

Case finding of lifestyle and mental health disorders in primary care: validation of the 'CHAT' tool

Felicity Goodyear-Smith, Nicole M Coupe, Bruce Arroll, C Raina Elley, Sean Sullivan and Anne-Thea McGill

ABSTRACT

Background

Primary care is accessible and ideally placed for case finding of patients with lifestyle and mental health risk factors and subsequent intervention. The short self-administered Case-finding and Help Assessment Tool (CHAT) was developed for lifestyle and mental health assessment of adult patients in primary health care. This tool checks for tobacco use, alcohol and other drug misuse, problem gambling, depression, anxiety and stress, abuse, anger problems, inactivity, and eating disorders. It is well accepted by patients, GPs and nurses.

Aim

To assess criterion-based validity of CHAT against a composite gold standard.

Design of study

Conducted according to the Standards for Reporting of Diagnostic Accuracy statement for diagnostic tests.

Setting

Primary care practices in Auckland, New Zealand.

Method

One thousand consecutive adult patients completed CHAT and a composite gold standard. Sensitivities, specificities, positive and negative predictive values, and likelihood ratios were calculated.

Results

Response rates for each item ranged from 79.6 to 99.8%. CHAT was sensitive and specific for almost all issues screened, except exercise and eating disorders. Sensitivity ranged from 96% (95% confidence interval [CI] = 87 to 99%) for major depression to 26% (95% CI = 22 to 30%) for exercise. Specificity ranged from 97% (95% CI = 96 to 98%) for problem gambling and problem drug use to 40% (95% CI = 36 to 45%) for exercise. All had high likelihood ratios (3–30), except exercise and eating disorders.

Conclusion

CHAT is a valid and acceptable case-finding tool for most common lifestyle and mental health conditions.

Keywords

lifestyle; mass screening; mental health; primary health care; risk reduction behavior; validation studies.

INTRODUCTION

Increasing emphasis on preventive practice in primary healthcare necessitates identifying patients with lifestyle and mental health risk factors. Many at-risk behaviours and conditions may not be identified in routine practice at present. For example, the recent Mental Health in General Practice Investigation study reported that a third of primary care patients had experienced a diagnosable mental health disorder according to the *Diagnostic and Statistical Manual of Mental Disorders, Fourth edition (DSM-IV)*.¹ Despite the prevalence of mental health disorders presenting in primary care settings, the World Health Organization reports that many of these disorders go undiagnosed and they estimate that less than one-third of those who need treatment receive it.²

General practice is highly accessible to patients requiring help with problem behaviours, and patients expect to receive preventive lifestyle advice from their GP.³ Research shows that 80% of the population consult with their GP at least once a year.⁴ However, opportunistic screening is likely to have a limited effect⁵ and, given consultation time restraints, compliance with routine screening

F Goodyear-Smith, MGP, FRNZCGP, associate professor; *B Arroll*, PhD, FRNZCGP, professor; *CR Elley*, PhD, FRNZCGP, doctor; *A-T McGill*, BSc, FRNZCGP, senior lecturer, Department of General Practice and Primary Health Care, School of Population Health, University of Auckland. *NM Coupe*, PhD, post-doctoral fellow, Nga Pae o te Maramatanga, Whariki, Massey University. *S Sullivan*, PhD, director, Abacus Counselling and Training Services Ltd., Auckland Mail Centre, Auckland, New Zealand.

Address for correspondence

Dr Felicity Goodyear-Smith, Department of General Practice and Primary Health Care, School of Population Health, Faculty of Medical and Health Sciences, University of Auckland, PB 92019, New Zealand.
E-Mail: f.goodyear-smith@auckland.ac.nz

Submitted: 10 May 2007; Editor's response: 30 July 2007; final acceptance: 3 September 2007.

©British Journal of General Practice 2008; 58: 26–31.

DOI: 10.3399/bjgp08X263785

regimes can be low for both patients and practitioners.^{6,7}

Some patients are embarrassed or object to being asked sensitive questions about their lives. For example, a number of studies examining women's acceptability of domestic violence screening show huge variability in the percentage of women who object — ranging from 15 to 57%.⁸ Results of such studies indicate there is a need for development of tools to help primary care better address this sector of practice for the population. Furthermore, any tool would have to be acceptable, reliable, and valid before widespread use.

