design was used and tested two different methods of mood disorder detection by comparing them to a gold standard procedure in a "tandem-testing" manner [18].

After routine assessment by the GP , we administered a protocol including an informed consent and the following self-administered screening instruments for mood disorders: the patient health questionnaire (PHQ-9) [19] for depression and the mood disorder questionnaire (MDQ) for bipolar disorder [25] . Then, blind to PHQ-9/MDQ screening results, a trained clinician applied the DSM-IV Structural Clinical Diagnostic Interview (SCID-I) to obtain mood diagnostic status for each participant. Patients were considered as mood disorders cases if they obtained a positive result given by SCID I depression module or SCID I bipolar disorder module.. Finally, medical records were reviewed looking for drugs prescription (antidepressants, anxyolitics, antibiotics, NSAIDS, others), comorbidities (hypertension, diabetes, COPD, epilepsy, drug/alcohol abuse, obesity) and

GP's mood detection (See appendix 1 Flow of patients). The study obtained ethical approval from the Institutional Review Board (IRB) at the Clinical Hospital of the University of Chile.

Instruments

Demographics: An demographic form was used to collect standardized data on age, gender, marital status, and education level.

Determination of GP Assessment**:** In the last month, any of the following findings in patients' medical records, were counted as positive detection of mood disorders:

1. Any explicit mood disorder diagnosis. The words that were used to define an accurate diagnosis of each disorder are as follow: "Depression, major depression, major depressive episode, major depressive disorder, depressive syndrome, depressive-anxious syndrome, bipolar, bipolar disorder, manic episode, hypomanic episode, mania, hypomania or bipolar affective disorder".
2. Any clinical description of mood symptoms in the medical history along with changes in treatment or management of the patient (i.e. antidepressant, mood stabilizers or neuroleptics prescriptions and/or mental health professional referral). The list of words included: "low mood, anxiety and low mood, sleep disturbances, insomnia, suicidal thoughts, worry, stress".

Patient Health Questionnaire (PHQ-9): The PHQ-9 was used to screen for depressive disorders. It is a 9-item self-administered measure of depression, with documented reliability and validity in the sample population [20]. This measure screens for elevated

depressive symptoms in the previous two weeks and can be used to measure either presence and/or severity of depression. A score ≥ 10 pts is indicative of major depressive syndrome [20]. Severity can be categorized as follows: healthy (1-5 pts), subclinical depressive symptoms (6-10 pts), mild depression (11-15 pts), moderate depression (16-20 pts), and severe depression (21-27 pts).

Mood Disorder Questionnaire (MDQ): The MDQ was used to screen for bipolar disorder [21]. It is a 15 item self-administered scale with a score from 1 to 13, also with demonstrated reliability and validity in local studies in primary care [22]. A score ≥ 7 indicates a positive screen for bipolar disorder.

Structured Clinical Interview for DSM IV Axis I Disorders (SCID-I)**[23]**: It was used as the gold-standard for the major depressive episode and manic/hypomanic episode current diagnosis (last month) . It was administered face-to-face to the patient after the GP consult. Raters were clinicians with extensive psychiatric training. This training included reviewing the instruments and the DSM-IV, listening to audio recordings of experienced interviewers, practicing via role play, observing experienced interviewers in person, and conducting interviews with a supervisor present until the interviewer is deemed competent to conduct interviews alone. Interviewers had ongoing (i.e., monthly) trainings to reduce rater drift. SCID-I raters also review all interviews with an experienced and accredited trainer and clinical psychiatrist (PAV), who also randomly reviewed cases' records to check accuracy of researchers assessment.

Data Analysis

Continuous variables were reported as means with standard deviations (SD). Categorical or binary variables were reported as frequencies (%) of the total sample. The primary outcome assessed the degree of disagreement/agreement between GP assessment and screening tools with the gold standard evaluation. A McNemar test was used because the same patient received two assessments at the same time, Cohen's Kappa statistic was used to assess the degree of agreement between detection methods. Statistical significance was set at $p \leq 0.05$ overall, which leads to 0.025 in simple comparisons. With β of 20% (1-β=80% power) and α at 0.05%, using the McNemar test, a sample of 140 patients was sufficient to detect a 20% difference between the two procedures (Screening/GP) and the SCID[24]. Sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV) and positive/negative likelihood ratio (LR+, LR-) were calculated.

Results

One hundred and ninety-seven patients were invited to take part in the study. All patients assented to participate; there were no excluded patients in the sample. Their demographic and clinical characteristics are presented in Table 1.

Consistent with prior studies, there was a 25% prevalence of MDD in this Chilean primary care sample [11], it was predominantly female (75%) as could be expected for patients attending primary care centers in Chile [25, 26], the mean age was 48.5 (SD=16.8). Less than 10% of the sample had reached graduated level education. Almost half of the enrolled patients lived with their children. The prevalence of chronic illness

appeared representative of a primary care population: hypertension was found in 46%, obesity in 23%, and diabetes in 13%. As shown in Table 2, screening tools showed a higher level of agreement with the gold standard compared to GP assessment, [Cohen's kappa 0.7 (SE=0.05 [95%CI 0.5, 0.7]), McNemar test= 2.2 df=1 (p-value=0.13) for screening tools; Cohen's Kappa 0.2 (SE=0.07 [95%CI 0.12, 0.27]: p-value<0.001), McNemar test=29.8 (df=1) (p-value<0.001) for GP assessment]. Screening tools obtained a sensitivity of 0.8 [95%CI 0.71, 0.81] and specificity of 0.9 [95%CI 0.86, 0.96], whereas GP assessment showed values of 0.2 [95%CI 0.12, 0.27], (p-value<0.0001, [95% CI 0.42, 0.59]) and 0.9 [95%CI 0.88, 0.97], (p-value=0.7, [95%CI -0.06, 0.04]) respectively. Regarding PPV and NPV values, screening tests obtained 0.83 [95%CI 0.82, 0.90] and 0.85 [95%CI 0.078, 0.89], while GP assessment obtained 0.64 [95%CI 0.50, 0.64] (p-value=0.001, [95%CI 0.1, 0.2]) and 0.67 [95%CI 0.56, 0.72] (p-value=0.001, [95%CI 0.09, 026]) respectively. (See tables 3 and 4)

Taking into account a prior probability of mood disorders of 37% (prevalence of all mood disorders in this sample), if GP assessment was positive, the posterior probability of mood disorders increased to only 51%. In contrast, this posterior probability increased to 81% with positive scores on mood disorder screening tools. Posterior probability of mood disorder changed little, to 33%, with a negative GP assessment, but decreased notably (14%) with a negative screening tool result (See table 4) (See Appendix 3 Likelihood Ratio Calculations).

Discussion

This study sought to compare the accuracy of detection of mood disorders by GP evaluators versus screening tools, in a primary care setting with a low-income population, using a structured research interview as a reference. Screening tools outperform GP assessments considerably, without increasing false positive cases. This is the first time this comparison has been done in low-income primary care settings. We found that almost 2/3 of true mood disorder cases were missed by GP assessment and the same proportion was correctly classified as having mood disorders by the screening tools. Compared with previous studies [27-29], the PHQ-9, one of the screening tools, showed a higher PPV in this sample. One possible explanation for this finding could be the high prevalence of mood disorders in this Chilean population, seen in previous national and international studies [15, 20]. The likelihood ratios (LR) in our study suggest that screening tools as seen in a nomogram (See appendix 2) could be useful for clinicians. Such a nomogram can show how diagnostic accuracy can improve from a prior probability of 37%, baseline prevalence, to more than 80% for positive posterior probability and 15% for negative posterior probability. This may be especially useful in primary care settings where scarce resources and time impede the ability of GPs to explore mental health.

Strengths of this study, include several methodological precautions to avoid biases, such as tandem testing diagnostic assessment [18] using a structured interview. Possible cases were approached after GP assessment in order to avoid a biased evaluation of mental health status by GPs. Assessment and accuracy of screening tools was

compared with a gold standard, the SCID I semistructured clinical interview, unlike almost all prior studies of these screening tools [30-32]. The gold standard SCID assessment was blind to screening tool or GP assessment results. The screening tools chosen are well validated and have been previously applied in primary care setting [33]. An expert (PAV) routinely reviewed randomly assessed cases to maintain accurate diagnosis.

Limitations of the study include the following: Recruitment bias is possible since patients with more visits to centers, who may have been more severely ill, were more likely to be enrolled. Moreover, the sample is clinically affected and not representative of the general population. This approach was deemed to be most appropriate due to time and budget constraints; it has been used in similar investigations [34, 35]. Also assessment of GP diagnostic skills was based only on medical records and conducted without their knowledge of the specific purposes of the study. This approach has a pragmatic and positive aspect of allowing for assessment of real-world practices but it may underestimate the actual diagnostic accuracy of GPs.

In conclusion, in primary care settings at low-income populations, screening tools for depression (PHQ9) and bipolar disorder (MDQ) improve mood disorder detection accuracy versus general practitioners evaluators; these screening tools may enhance detection and general management of mood disorders in low-income populations. Further research is needed to support this claim.

Table 1 Sample Demographics and Clinical Characteristics (n=197)

Variable		n=197
Age	Mean (SD)	48.5(16.8)
Gender	n(%)	
	Women	147(75)
	Men	49(25)
Marital Status	n (%)	
	Married	83(43)
	Single	61(31)
	Divorced/separated	22(11)
	Widow	15(7)
	LTR with significant other without being married	16(8)
Education	n (%)	
	No graduate	180(91)
	Graduate	17(9)
Occupation	n (%)	
	At home	69(35)
	Working	67(34)
	Retired	30(15)
	Unemployed	13(7)
	Occasional work	18(9)
Lives with*	n (%)	
	Significant other	99(50)
	Children	95(48)
	Other family	67(34)
	Alone	14(7)
	Friends	6(3)
Children	n (%)	
	Yes	160(81)
	No	37(19)
Chronic conditions*	n (%)	
	Absent	69(35)
	Hypertension	91(46)
	Obesity	45(23)
	Smoke	28(14)
	Diabetes	26(13)
	Epi-COPD	10(5)
Medication*	n (%)	
	No drug	80(41)
	hypoglycemic agent	12(6)
	antihypertensives	47(24)
	NSAID	32(16)

	Antibiotics	12(6)
	Anxiolytic	18(9)
	Antidepressants	20(10)
PHQ-9	mean (SD)	9.97(6)
MDQ	mean (SD)	4.74(2.7)
SCID-I	n (%)	
	Depression	49(25)
	Bipolar Disorder	10(5)
	Both	13(7)
	Mood Disorder (either)	72(37)
Screening tools	n (%)	
	Depression	39(20)
	Bipolar disorder	14(7)
	Both	11(6)
	Mood disorders (either)	64(32)
General practitioner	n (%)	
	Mood disorders (either)	26(13)
Endorsed any mood disorder	n (%)	
	General Practitioner	26(13)
	Screening tools	64(32)
	SCID-I	72(37)

Abbreviations: LTR: long term relationship; SCID-I=Structured Clinical Interview for DSM-IV Axis I Disorders, PHQ-9=9-item Patient health Questionnaire, MDQ= Mood Disorders Questionnaire, NSAID=Non steroidal anti-inflammatory drugs, EPI= epilepsy, COPD= Chronic obstructive pulmonary disease. *: more than one category could apply. Total sum more than 100%. SD: standard deviation.

Table 2. GPs and screening tools agreement/disagreement versus gold standard

	Gold Standard SCID for Mood Disorders					
	Positive	Negative	Statistic			
	(n=72) %	(n=125) %	McNemar	*p-value*	Cohen's Kappa	(95%CI)
General practitioners						
Positive	9	5	29.8	<0.0001	0.2	0.05, 0.30
Negative	28	58			(SE=0.07)	
Screening tools						
Positive	26	6	2.20	0.21	0.7	0.54, 0.77
Negative	10	58			(SE=0.05)	
					p-value <0.001	

SCID: Structured Clinical Interview for DSM-IV. CI: Confidence interval. SE: standard error. McNemarX^2: Chi square value of test.

Table 3. Validity and positive predictive value for GP diagnosis and screening questionnaires compared with SCID-I as gold standard

	Patients screened positive		Patients screened negative		Positive predictive value (95%CI)	Negative predictive calue (95%CI)
Method of detection	True positive%	False positive%	True negative%	False negative%		
General practitioners	9	5	58	28	0.64 (0.50-0.64)	0.67 (0.56, 0.64)
Screening tools	26	6	58	10	0.83 (0.82-0.90)	0.85 (0.78, 0.89)
					p-value <0.001	p-value <0.001

Table 4. Sensitivity, specificity, and likelihood ratio with SCID-I as gold standard

			Likelihood ratio (LR)	
Method of detection	Sensitivity (95% CI)	Specificity (95% CI)	Positive LR (95%CI)	Negative LR (95%CI)
General practitioners	0.23 (0.12-0.15)	0.93 (0.88-0.97)	2.0 (1.28-2.72)	0.92 (0.87-0.96)
Screening tools	0.74 (0.71-0.79)	0.92 (0.86-0.96)	8.0 (6.97-9.43)	0.30 (-0.43-1.03)
P-value	<0.0001	0.7	<0.0001	<0.0001