The short self-administered Case-finding and Help Assessment Tool (CHAT) was developed for lifestyle and mental health assessment of adult patients (aged ≥ 16 years) in primary health care. The CHAT assesses for tobacco use, alcohol and other drug misuse, problem gambling, depression, anxiety and stress, abuse, anger problems, inactivity, and eating disorders. The tool was designed by a team of GPs, university researchers, a psychologist, and a community-based brief-intervention educator of primary healthcare providers.

The acceptability and feasibility of the CHAT has been assessed previously in 2543 patients of GPs and practice nurses in 41 rural and urban New Zealand general practices.⁹ Patients came from diverse ethnic, geographical, and socioeconomic backgrounds. The sample prevalence of positive responses ranged from 2.8% (gambling) to 42.7% at risk of depression. The number of patients requesting immediate assistance with these responses (from 0.5% for gambling through to 13.5% for depression and anxiety) did not overwhelm clinicians. All patients and practitioners in the evaluation study⁹ completed feedback forms, which recorded objections to any of the screened topics and positive and negative responses to the tool. The tool was well accepted by patients, with few objections to specific questions (0.1–0.8%) and was not considered overly burdensome by practitioners.

Given that the tool is largely based on previously reported measures and has undergone considerable practitioner and patient assessment, it can be assumed that it has good content validity.

Subsequent to this study, the CHAT has been used in a number of clinical settings, including the general practice of one of the co-authors, with no complaints reported.

While many of the individual items in the CHAT have been validated against more comprehensive tools, the CHAT had not been validated as a whole. The aim of this study was to conduct a criterion-based validation of the tool against a group of gold standard instruments.

How this fits in

Detecting lifestyle and mental health risk factors in primary healthcare patients can have valuable health outcomes, but frequently does not occur in the context of busy consultation schedules. The short self-administered Case-finding and Help Assessment Tool (CHAT) identifies tobacco use, alcohol and other drug misuse, problem gambling, depression, anxiety and stress, abuse, anger problems, inactivity, and eating disorders. It is well accepted by patients, GPs, and nurses. On validation against a composite gold standard, the CHAT has proved to be a valid and acceptable case-finding tool for most common lifestyle and mental health disorders.

METHOD

The tool was validated in primary healthcare practices in a primary health organisation in South Auckland (a socioeconomically deprived region) and a primary health organisation in the North Shore (a socioeconomically advantaged region) of Auckland.

The gold standard tools were selected using a pragmatic approach (Table 1). While DSM-IV diagnostic interviews might be ideal, conducting these in combination would have been too time-consuming to be practical in the primary healthcare setting; and requiring each participant to complete one diagnostic instrument on a random basis would have required a prohibitively large sample size.

All consecutive primary healthcare patients aged 16 years and over attending the practices were invited to complete the CHAT and a composite gold standard. Exclusion criteria were inability to understand English, or mental impairment that precluded meaningful participation. Recruitment ceased when 1000 patients had been recruited.

The CHAT and composite gold standard forms were self-administered by patients in the waiting room. There was a research assistant present to assist with consent and collection who was advised not to look at the screening tool answers when patients were completing the gold standard form. The study was conducted according to the Standards for Reporting of Diagnostic Accuracy (STARD) statement for diagnostic tests.¹⁰ Where the tool detected a risk factor that the patient wanted addressed, the GP could either deal with the problem at the time of the consultation or schedule a later appointment.

Data analysis was conducted using Microsoft Excel. Scores on the gold standard forms were dichotomised as 'case' or 'not-a-case'. Sensitivities, specificities, positive and negative predictive values (PPV and NPV), and likelihood ratios were calculated using the online statistical calculator from the Centre for Evidence-Based Medicine (www.cebm.utoronto.ca/practise/ca/statscal/).

Table 1. Gold standards used for each component of the CHAT.