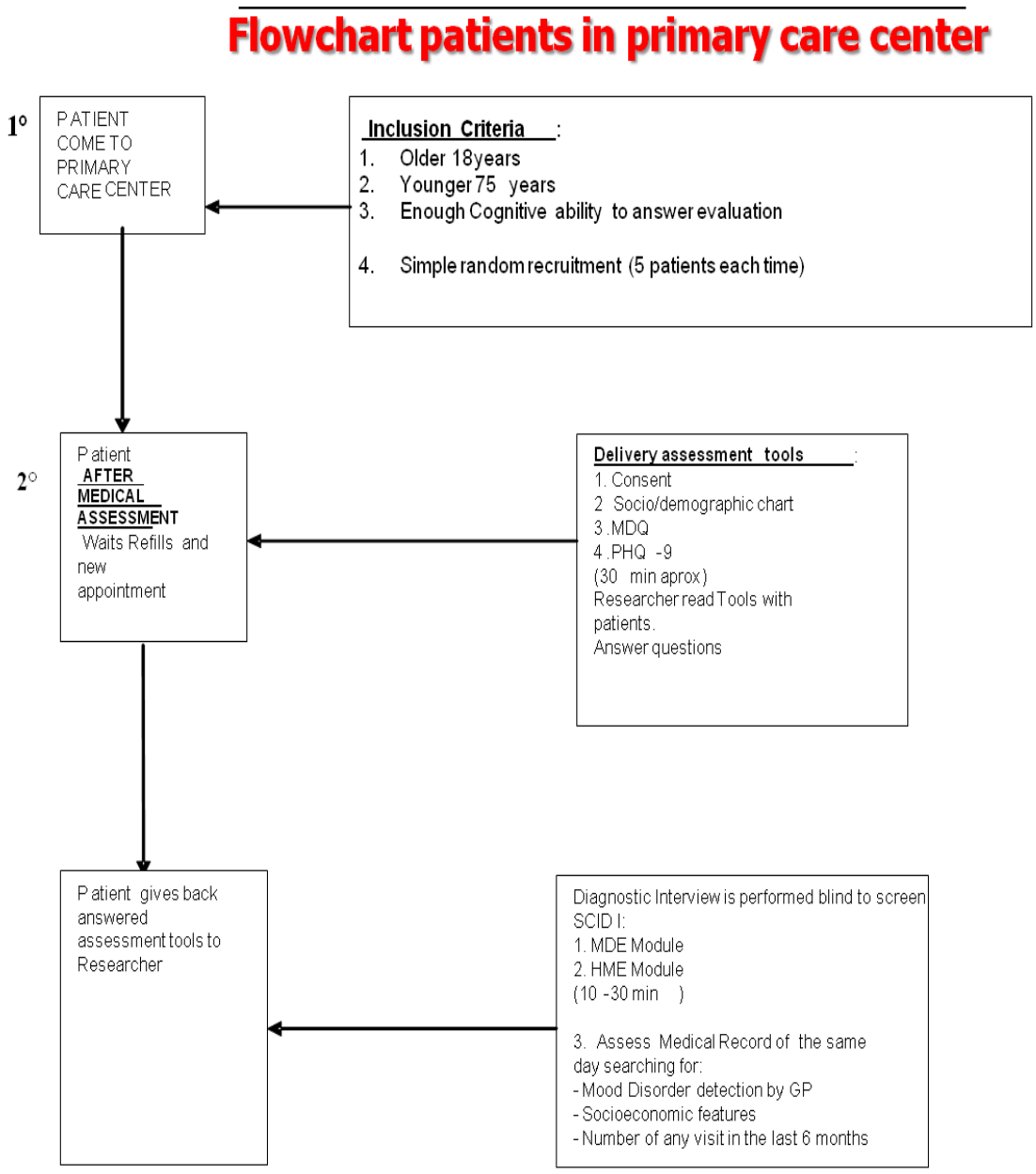
References

1. Kessler, R.C., et al., *Prevalence and treatment of mental disorders, 1990 to 2003.* N Engl J Med, 2005. **352**(24): p. 2515-23.
2. Verhaak, P.F.M., ed. *Mental Disorder in the Community and in General Practice: Doctors' Views and Patients' Demands*. 1st ed. 1995, Avebury: London. 180 pages.
3. Mechanic, D., *Barriers to help-seeking, detection, and adequate treatment for anxiety and mood disorders: implications for health care policy.* J Clin Psychiatry, 2007. **68**(Suppl 2): p. 20-6.
4. Lecrubier, Y., *Widespread underrecognition and undertreatment of anxiety and mood disorders: results from 3 European studies.* J Clin Psychiatry, 2007. **68**(Suppl 2): p. 36-41.
5. Terluin, B.B., EP.; van Marwijk, H.W.; Verhaak, P.F.; van der Horst, H.E., *Detecting depressive and anxiety disorders in distressed patients in primary care; comparative diagnostic accuracy of the Four-Dimensional Symptom Questionnaire (4DSQ) and the Hospital Anxiety and Depression Scale (HADS).* BMC Fam Pract, 2009. **Aug**(23): p. 10:58.
6. Klinkman, M., *Competing demands in psychosocial care. A model for the identification and treatment of depressive disorders in primary care.* Gen Hosp Psychiatry. , 1997. **Mar;19**(2): p. 98-111.
7. Ormel, J.K., V.; Ustun, M.; et al., *Common mental disorders and disability across cultures. Results from the WHO Collaborative study on psychological problems in general health care.* . JAMA, 1994. **272**: p. 1741-1748.
8. Araya R, R.G., Fritsch R, Acuña J, Lewis G, *Common mental disorders in Santiago, Chile: prevalence and socio-demographic correlates.* The British Journal of Psychiatry, 2001. **178**: p. 228-33.
9. Coyne JC, F.-B.S., Schwenk TL, *Prevalence, nature, and comorbidity of depressive disorders in primary care.* Gen Hosp Psychiatry, 1994. **4**(16): p. 267-76.
10. Canada, S. *Canadian community health survey:mental health and well-being* http://www.statcan.gc.ca/pub/82-617-x/index-eng.htm. Canada Catalogue 2006 [cited 2006 November 1, 2006].
11. Simon, G.V.M., Piccinelli M, Fullerton C, Ormel J, *An international study of the relation between somatic symptoms and depression.* N Engl J Med, 1999. **18**(341): p. 1329-35.
12. Murray, C.L., AD, *Mortality by cause for eight regions of the world: Global Burden of Disease Study.* Lancet, 1997. **349**(9061): p. 1269-76.
13. Weissman, M.M., et al., *Cross-national epidemiology of major depression and bipolar disorder.* Jama, 1996. **276**(4): p. 293-9.
14. Weissman, M.B., LM; Leaf, PJ; et al., *Affective disorders*, in *Psychiatric disorder in America: The epidemiological Catchment Area Study*, L. Robins, Editor. 1991, Free press: New York. p. 53-80.
15. Brieger P, E.U., Marneros A, *Frequency of comorbid personality disorders in bipolar and unipolar affective disorders.* Compr Psychiatry 2003. **44**(1): p. 28-34.

16. Vicente, B.K., R; Rioseco, P; Saldivia, S; Levav, I; Torres, S., *Lifetime and 12-Month Prevalence of DSM-III-R Disorders in the Chile Psychiatric Prevalence Study*. Am J Psychiatry, 2006. **163**(8): p. 1362-70.
17. Pearson, S.K., DJ; Simon, GE; Manning, WG; Helstad, CP; Henk, HJ., *Depression among high utilizers of medical care.* J Gen Intern Med, 1999. **14**(8): p. 461-8.
18. Hulley, S.a.S.C., *Designing Clinical Research*. 3rd edition ed. 2007, Baltimore: Lippincott Williams & Wilkins.
19. Kroenke, K.S., RL; Williams, JB. , *The PHQ-9: Validity of a Brief Depression Severity Measure.* . J Gen Intern Med, 2001. **16**(9): p. 606-613.
20. Kroenke, K.S.S., R.L.; Williams, JB.; Löwe, B., *The Patient Health Questionnaire Somatic, Anxiety, and Depressive Symptom Scales: a systematic review.* Gen Hosp Psychiatry, 2010. **32**(Jul-Aug 4): p. 345-59.
21. Hirschfeld, R., et al., *Development and validation of a screening instrument for bipolar spectrum disorder: The Mood Disorder Questionnaire.* American Journal of Psychiatry, 2000. **157**(11): p. 1873-1875.
22. Hirschfeld, R.M., et al., *Screening for bipolar disorder in the community.* Journal of Clinical Psychiatry, 2003. **64**(1): p. 53-9.
23. Spitzer , R.L., et al., eds. *Structured Clinical Interview for DSM-III-R--Patient Version (SCID-P, 9/1/1989 version)*. 1989, Biometrics Research Department, New York State Psychiatric Institute: New York, NY.
24. M, V., Primary Care Companion. CNS disord, 2011. **13**(2).
25. Vega J, M.M., George M, Monreal V, *La Depresión, detección diagnóstico y tratamiento. Guía Clínica para la Atención Primaria*, M.d.S.d. Chile, Editor. 2001, Ministerio de Salud de Chile: Santiago.
26. Ruiz T, A., H. Silva I, and Miranda C, *Clinical and psychometric diagnosis of depression in general medical patients.* Rev Med de Chile, 2001. **129**: p. 627-633.
27. Zuithoff, N.P., et al., *The Patient Health Questionnaire-9 for detection of major depressive disorder in primary care: consequences of current thresholds in a crosssectional study.* BMC Fam Pract. **11**: p. 98.
28. Gilbody, S., et al., *Screening for depression in medical settings with the Patient Health Questionnaire (PHQ): a diagnostic meta-analysis.* J Gen Intern Med, 2007. **22**(11): p. 1596-602.
29. Fann, J.R., et al., *Depression screening using the Patient Health Questionnaire-9 administered on a touch screen computer.* Psychooncology, 2009. **18**(1): p. 14-22.
30. Kroenke, K., et al., *The Patient Health Questionnaire Somatic, Anxiety, and Depressive Symptom Scales: a systematic review.* Gen Hosp Psychiatry. **32**(4): p. 345-59.
31. Huang, F.Y., et al., *Using the Patient Health Questionnaire-9 to measure depression among racially and ethnically diverse primary care patients.* J Gen Intern Med, 2006. **21**(6): p. 547-52.
32. Martin, A., et al., *Validity of the Brief Patient Health Questionnaire Mood Scale (PHQ-9) in the general population.* Gen Hosp Psychiatry, 2006. **28**(1): p. 71-7.
33. Arnow, B.A., et al., *Comorbid depression, chronic pain, and disability in primary care.* Psychosom Med, 2006. **68**(2): p. 262-8.

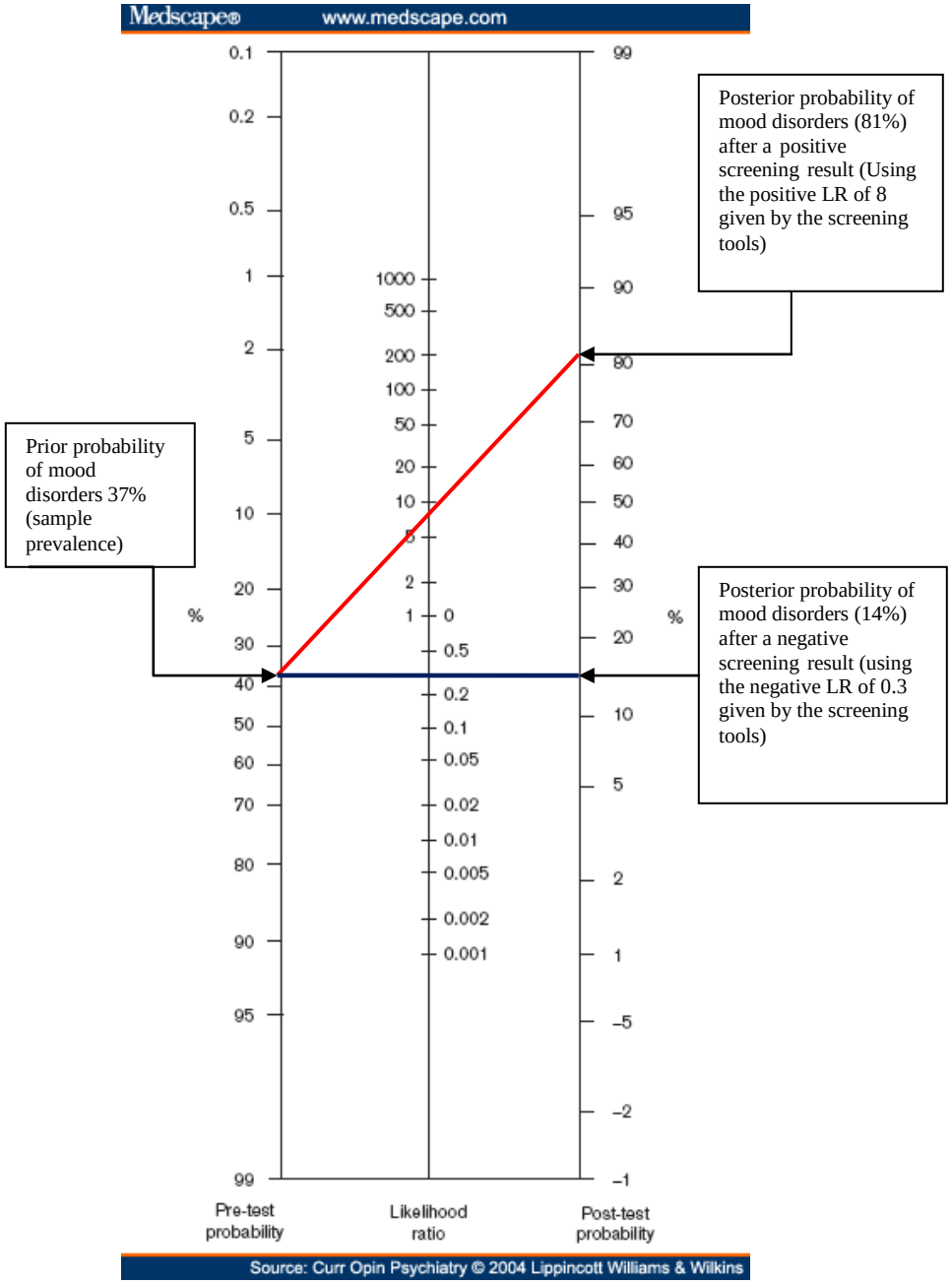
34. Linzer, M., et al., *Gender, quality of life, and mental disorders in primary care: results from the PRIME-MD 1000 study.* Am J Med, 1996. **101**(5): p. 526-33.
35. Ansseau, M., et al., *High prevalence of mental disorders in primary care.* J Affect Disord, 2004. **78**(1): p. 49-55.