Risk factor	CHAT questions ^a	Positive CHAT	Gold standard tool
Smoking	How many cigarettes do you smoke every day? Do ever feel the need to cut down or stop your smoking?	Yes >10 cigarettes/day or Yes to second question	Heavy Smoking Index
Alcohol	Do you feel the need to cut down on your drinking alcohol? In the past year, have you drunk more alcohol than you meant to?	Yes to either question	Alcohol Use Disorders Identification Test
Other drugs	Do you ever feel the need to cut down on your non-prescription or recreational drug use? In the past year, have you ever used non-prescription or recreational drugs more than you meant to?	Yes to either question	Drug Abuse Screening Test
Gambling	Do you sometimes feel unhappy or worried after a session of gambling? Does gambling sometimes cause you problems?	Yes to either question	South Oaks Gambling Screen
Depression	During the past month have you often been bothered by feeling down, depressed, or hopeless? During the past month have you often been bothered by having little interest or pleasure in doing things?	Yes to either question	Patient Health Questionnaire depression scale
Anxiety	During the past month have you been worrying about a lot of different things?	Yes	Hospital Anxiety and Depression Scale
Abuse/violence	Is there anyone in your life of whom you are afraid or who hurts you in any way? Is there anyone in your life who controls you and prevents you from doing what you want?	Yes to either question	Conflict Tactics Scale (CTS-1) and Hurts, Insults, Threatens, Screams tool
Anger	Is controlling your anger sometimes a problem for you?	Yes	CTS-1
Physical inactivity	As a rule, do you do at least 30 minutes of moderate or vigorous exercise (such as walking or a sport) on 5 or more days of the week?	No	Aerobics Center Longitudinal Study – physical activity questionnaire
Eating disorders (anorexia nervosa, bulimia, or bingeing)	Do you often feel that you can't control what or how much you eat? Does your weight affect the way you feel about yourself?	Yes to either question	'Sick, Control, One, Fat, Food' questionnaire Eating Disorder Screen for Primary Care

^aFor each item, patients are asked: 'If yes, do you want help with this?', with the options 'Yes', 'Yes but not today', or 'No'; except for the exercise question, which asks: 'If no, do you want help with this?'. CHAT = Case-finding and Help Assessment Tool.

RESULTS

There was a 2% decline rate from consecutive eligible patients invited to participate. Sets of completed CHAT and gold standard forms were available from 995 patients, although not all CHAT questions and individual gold standard questionnaires were completed by all participants. Response rates ranged from 99.8% for smoking to 79.6% for anger questions (Table 2). Where response was incomplete, generally the CHAT was completed but not the relevant section of the composite gold standard.

Sixty per cent of the participants were New Zealand European, 16% Maori, 4% Pacific people, and 20% Asian or 'other' ethnicity. This is representative of the New Zealand population. Seventy-one per cent were female, a proportion commonly found in general practice adult-attending patient populations. The age distribution was slightly skewed to the older age groups: 21% were aged 16–29 years, 37% were in the 30–49 year age bracket, and 42% were aged ≥50 years.

Sensitivities, specificities, PPV, and NPV are recorded in Table 3. PPVs ranged from 68% for nicotine dependency and 44% for problematic drinking, which were high prevalence conditions (14.7% and 12% respectively) to 1% for physical violence (a low prevalence condition of 0.6%). NPVs were all between 97 and 100%, except for the exercise question, which produced flawed results in all probability due to systematic error.

The positive and negative likelihood ratios (LR+ and LR–), condition prevalence, and response rates are reported in Table 2. Likelihood ratio incorporates both sensitivity and specificity and is a direct estimate of how much the test result changes the odds of having the condition. The questions for smoking, problematic drug use, gambling, and abuse all had a LR+ >10, which indicates that it is a very good test for 'ruling in' the condition. Alcohol dependency and major depression had LR– <0.1 which indicates that it is a very good test for 'ruling out' these conditions.

Table 2. CHAT positive and negative likelihood ratios, condition prevalence, and response rate.