Appendix 1 Flowchart of patients in primary care centers



Patients visited their usual primary care center. According to inclusion criteria, researchers approached consecutive patients after general practitioner evaluation (Step 1). After informed consent was given, patients filled out the sociodemographic questionnaire, MDQ (Mood disorder questionnaire) and PHQ-9: (Patient Health Questionnaire) (Step 2). Finally the Semistructured Clinical Diagnostic interview for DSM-IV (SCID I) was applied to obtain the following psychiatric diagnoses: MDE (major depressive episode) or HME (hypomanic/manic episode).

Appendix 2 Fagan Nomogram (1975)



In a Fagan Nomogram (1975), the left hand side column shows the prior probability of a determined condition, usually the prevalence in a given population. (Pre-test probability column). The central column shows likelihood ratio (LR) values and the right hand side column shows the posterior probability of condition after the LR has been applied (post-test probability). The red line shows a positive result given by screening test (positive LR of 8) increasing prior probability of mood disorders from37% to a posterior probability of 81%. The blue line shows a negative result given by screening test (negative LR of 0.32) decreasing prior probability from 37% to a posterior probability of 14%.

Appendix 3. Likelihood Ratio (LR) Calculations

A positive LR+ or negative LR- value, changes the likelihood that a patient has the disease (posterior probability)

Formula to calculate posterior probability using LR:

Prior probability x LR = Posterior probability
(Prevalence of disease)

The positive likelihood ratio reflects both sensitivity and specificity, as the ratio of the probability of being identified as a case among true cases compared with among true non-cases
The negative likelihood ratio reflects sensitivity and specificity, as the ratio of the probability of being identified as non case among true cases compared with among the false cases.

Clinical LR useful values
Positive Likelihood Ratio (LR+) > 5 would be clinically useful
Negative Likelihood Ratio (LR-) < 0.5 would be clinically useful

Operational formulas to calculate LR

LR (+) = P [Result (+)| Disease] / P [Result (+)| No Disease] or = sensitivity/ (1–specificity)

LR (-) = P [Result (-)| Disease] / P [Result (-)| No Disease] or = (1–sensitivity) / specificity

Transformation from Odds to probability and vice versa
Odds= Probability/ (1-probability)
Probability= Odds/ (1+ Odds)

Calculation of Likelihood ratios for GP assessment
LR (+)=0.23 / (1 – 0.93) = **2** [95%CI(1.28, 2.72)]
LR (-)=(1 – 0.23) / 0.93 = **0.92** [95%CI(0.87, 0.96)]
Prior Disease Probability = ***37%*** (0.37) (Prevalence)
[0.37/(1-0.37)] x 2 = 1.07 Posterior Odds
[1.07/(1+ 1.07)]= ***51%*** Posterior probability to have the disease with a positive GP detection
[0.37/(1-0.37)] x 0.92 = 0.50 Posterior Odds
[0.50/1+0.50)] = ***33***% Posterior Probability to have the disease with a negative GP detection

Calculation of Likelihood ratios for screening tests
LR (+)=0.74 / (1 – 0.92) = **8** [95%CI(6.97, 9.43)]
LR (-) =(1 – 0.74) / 0.92 = **0.3** [95%CI(-0.43, 1.03)]
Prior Disease Probability = ***37%*** (0.37) (Prevalence)
[0.37/(1-0.37)] x 8 = 4.30 (Posterior Odds)
[4.30/ 1 + 4.30]= ***81%*** Posterior probability to have the disease with a screening Test (+)
[0.37/(1-0.37)] x 0.3 = 0.16 Posterior Odds
[0.16/ 1+0.16)] **14%** Posterior Probability to have the disease with a screening Test (-)

Word count: 2896

A Clinical Predictive Score for Mood Disorder Risk in Low-Income Primary Care Settings

Vöhringer PA[1,3,5], Jimenez MI[1], Igor MA[1], Fores GA[6], Alvear K[10], Correa MO[7], Sullivan M[3], Holtzman NS[3], Whitham EA[3], Barroilhet SA[2,3], Logvinenko T[9], Kent DM[8], Ghaemi SN[3,4].

1. Hospital Clínico Universidad de Chile. Unidad de Trastornos del Ánimo, Clínica Psiquiátrica Universitaria. Facultad Medicina, Universidad de Chile, Santiago, Chile
2. Escuela de Psicología, Universidad de los Andes, Santiago, Chile
3. Mood Disorders Program, Tufts Medical Center, Boston, MA
4. Tufts University School of Medicine
5. Graduate Program, Clinical and Translational Science Institute (CTSI), Sackler School of Graduate Biomedical Sciences, Tufts University, Boston, MA, United States
6. Departamento Neurología. Universidad Mayor. Santiago, Chile.
7. Departamento Psiquiatría, Facultad Medicina, Universidad de los Andes, Santiago, Chile
8. Institute for Clinical Research and Health Policy Studies, Tufts Medical Center, Boston, MA, USA
9. Biostatistics Research Center at the Institute for Clinical Research and Health Policy Studies, Tufts Medical Center, Boston, MA
10. Universidad Diego Portales, Santiago, Chile

Correspondence to:

Paul Vöhringer MD
Mood Disorder Program
Tufts Medical Center
800 Washington Street #1007
Boston, MA 02111
Telephone: 617-636-5735
Facsimile: 617-636-7795
Email: PVohringer@tuftsmedicalcenter.org

Abstract (Word count: 332)

Introduction: Although validated screening tests for mood disorders are available, general practitioners (GPs) have shown resistance to using them routinely. More feasible detection procedures are needed at this level of care. This study aimed to explore whether a simplified clinical predictive score could be developed to help screen for current presence of mood disorder in low-income primary care settings.

Methods: In a cross-sectional design, 197 patients seen at 10 primary care centers in Santiago, Chile, completed self-administered screening tools for mood disorders; the Patient Health questionnaire (PHQ-9) and the Mood Disorder Questionnaire (MDQ). Trained clinicians applied a gold-standard research diagnostic interview (SCID I), assessing current point-prevalence (not lifetime prevalence) of mood disorders. A simplified clinical predictive model (CM) was developed based on clinical features and selected questions from the screening tools, then a clinical predictive score (PS) was developed based on CM. Full PHQ-9 and GP assessment were compared with PS.

Results: The sample was 75% female, with mean age 48.5 years (SD=16.8). In multivariate logistic regression, the following variables were predictive of the presence of current mood disorder : female gender (OR=2.3 p-value=0.055), presence of long-term relationship with a significant other without being married (OR=8.4 p-value=0.003), presence of untreated chronic conditions (diabetes, hypertension, obesity, and chronic smoking) (OR=1.8 p-value=0.037), treatment with psychotropic drugs (OR=12.4 p-value=0.004), and anhedonia in the previous two weeks (OR=8.5 p-value<0.001). Predictive error for this model was 0.046 [95% CI-0.02, 0.11]. Based on these variables a

simple clinical predictive 8-point score was developed. PS had better discrimination than GP assessment (auROC-statistic=0.80 [95% CI 0.72, 0.85] vs. 0.58 [95% CI 0.52, 0.62] p-value<0.0001), but not as good as the full PHQ-9 (0.89 [95% CI 0.85, 0.93], p-value=0.03). Compared with GP assessment, PS increased sensitivity by 50% at a fixed sensitivity of 90%. Administered in primary care in a typical clinical population, it correctly predicted almost 80% of cases.

Conclusions: An easily-administered clinical predictive score determined with reasonable accuracy, the current risk of mood disorders in low-income primary care settings.

Introduction

In Chile, diagnosis and treatment of mood disorders in the primary care settings (PCS) is based on non-standardized clinical assessment by general practitioners (GP), despite the fact that Santiago, its capital, has the largest proportion of depressed people abroad [1]. GPs are usually overwhelmed, with numerous patients to treat daily, creating an environment in which a complete mental health evaluation is unlikely to be administered. Most GP spend only 5-10 minutes with each patient, making assessment of mood disorders even more difficult.

While screening tests for mood disorders in primary care settings have proven useful, applying and interpreting the results typically takes several minutes [2], a prohibitive amount of time for busy clinicians to apply universally when they are also responsible for screening numerous other conditions. Even when screening tools are attached to electronic records, a considerable proportion of GPs still do not use them extensively [3]. When applied, GPs mostly used screening tools to confirm mood disorder diagnoses or to follow-up its evolution, not for routine screening [4]. These problems are worsened in low-income countries, where electronic records are not available and early prediction could be especially helpful to preserve scarce economic resources. Simpler ways to screen for mood disorder risk in low-income primary care settings are badly needed. In this study, we explored whether a substantially simplified screening instrument using easily obtainable clinical and demographic information with selected key questions from the screening instrument, might have potential as a clinically useful tool to screen for mood disorders in resource-limited primary care settings.

Methods

Using a cross-sectional design, we aimed to develop a logistic regression model, then a score based on this model, supported on a substantially reduced subset of questions used in current screening tools, plus easily available clinical and demographic information as independent variables to predict the presence versus the absence of a current mood disorder as determine by a full diagnostic interview. The sample includes 197 patients enrolled between 2009 and 2011 for a study undertaken to improve the detection of mood disorders in primary care settings.

Participants

To minimize sample bias, subjects were recruited consecutively for a given clinic session from the general medical clinics ("Programas de medicina general") in ten different low-income primary care centers in Santiago, Chile. Patients from 18 to 75 years old seeking primary care medical evaluation for common illnesses were included. They were asked to voluntary participate in a study for mood disorder screening. In order to participate they were required to have a cognitive status compatible with the assessment and to offer informed consent. Having a mood diagnosis in the last 6 months was the only exclusion criterion. No patient refused to participle in the study.

Procedure

After the GP regular assessment, we administered a protocol that included informed consent and self-administered screening instruments for mood disorders: the patient health questionnaire (PHQ-9) [5] for depression and the mood disorder questionnaire (MDQ) for bipolar disorder [6] . Next, blind to PHQ-9/MDQ screening results, a psychiatric trained clinician applied the DSM-IV Structural Clinical Diagnostic Interview (SCID-I) [7] to obtain current mood diagnostic status for each participant (point prevalence, not lifetime prevalence). Finally, medical records were reviewed looking for drugs prescription (antidepressants, anxyolitics, antibiotics, NSAIDS, others), comorbidities (hypertension, diabetes, COPD, epilepsy, drug/alcohol abuse, obesity) and diagnosis of mood disorder by GPs. The study obtained ethical approval from the IRB of main hospital of the University of Chile (Hospital Clínico de Universidad de Chile).

Instruments

Demographics**:** A form was used to identify gender, marital status, and education level.

Determination of GP Assessment: Diagnosis of mood disorders by GPs was based on any of the following findings in patients' medical records within the past month:

3. Any explicit mood disorder diagnosis. Words used to define an accurate diagnosis of each disorder are as follows: "Depression, major depression, major depressive episode, major depressive disorder, depressive syndrome, depressive-anxious syndrome, bipolar, bipolar disorder, manic episode, hypomanic episode, mania, hypomania or bipolar affective disorder".

4. Any clinical description of mood symptoms in the medical history along with changes in treatment or management of the patient (i.e. antidepressant, mood stabilizers or neuroleptics prescriptions and/or mental health professional referral). Applicable words included: “low mood, anxiety and low mood, sleep disturbances, insomnia, suicidal thoughts, worry, stress”.

Patient Health Questionnaire (PHQ-9): The PHQ-9 was used to screen for depressive disorders. It is a 9 item self-administered measure of depression, with documented reliability and validity in the sample population [8]. It screens for elevated depressive symptoms in the previous two weeks and can be used to measure presence and/or severity of depression. A score ≥ 10 pts is indicative of major depressive syndrome [8]. Severity is categorized as follows: healthy (1-5 pts), subclinical depressive symptoms (6-10 pts), mild depression (11-15 pts), and moderate depression (16-20 pts), and severe depression (21-27 pts).

Mood Disorder Questionnaire (MDQ): The MDQ was used to screen for bipolar disorder [6]. It is a 15 item self-administered scale with a score from 1 to 13, with demonstrated reliability and validity in primary care settings [9]. A score ≥ 7 indicates a positive screen for bipolar disorder.

Structured Clinical Interview for DSM IV Axis I Disorders (SCID-I) **[7]** The SCID-I was used as the gold-standard to make the current diagnosis (in the previous month) of major depressive episode and manic/hippomanic episode. It was administered to the patient face-to-face after the GP consult. Raters were psychiatric clinicians, trained by accredited trainers. (Two MD psychiatrist-researchers with expertise on mood disorders, who completed an accredited training on the SCID I)

Data description

Primary outcome variable

The primary outcome was defined as the presence or absence of current (previous month) mood disorder: A major depressive episode (MDE) or a manic/hypomanic episode. The outcome was assessed by trained, experienced clinicians using the diagnostic interview of the diagnostic manual of the American Psychiatric Association in its fourth version (SCID I of DSM IV).

Predictor variables of mood disorders

The main predictor variables were obtained from self-report of symptoms of mood disorders from the PHQ 9 and MDQ. In addition to these, there are well established risk factors for mood disorders in the literature including: female gender [10], low socio-economic status [11], employment status [12], loneliness [13], poor physical and mental health [14], chronic illness [15], marital status and depressive symptomatology [16]. As part of our study, we also collected the following variables that were included as candidate predictor variables: gender; age; educational status: primary school, high school, college; marital status: married, single, separated/divorced, long term relationship with a significant other (LTR), widow; employment status: working at home, paid work, unemployed; living alone or with others; current medications and chronic comorbidity: hypertension, diabetes, chronic smoking, alcohol abuse or use of illicit drugs.

Statistical Analysis

Procedures

Clinical data were reported in a stratified descriptive analysis as means with standard deviations (SD) for continuous variables and percentages for categorical variables, along 95% confidence intervals (table 1). T-tests for independent samples assuming unequal variances were applied for continuous variables, chi-square or Fisher exact test were applied for binary variables (see appendix 1). All clinical and demographics variables that were statistically significant (p-values<0.05), were analyzed along with items from PHQ-9 and MDQ in a univariate fashion with mood disorder status as the primary outcome. All screening tests item with p-values 0.1 or less where included in logistic regression modeling using a backwards selection procedure (with AIC criterion) to obtain the best fitting model. [17]. Robustness of model and logistic model assumptions were tested, as were the presence of collinearity and interactions. Diagnostic evaluations of the model were conducted by removing influential points. Because it is well known that prediction models from multivariable regression analysis usually *overestimate* their regression coefficients, which might result in extreme predictions when applied to new patients [18, 19], a penalized log likelihood shrinkage factor was applied to regression coefficients and the area under the receiver operating curve (auROC)-statistic values to improve internal validity of the classifier. [20]

General modeling approach

A logistic regression model was built using those variables given by the backwards selection procedure, plus clinically based knowledge. This model included mood disorder

status as the primary outcome (binary) and the clinical and demographical predictors along with screening tool items. This classifier was denominated as the "Clinical predictive Model" (CM). Model calibration was assessed using a Hosmer-Lemershow test, which determines whether observed and predicted outcome rates across deciles are statistically different.

Internal Validity assessment of clinical predictive model

A ten-fold cross-validation procedure was applied in order to obtain an internal validation assessment of the CM. [20] The whole sample was randomly partitioned into 10 subsamples. Of these, a single subsample was retained as the testing subsample, and the remaining subsamples were used as training data. This procedure was repeated 200 times and each iteration produced an overall measure of the variability explained by the model in every subsample (Nagelkerke pseudo-R-square) [17]. Then, an average pseudo-R-square for each subset was compared. Comparisons between original, training, and test subsamples were made. Predictive error was computed along with its 95% confidence interval, adjusting original value with those given by the test subsample. (appendix 2)

Clinical predictive score development

After the final clinical predictive model (CM) was built and tested, shrunken regression coefficients of the predictors were transformed into rounded score points (table 3). An easy-to-use clinical predictive score (PS) was constructed (table 4). Total scores were linked to levels of risk of mood disorder. Risk of mood disorders strata were computed

for each score stratum (table 5, appendix 4) [21]. The PS was compared with CM using a chi square test.

Assessment of clinical predictive score performance

Model discrimination was assessed by the auROC statistic. The auROC statistic, along with its 95% confidence interval, was used to ascertain possible increases in predictive information given by the clinical predictive score (PS) and its comparators (see figure 1). Hypothesis testing analyses between them were made using a chi square test. To detect possible differences gained in sensitivity, PS was compared with GPs assessment. In this analysis, specificity was fixed to detect possible differences in sensitivity (appendix 3). To report how PS would perform in usual epidemiologic primary care conditions, a predictive performance analysis was conducted with a predictive threshold set at 0.4 (similar to prevalence of mood disorders in the sample) (appendix 3)

Analyses were completed with Stata 11[20] and R statistical package [22].