	LR+ (95% CI)	LR- (95% CI)	Prevalence, % (positive gold standard)	Response rate, ^a %
Nicotine dependency	13.13 (10.22 to 16.87)	0.11 (0.07 to 0.18)	14.7 (142/963)	99.8
Alcohol				
Problematic drinking	5.75 (4.76 to 6.93)	0.22 (0.15 to 0.32)	12 (116/967)	97.2
Alcohol dependency	4.52 (3.80 to 5.36)	0.07 (0.01 to 0.50)	1.8 (17/967)	97.2
Problematic drug use	16.24 (9.34 to 28.26)	0.57 (0.41 to 0.79)	3.0 (29/971)	97.6
Problematic gambling	30.05 (17.71 to 45.81)	0.13 (0.02 to 0.23)	1.8 (16/909)	91.4
Major depression	3.09 (2.76 to 3.45)	0.06 (0.02 to 0.23)	5.3 (50/947)	95.2
Anxiety	3.32 (2.91 to 3.80)	0.16 (0.10 to 0.28)	11 (101/927)	93.2
Abuse				
Verbal abuse CTS	12.27 (7.79 to 19.32)	0.40 (0.23 to 0.69)	2.6 (21/794)	79.8
Physical abuse CTS	13.15 (7.84 to 22.05)	0.21 (0.04 to 1.23)	0.6 (5/794)	79.8
Victim according to HITS	11.11 (7.08 to 17 to 44)	0.54 (0.40 to 0.72)	4.9 (41/838)	84.2
Anger				
Verbal abuse	5.34 (3.62 to 7.96)	0.42 (0.23 to 0.75)	2.4 (19/792)	79.6
Physical abuse	3.90 (1.44 to 10.57)	0.57 (0.22 to 1.53)	0.5 (4/792)	79.6
Exercise	0.43 (0.36 to 3.08)	1.83 (1.62 to 2.08)	46 (401/874)	87.8
Eating disorder	2.75 (2.46 to 3.08)	0.14 (0.08 to 0.24)	14 (127/932)	93.7

^aTotal n = 995. CHAT = Case-finding and Help Assessment Tool. CTS = Conflict Tactics Scale. HITS = 'Hurts, Insults, Threatens, Screams' tool. LR+ = likelihood ratio of positive test. LR- = likelihood ratio of negative test.

While the 'eating disorder' questions in the CHAT had good test properties to exclude an eating disorder (NPV = 98%; LR- = 0.14) there was low LR+ of 2.75, specificity (67%), with 14% 'prevalence'. Given that less than 5% of primary healthcare patients are likely to meet DSM-IV criteria for eating disorders,¹ it seems apparent that these questions were measuring something significantly broader than eating disorder itself.

DISCUSSION

Summary of main findings

The CHAT tool is both valid and acceptable for lifestyle and mental health disorder screening in primary care. All items showed good sensitivity, specificity, and likelihood ratios when compared with gold standard instruments, except for exercise and eating disorders. The validation of this tool used a pragmatic approach, given that a battery of gold standard tests needed to be administered in a waiting-room environment. This meant that brief, but well validated, tools such as the Alcohol Use Disorders Identification Test, Patient Health Questionnaire depression scale (PHQ-9), and the Hospital Anxiety and Depression Scale were used rather than longer instruments, such as a full Composite International Diagnostic Interview, which would have been too time-consuming and impractical in this setting.

The low response rates (79.6–84.2%) for the abuse and anger questions reflect reluctance of responders

to complete the gold standard (Conflict Tactics Scale; CTS-1) for these items. This tool takes a long time to complete (and is therefore not appropriate for use in a waiting-room setting), and also asks particularly sensitive questions. The other lengthy gold standard was the Aerobics Center Longitudinal Study for exercise which was completed by only 87.8% of responders, whereas response rates for the other conditions were 91.4–99.8%.

Sensitivity and specificity of the single exercise question were extremely low (26% and 40% respectively), with the PPV only 27%, compared with 81% in a previous primary care study which validated it as a screening question for being sedentary.¹¹ This was probably due to the way the question was presented in the CHAT, causing confusion to the responders. With all the other questions, a 'yes' response indicated a possible condition, whereas with the exercise response 'no' indicated probable sedentary behaviour. With the exception of the exercise question, 'no' responses were down the left-hand column of the tool. This was to aid the clinician by running the eye down the left-hand column to check for absence of conditions needing further enquiry. Examining the gold standard responses of those who ticked 'no' for the exercise question, it is apparent that many actually were very physically active, and it seems probable that the format caused them to invert their replies.