Results

Characteristics of the study sample (n=197; 75% women; with mean age of 48.5 and SD of 16.8 years, respectively) are presented with effect estimates along 95% confidence intervals by mood disorder status in Table 1. After GP assessment, all patients consecutively invited consented to participate. Clinical predictors that reached statistical significance at the univariate level were included in the multivariate analysis: Female gender, age, being married, work at home, long term relationship with a significant other

without being married, use of psychotropic drug, untreated chronic conditions. Four PHQ-9/MDQ items that reached 0.10 significance level, were also included in the multivariate analysis.: PHQ-9 item #1 “presence of anhedonia in the last two weeks”; PHQ-9 item #2“presence of low mood in the last two weeks”; MDQ, item #2 “presence of irritability in the last two weeks”; and MDQ, item #16 “familiar background of mood disorders” (appendix 1).

Multivariate Model

The following final variables were selected from a backwards variable selection procedure plus clinical evidence:

1. PHQ 9 question #1: Presence of anhedonia, more than half the days for the last two weeks
2. Presence of long term relationship with a significant other without being married.
3. Presence of untreated chronic conditions.
4. Presence of psychotropic drug treatment.
5. Female gender : This predictor was forced into the model despite the fact it had not been selected with the stepwise procedure, because of strong evidence in the scientific literature [10].

After the shrinkage procedure was applied, final output of the clinical predictive model (CM) model is shown in table 2. Goodness-of-fit shows that CM is a well calibrated model (Hosmer-Lemeshow test=2.69; numbers of groups=10; p-value=0.74 > 0.05). Internal validity results of CM show that it produces a predictive error of 0.046 [95% CI -0.02, 0.11] (appendix 2).

Clinical predictive score results

Clinical predictive score (PS) and mood disorder risk strata along with the clinical predictive model (CM) regression coefficients are shown in table 3. Final PS and mood disorder risk strata are shown in table 4 -5 and appendix 4

Figure 1 shows comparison between areas under the receiver operating curve (auROC) for PS and its comparators. The PS as a simplified screening tool had considerably better discrimination that general practitioner assessment (0.78 [95% CI 0.72, 0.85]; versus 0.58 [95% CI 0.50, 0.64] p-value <0.0001) but not quite as good as the full PHQ-9 (0.89 [95% CI 0.85, 0.93], p-value=0.03). There was no statistical difference between PS and the CM (auROC: 0.79 [95% CI 0.72, 0.84] versus 0.78 [95% CI 0.72, 0.85] X^2=0.02, p-value=0.9). Of note, GP assessment did not show statistical difference compared with a prediction produced totally at random (auROC= 0.50) (X^2= 2.95, p-value=0.09). Full MDQ score was not included in the comparison because it presented high collinearity with full PHQ-9 results, and the latter obtained higher predictive value

Given a fixed specificity of 90%, the PS yielded greater sensitivity than GP assessment (74% versus 23% (appendix 3)).When the PS model was applied with usual epidemiologic features in primary care settings (predictive threshold of 0.4, similar to current mood disorder point prevalence in the sample), it correctly classified mood disorder in almost 80% of the cases (appendix 3).

Discussion

We developed and internally validated a clinical predictive score (PS) for mood disorder risk in the primary care setting which predict current mood disorders risk with reasonably accuracy. The score was based on a clinical predictive model (CM) built using the following easily obtainable current clinical features: long term relationship with a significant other, psychotropic drug treatment, untreated chronic conditions, female gender and anhedonia.

Our results show statistically significant differences in discriminative capacity (auROC statistic) between the full PHQ-9 and the PS and between the PS and GP prediction. Statistical difference was not found neither between the CM and the PS, nor between GPs prediction and a prediction produced totally at random (equivalent to tossing a coin).

Of the items in the Patient Health Questionnaire 9 (PHQ-9), we found that the most predictive item was the "presence of anhedonia in the last two weeks"; this was the only item from the screening instruments that was included in our predictive model.

This PS may enhance detection of mood disorders in low-income countries like Chile, even with primary care constraints such as limited appointment time and limited clinical diagnostic skills. This score is similar to depression symptom scales, like PHQ-9, but unlike the PHQ-9, it is easier to use and requires little, if any, extra clinical work beyond usual GP assessment due to its simplicity, including only one item from PHQ-9 and four likely evident clinical features of the patients. One study found that screening tests like PHQ-9 accurately detect depression, but it required several extra minutes to add and interpret even after being self-administered. This extra time required might not be

feasible for GPs, especially in low-income primary care settings [2]. Moreover, about one half of GPs show resistance to using any kind of standardized assessment [3, 23, 24] despite national guidelines to the contrary for mental illnesses [25].

Another advantage to the clinical predictive score is that it is the first to be developed for mood disorder risk in low-income countries. For the most part, research in mood disorder screening has been produced in advanced countries; one study published a risk prediction algorithm for episodes of major depression in primary care in developed countries (Europe), and later validated in Chilean population. [26]. This tool was built using ten factors which required specific assessment by GPs outside of their usual routine. In contrast, our clinical predictive score has five factors, only one of which, anhedonia, requires specific questioning outside GPs' usual routine. Since scarce resources are a fundamental problem in low-income countries, a simple clinical predictive score like ours may be especially useful.

Limitations of this study include potential bias from overly-highcompliance in a sample in which no patient refused consent. However, such high compliance with research is not uncommon in underdeveloped countries like Chile where medical resources are scarce. Another limitation might be sample size, which, though not small, was not huge, thus potentially impairing accuracy of the results. Furthermore, generalizability could be at issue, if the score is eventually applied to populations with different mood disorders prevalence rates. In that case adjustment of cut-offs values might be needed. Another independent dataset is required to assess predictive external validity of this clinical score. We have begun to develop that replication dataset in Chilean primary care centers. The methods used assessed point-prevalence, not lifetime

prevalence, of mood disorders, and may not generalize to the ability of GPs to make lifetime diagnosis.

Conclusions

A clinical predictive score of risk of current presence of mood disorders in low-income primary care settings has been developed based on a predictive model built with five easily-obtainable clinical features. Its predictive capacity is lower than screening test results, but it appear more feasible for use by general practitioners. External validation is required. If proven generalizable, this clinical predictive score may be useful in detection of mood disorders in the primary care settings.

Table 1.Stratified descriptive analysis of demographic and clinical characteristics of the sample. (n=197)

Variable	Overall sample	Mood Disorders (+) (n=72) %	Mood Disorders(-) (n=125) %	RR or mean difference (95% CI)
Age [Mean (SD)]	48.5(16.8)	47.17(16.48)	53.20(16.45)	-6.03 (-10.76, -1.29)
Gender n(%)				
Women	147(75)	88	68	2.31 (1.24, 4.30)
Men	49(25)	12	31	0.43 (0.23, 0.80)
Marital Status n(%)				
Married	83(43)	32	48	0.64 (0.42, 0.96)
Single	61(31)	33	29	1.11 (0.75, 1.63)
Divorced/separated	22(11)	10	12	0.85 (0.45, 1.62)
Widow	15(7)	7	8	0.90 (0.43, 1.89)
LTR with a significant other without being married	16(8)	18	2	1.42 (1.01, 2.48)
Education n(%)				
No graduate	180(91)	90	92	0.87 (0.48, 1.59)
Graduate	17(9)	10	8	1.14 (0.62, 2.07)
Occupation n(%)				
At home	69(35)	44	29	1.51 (1.05, 2.17)
Working	67(34)	32	37	0.88 (0.59, 1.31)
Retired	30(15)	13	18	0.76 (0.42, 1.36)
Unemployed	13(7)	7	6	1.14 (0.57, 2.24)
Occasional work	18(9)	5	9	0.52 (0.18, 1.46)
Lives with* n(%)				
Significant other	99(50)	47	52	0.90 (0.62, 1.30)
Children	95(48)	53	46	1.20 (0.82, 1.73)
Other family	67(34)	26	37	0.71 (0.46, 1.09)
Alone	14(7)	8	6	1.18 (0.62, 2.24)
Friends	6(3)	3	7	0.55 (-3.64,-1.40)
Children n(%)				
Yes	160(81)	82	85	0.94 (0.58, 1.51)
No	37(19)	18	15	1.05 (0.65, 1.69)