The eating disorder prevalence in this study was 14%, which is more than seven times higher than

Table 3. CHAT sensitivity, specificity, positive and negative predictive values.

	Sensitivity, % (95% CI)	Specificity, % (95% CI)	PPV, % (95% CI)	NPV, % (95% CI)
Nicotine dependency	89 (83 to 93)	93 (91 to 94)	68 (61 to 74)	98 (97 to 99)
Alcohol				
Problematic drinking	81 (73 to 87)	86 (83 to 88)	44 (37 to 51)	97 (96 to 98)
Alcohol dependency	94 (73 to 99)	79 (77 to 82)	8 (5 to 12)	100 (99 to 100)
Problematic drug use	45 (28 to 62)	97 (96 to 98)	33 (21 to 49)	98 (97 to 99)
Problematic gambling	88 (64 to 97)	97 (96 to 98)	35 (21 to 50)	100 (99 to 100)
Major depression	96 (87 to 99)	69 (66 to 72)	15 (11 to 19)	100 (99 to 100)
Anxiety	88 (80 to 93)	74 (70 to 76)	29 (24 to 34)	98 (97 to 99)
Abuse				
Verbal abuse CTS	62 (41 to 79)	95 (93 to 96)	25 (15 to 38)	99 (98 to 100)
Physical abuse CTS	80 (38 to 99)	94 (92 to 95)	1 (0 to 2)	100 (99 to 100)
Victim according to HITS	49 (34 to 64)	96 (94 to 97)	4 (2 to 5)	97 (96 to 98)
Anger				
Verbal abuse	63 (41 to 81)	88 (86 to 90)	12 (1 to 19)	99 (98 to 100)
Physical abuse	50 (15 to 85)	87 (85 to 89)	1 (0 to 1)	100 (99 to 100)
Exercise	26 (22 to 30)	40 (36 to 45)	27 (23 to 32)	39 (35 to 44)
Eating disorder	91 (84 to 95)	67 (64 to 70)	30 (26 to 36)	98 (96 to 99)

CHAT = Case-finding and Help Assessment Tool. CTS = Conflict Tactics Scale. HITS = 'Hurts, Insults, Threatens, Screams' tool. NPV = negative predictive value. PPV = positive predictive value.

that reported in other studies.¹ This study has demonstrated that the gold standard is not precise enough as it was probably identifying participants who had concerns about being overweight and eating patterns, rather than a formal eating disorder. As eating disorders are relatively rare in general practice, it is planned to remove this question from future screening tools.

Strengths and limitations of the study

The limitations of this study firstly are that some of the conditions had very small numbers of responders with the condition. Secondly, pragmatic gold standard instruments were used because administering diagnostic interviews for all conditions would have been excessively time-consuming to conduct in the primary healthcare setting. Some of the gold standards used are diagnostic tools and some are primarily screening instruments that have

been validated against a gold standard, such as a DSM-IV interview. For example, while the South Oaks Gambling Screen might be considered a screening tool, there is no epidemiological gold standard in the area of disordered gambling prevalence.¹² The South Oaks Gambling Screen was used as the gold standard as it has been the primary method used to identify problem and pathological gambling since the late 1980s.¹³

A strength of this study is that there was a very low decline rate and a consecutive group of patients. This means these results are generalisable to other general practices for consecutive patients. For the high prevalence conditions adequate numbers were found which produced narrow confidence intervals. A further strength is that patients completed the gold standards for all the conditions.

Implications for future research and clinical practice

The CHAT provides an important tool for routine use in primary healthcare settings in lifestyle and mental health domains where strong argument can be made for case finding and subsequent intervention.

It is envisaged that CHAT will be used in different ways in different settings. Some practices are already using it with all new patients, and asking adult patients to complete it if it has been more than 2 years since their last visit. Moves are under way to have the CHAT entered electronically and self-administered on a waiting-room touch screen where this is available. Because patients complete the CHAT before their consultation, they indicate whether or not it has brought up any issues that they wish to address in their consultation. This means that it is unlikely to inhibit or hinder a patient discussing their own agenda rather than their doctor's prevention agenda.