Chronic conditions* n(%)				
Absent	69(35)	36	42	0.62 (0.68, 1.48)
Hypertension	91(46)	42	49	0.83 (0.57, 1.21)
Obesity	45(23)	25	22	1.12 (0.74, 1.70)
Chronic smoking	28(14)	17	13	1.20 (0.75, 1.93)
Diabetes	26(13)	13	12	1.03 (0.56, 1.71)
Epilepsy-COPD	10(5)	4	5	0.90 (0.35, 2.33)
Medication* n(%)				
No drug	80(41)	32	44	0.62 (0.92, 1.91)
hypoglycemic	12(6)	4	7	0.90 (0.28, 2.86)
agents	47(24)	18	28	0.67 (0.42, 1.02)
Antihypertensives	32(16)	13	19	0.73 (0.40, 1.32)
NSAIDs	12(6)	8	5	1.40 (0.77, 2.54)
Antibiotics	18(9)	15	6	1.79 (1.17, 2.73)
Anxyolitics	20(10)	17	6	1.87 (1.25, 2.79)
Antidepressants				
Untreated chronic conditions § n(%)	128(65)	74	52	1.42 (-0.34, -0.9)
PHQ-9 [Mean SD)]	9.97(6)	15.30(5.27)	6.53(4.27)	8.83 (8.79, 8.86)
MDQ [Mean SD)]	4.74(2.7)	5.43(2.51)	3.92(2.72)	1.51 (0.97, 1.04)
SCID-I n(%)				
Depression	49(25)	86	0	13.5 (7.43, 24.5)
Bipolar disorder	10(5)	32	0	3.55 (2.80, 4.50)
Both	13(7)	18	0	3.11 (2.52, 3.84)
Mood disorders (either)	72(37)	100	0	
Screening tools n(%)				
Depression	39(20)	63	4	4.9 (3.44, 6.97)
Bipolar disorder	14(7)	24	7	2.12 (1.50, 3.00)
Both	11(6)	13	1	2.72 (2.06, 3.59)
Mood disorders (either)	64(32)	72	9	5.53 (3.63, 8.41)
General practitioner Diagnosis of mood disorders (either) n(%)	26(13)	24	7	2.02 (1.41, 2.88)
Endorsed any mood disorder n(%)				
General practitioner	26(13)			
Screening tools	64(32)			
SCID-I	72(37)			

SCID-I: Structured Clinical Interview for DSM-IV Axis I Disorders; LTR: long term relationship; PHQ-9: Patient Health Questionnaire 9; MDQ: Mood Disorders Questionnaire; NSAIDs: Non steroidal anti-inflammatory drugs; COPD: Chronic obstructive pulmonary disease; "*": more than one category could apply. Total sum more than 100%; * RR: Relative risk for categorical variables, mean differences for continuous variables; CI: confidence interval; §: Diabetes, hypertension, chronic smoking and obesity.

Table 2. Final CM after shrinkage procedure

Predictor	Coefficient(β)	OR	SE	P value	95% CI
PHQ-9 question #1: anhedonia	2.15	8.58	5.63	0.000	3.43, 32.72
LTR with a significant other without being married	2.14	8.49	7.64	0.003	2.43, 49.52
Untreated chronic conditions	0.63	1.87	0.54	0.037	1.18, 4.47
Psychotropic drug treatment	2.52	12.42	13.29	0.004	2.92, 53.4
Female	0.84	2.31	0.92	0.055	-0.98, 6.52
Constant	-3.69	0.03	0.4	0.000	-6.13, -2.76

CM: Clinical predictive model; PHQ-9: Patient Health Questionnaire 9; LTR: long term relationship; OR: Odds Ratio; SE: standard error; CI: confidence interval

Table 3. Relation between PS and shrunken CM regression coefficients

Predictor	*CM Shrunken Coefficient(β)*	*Predictive score*
PHQ-9 question #1 anhedonia	*2.15*	*2*
LTR with a significant other without being married	*2.14*	*2*
Psychotropic drug treatment	*2.52*	*2*
Untreated chronic conditions	*0.63*	*1*
Female	*0.84*	*1*
TOTAL		*0-8*

PS: clinical predictive score; CM; clinical predictive model; LTR: Long term relationship

Table 4. Final clinical prediction score for mood disorders risk

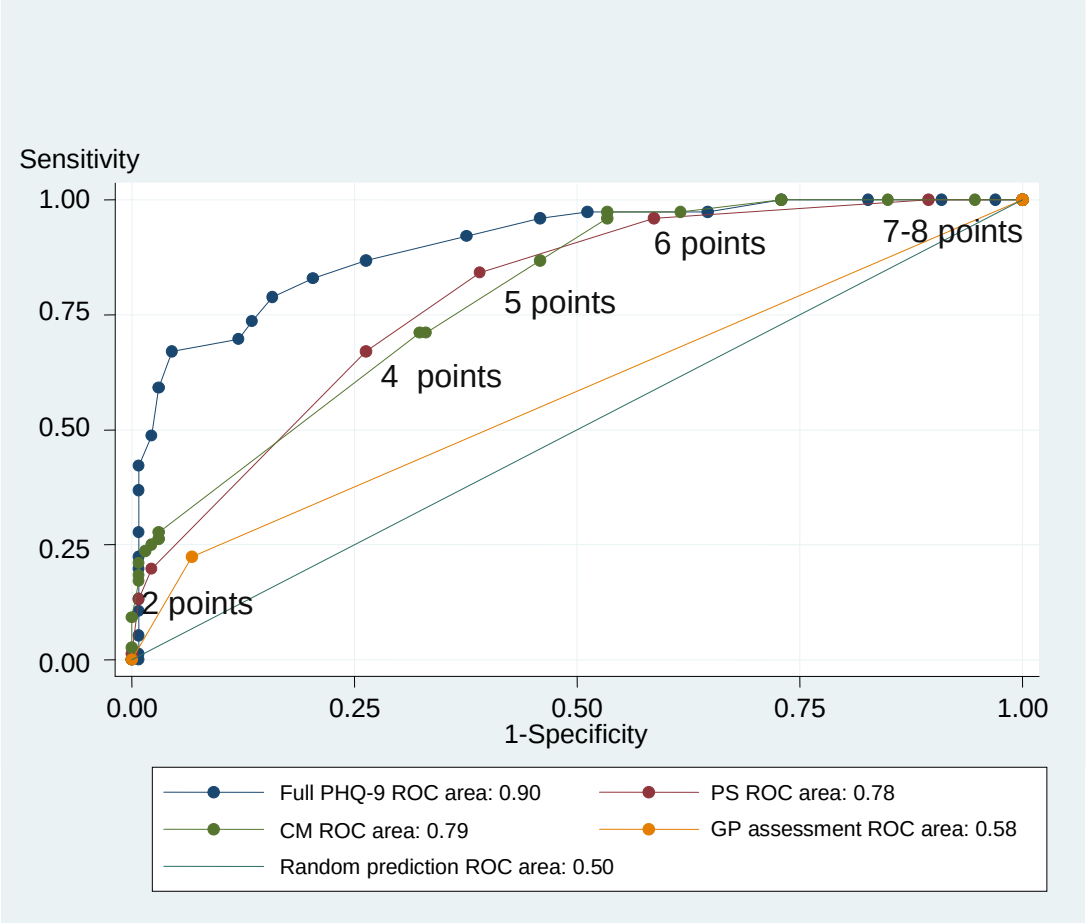
Score 0: if absent; 2 if present	***SCORE***
1. Current LTR with a significant other without being married	
2. Presence of anhedonia in the last two weeks	
3. Current treatment with psychotropic drugs	
Score 0: if absent; 1 if present	
4. Current untreated chronic conditions (HT, DM, obesity, chronic smoking)	
5. Female gender	
Total Score =	
LTR: long term relationship; HT: Hypertension; DM: Diabetes type II.	Range (0-8 points)

Table 5. Clinical predictive score and mood disorder risk strata (n=197)

If TOTAL PS SCORE is:	Likelihood of mood disorder	Mood disorder risk % (cases/total patients)
0-1	**Low**	5% risk (3/55)
2-3	**Moderate**	32% risk (21/66)
4-5	**High**	60% risk (39/66)
6-8	**Very High**	90% risk (9/10)

PS: Clinical predictive score

Figure 1. Discriminative capacity of PS compared with full PHQ-9, general practitioner assessment and CM.



PHQ-9: Patient Health Questionnaire 9; PS: clinical predictive score; CM: clinical predictive model; GP: general practitioner; ROC area: area under the receiver operating curve; "Points": score levels given by the PS.

REFERENCES

1. Simon, G.E.V., M; Piccinelli, M; Fullerton, C; Orme,l J, *An international study of the relation between somatic symptoms and depression.* N Engl J Med, 1999. **341**(18): p. 1329-35.
2. Spitzer, R.L., K. Kroenke, and J.B. Williams, *Validation and utility of a self-report version of PRIME-MD: the PHQ primary care study. Primary Care Evaluation of Mental Disorders. Patient Health Questionnaire.* JAMA, 1999. **282**(18): p. 1737-44.
3. Gill JM YX, G.A., Diamond JJ, Lieberman MI, Klinkman MS, *Electronic clinical decision support for management of depression in primary care: a prospective cohort study.* The Primary Care Companion for CNS Disorders, 2012. **14**(1).
4. Rost, K., et al., *The role of competing demands in the treatment provided primary care patients with major depression.* Arch Fam Med, 2000. **9**(2): p. 150-4.
5. Kroenke, K.S., RL; Williams, JB. , *The PHQ-9: Validity of a Brief Depression Severity Measure.* . J Gen Intern Med, 2001. **16**(9): p. 606-613.
6. Hirschfeld, R., et al., *Development and validation of a screening instrument for bipolar spectrum disorder: The Mood Disorder Questionnaire.* American Journal of Psychiatry, 2000. **157**(11): p. 1873-1875.
7. Spitzer, R., et al., *Structured Clinical Interview for DSM-III-R - Patient Edition (SCID-P, Version 1.0)*. 1990, Washington, D.C.: American Psychiatric Press.
8. Kroenke, K.S.S., R.L.; Williams, JB.; Löwe, B., *The Patient Health Questionnaire Somatic, Anxiety, and Depressive Symptom Scales: a systematic review.* Gen Hosp Psychiatry, 2010. **32**(Jul-Aug 4): p. 345-59.
9. Hirschfeld, R.M., et al., *Screening for bipolar disorder in the community.* Journal of Clinical Psychiatry, 2003. **64**(1): p. 53-9.
10. Weich, S., A. Sloggett, and G. Lewis, *Social roles and gender difference in the prevalence of common mental disorders.* Br J Psychiatry, 1998. **173**: p. 489-93.
11. Weich, S. and G. Lewis, *Material standard of living, social class, and the prevalence of the common mental disorders in Great Britain.* J Epidemiol Community Health, 1998. **52**(1): p. 8-14.
12. Weich, S. and G. Lewis, *Poverty, unemployment, and common mental disorders: population based cohort study.* BMJ, 1998. **317**(7151): p. 115-9.
13. Prince, M.J., et al., *Social support deficits, loneliness and life events as risk factors for depression in old age. The Gospel Oak Project VI.* Psychol Med, 1997. **27**(2): p. 323-32.
14. Bruce, M.L. and R.A. Hoff, *Social and physical health risk factors for first-onset major depressive disorder in a community sample.* Soc Psychiatry Psychiatr Epidemiol, 1994. **29**(4): p. 165-71.
15. Weich, S., *Risk Factors for the Common Mental Disorders in Primary Care*. 2001, Cambridge, England: University of Cambridge.
16. Salokangas, R.K. and O. Poutanen, *Risk factors for depression in primary care. Findings of the TADEP project.* J Affect Disord, 1998. **48**(2-3): p. 171-80.
17. Steyerberg, E., *Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating*. 2009, New York: Springer. 50.

18. Harrell, F.E., Jr., K.L. Lee, and D.B. Mark, *Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors.* Stat Med, 1996. **15**(4): p. 361-87.
19. Spiegelhalter, D.J., *Probabilistic prediction in patient management and clinical trials.* Stat Med, 1986. **5**(5): p. 421-33.
20. LP, S., *Stata 11*. 2009: College Station, Texas.
21. Zuithoff, N.P., et al., *A clinical prediction rule for detecting major depressive disorder in primary care: the PREDICT-NL study.* Fam Pract, 2009. **26**(4): p. 241-50.
22. *R Statistical Package*. 2011.
23. Zimmerman, M.G., J., *Psychiatrists' and nonpsychiatrist physicians' reported use of the DSM-IV criteria for major depressive disorder.* J Clin Psychiatry., 2010. **73**(3): p. 235-8.
24. Zimmerman, M.M., J.B., *Why don't psychiatrists use scales to measure outcome when treating depressed patients?* J Clin Psychiatry. , 2008. **69**(12): p. 1916-9.
25. Panel, D.G., *Depression in Primary Care: Treatment of Major Depression*, in *AHCPR Publication No. 93-0551*. 1993, US Dept of Health and Human Services: Rockville, MD.
26. King, M., et al., *Development and validation of an international risk prediction algorithm for episodes of major depression in general practice attendees: the PredictD study.* Arch Gen Psychiatry, 2008. **65**(12): p. 1368-76.

Appendix 1 Univariate analysis results*

Predictor	Test Value (chi^2, t test)	P-value
Age	3.14	0.01
Married	-2.89	0.02
Gender	9.39	<0.0001
LTR with a significant other without being married	15	<0.0001
Untreated chronic conditions	4.64	0.001
Psychotropic drug treatment	10.29	<0.0001
Working at home	-2.75	0.03
PHQ1	14.05	<0.0001
PHQ2	5.36	0.001
PHQ3	-0.17	0.86
PHQ4	-0.17	0.86
PHQ5	-1.2	0.22
PHQ6	1.72	0.18
PHQ7	1.95	0.16
PHQ8	2.02	0.13
PHQ9	1.12	0.28
MDQ1	0.02	0.9
MDQ2	4.33	0.001
MDQ3	0.75	0.45
MDQ4	0.97	0.36
MDQ5	-0.04	0.9
MDQ6	2.02	0.13
MDQ7	1.92	0.17
MDQ8	0.99	0.34
MDQ9	1.1	0.27
MDQ10	1.57	0.21
MDQ11	1.02	0.32
MDQ12	0.74	0.46
MDQ13	0.52	0.51
MDQ14	0.93	0.38
MDQ15	0.34	0.65
MDQ16	4.33	0.01
MDQ17	1.9743	0.18

*: Appendix 5 and 6 show Patient Health Questionnaire 9 and Mood Disorder Questionnaire scales.

.

Appendix 2 Internal validation assessment of CM

	Original sample	Training sample	Test sample	Difference	Original adjusted	Predictive error (95%CI)
Nagelkerke Pseudo R-squared	0.369	0.383	0.335	0.047	0.321	0.046 (-0.02, 0.11)

CM: clinical predictive model; Original adjusted= Original – training + test; CI: Confidence interval

Appendix 3. Comparison of discriminative capacity between PS, GP assessment, and PS model given current (last month) clinical point prevalence of mood disorders in primary care settings

Model	Sensitivity %	Specificity %	PPV %	NPV %	Given PT	Pts. Correctly Classified %
GP assessment	23	92	64	67	0.5	57
PS model	74	92	83	85	0.5	75
PS model at PCS*	78	80	85	86	0.4*	79

PS: Clinical predictive score; PPV; positive predictive value; NPV: negative predictive value; PT: predictive threshold; Pts: patients; PCS*: Current (last month) mood disorder clinical point prevalence in the Chilean primary care settings (0.4).

Appendix 4. Predicted risk of mood disorder when applying the score to the developing data set (n=197)

Predictive score	Mood disorder risk %	Mood disorder cases/total patients
0	0	0/13
1	7	3/42
2	27	9/33
3	36	12/33
4	58	34/59
5	71	5/7
6	89	8/9
7-8	100	1/1

PS: clinical predictive score

Appendix 5. Patient Health Questionnaire 9 (PHQ-9)

PATIENT HEALTH QUESTIONNAIRE (PHQ-9)

NAME:_____ DATE:_____

Over the last 2 *weeks*, how often have you been bothered by any of the following problems? *(use "✓" to indicate your answer)*	Not at all	Several days	More than half the days	Nearly every day
1. Little interest or pleasure in doing things	0	1	2	3
2. Feeling down, depressed, or hopeless	0	1	2	3
3. Trouble falling or staying asleep, or sleeping too much	0	1	2	3
4. Feeling tired or having little energy	0	1	2	3
5. Poor appetite or overeating	0	1	2	3
6. Feeling bad about yourself—or that you are a failure or have let yourself or your family down	0	1	2	3
7. Trouble concentrating on things, such as reading the newspaper or watching television	0	1	2	3
8. Moving or speaking so slowly that other people could have noticed. Or the opposite —being so figety or restless that you have been moving around a lot more than usual	0	1	2	3
9. Thoughts that you would be better off dead, or of hurting yourself	0	1	2	3
	add columns	+	+	
(Healthcare professional: For interpretation of TOTAL, please refer to accompanying scoring card).	TOTAL:			

10. If you checked off *any problems*, 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?	Not difficult at all _____ Somewhat difficult _____ Very difficult _____ Extremely difficult _____

A2663B 10-04-2005

Appendix 6. Mood Disorder Questionnaire (MDQ)

THE MOOD DISORDER QUESTIONNAIRE

Instructions: Please answer each question to the best of your ability.

	YES	NO
1. Has there ever been a period of time when you were not your usual self and...		
...you felt so good or so hyper that other people thought you were not your normal self or you were so hyper that you got into trouble?	○	○
...you were so irritable that you shouted at people or started fights or arguments?	○	○
...you felt much more self-confident than usual?	○	○
...you got much less sleep than usual and found you didn't really miss it?	○	○
...you were much more talkative or spoke much faster than usual?	○	○
...thoughts raced through your head or you couldn't slow your mind down?	○	○
...you were so easily distracted by things around you that you had trouble concentrating or staying on track?	○	○
...you had much more energy than usual?	○	○
...you were much more active or did many more things than usual?	○	○
...you were much more social or outgoing than usual, for example, you telephoned friends in the middle of the night?	○	○
...you were much more interested in sex than usual?	○	○
...you did things that were unusual for you or that other people might have thought were excessive, foolish, or risky?	○	○
...spending money got you or your family into trouble?	○	○
2. If you checked YES to more than one of the above, have several of these ever happened during the same period of time?	○	○
3. How much of a problem did any of these cause you – like being unable to work; having family, money or legal troubles; getting into arguments or fights? *Please circle one response only.* No Problem Minor Problem Moderate Problem Serious Problem		
4. Have any of your blood relatives (i.e. children, siblings, parents, grandparents, aunts, uncles) had manic-depressive illness or bipolar disorder?	○	○
5. Has a health professional ever told you that you have manic-depressive illness or bipolar disorder?	○	○