As well as being a self-administered waiting-room tool, the CHAT can also be administered by the GP and by other health professionals, such as the practice nurse. Once available electronically, the CHAT will be integrated into the patient's electronic record (patient management system). This will also allow for 'second tier' tools (for example the Alcohol Use Disorders Identification Test) to be available for administering from the practitioner's computer (including automatic scoring) should a patient have a positive response on the CHAT.

Because it is quick to use and well accepted, the CHAT can be used for follow-up after intervention for identified problems. The next step in the development of this tool is to conduct a trial to test against clinical outcomes. This will establish whether systematic use of the CHAT in primary healthcare setting leads to better health outcomes for patients.

Funding body

The Oakley Mental Health Foundation (project grant 3606867) and the Charitable Trust of the Auckland Faculty of the Royal New Zealand College of General Practitioners (project grant 3601614)

Ethical approval

Approval was obtained from the Auckland Ethics Committee (AKY/04/04/079)

Competing interests

The authors have stated that there are none

Acknowledgements

The study involved collaboration between primary healthcare researchers with specific lifestyle or mental health interests and expertise in the Department of General Practice and Primary Health Care, the University of Auckland in the development of the tool. We thank all those who have made contributions to this work over the several years of its development

Discuss this article

Contribute and read comments about this article on the Discussion Forum: <http://www.rcgp.org.uk/bjgp-discuss>

REFERENCES

1. MaGPIe Research Group. The nature and prevalence of psychological problems in New Zealand primary healthcare: a report on Mental Health and General Practice Investigation (MaGPIe). *N Z Med J* 2003; **116**(1171): 1171.
2. Andrews G, Sanderson K, Slade T, Issakidis C. Why does the burden of disease persist? Relating the burden of anxiety and depression to effectiveness of treatment. *Bull World Health Organ* 2000; **78**(4): 446–454.
3. Johansson K, Bendtsen P, Akerlind I. Advice to patients in Swedish primary care regarding alcohol and other lifestyle habits: how patients report the actions of GPs in relation to their own expectations and satisfaction with the consultation. *Eur J Public Health* 2005; **15**(6): 615–620.
4. Ministry of Health. *Taking the pulse: 1996/97 New Zealand health survey*. Wellington: Ministry of Health, 1999.
5. Norman P, Fitter M. The potential and limitations of opportunistic screening: data from a computer simulation of a general practice screening programme. *Br J Gen Pract* 1991; **41**(346): 188–191.
6. Smith HE, Herbert CP. Preventive practice among primary care physicians in British Columbia: relation to recommendations of the Canadian Task Force on the Periodic Health Examination. *CMAJ* 1993; **149**(12): 1795–1800.
7. Shepherd R-M. Clinical obstacles in administrating the South Oaks Gambling Screen in a methadone and alcohol clinic. *J Gamb Stud* 1996; **12**(1): 21–32.
8. Goodyear-Smith F, Arroll B. Screening for domestic violence in general practice: a way forward? *Br J Gen Pract* 2003; **53**(492): 515–518.
9. Goodyear-Smith F, Arroll B, Sullivan S, *et al*. Lifestyle screening: development of an acceptable multi-item general practice tool. *N Z Med J* 2004; **117**(1205): U1146.
10. Bossuyt PM, Reitsma JB, Bruns DE, *et al*. Toward complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative [comment]. *Acad Radiol* 2003; **10**(6): 664–669.
11. Elley CR, Kerse NM, Arroll B. Why target sedentary adults in primary health care? Baseline results from the Waikato Heart, Health, and Activity Study. *Prev Med* 2003; **37**(4): 342–348.
12. Rush A. *Handbook of psychiatric measures*. Washington, DC: American Psychiatric Association, 2000.
13. Shaffer H, Freed C. Assessment of Gambling-Related Disorders. In: Donovan D, Marlatt C (eds). *Assessment of addictive behaviors*. 2nd edn. New York: Guilford Press. 2005: 334–391.