Difference in prevalence of common mental disorder as measured using four questionnaire delivery methods among young people in rural Zimbabwe

Lisa F Langhaug[1], **Yin Bun Cheung**[2], **Sophie Pascoe**[3], **Richard Hayes**[3], and **Frances M Cowan**[1,4]

[1] Royal Free and University College Medical School, London

[2] Singapore Clinical Research Institute, Singapore

[3] London School of Hygiene & Tropical Medicine, London

[4] University of Zimbabwe

Abstract

Background—Previous studies have suggested that interviewer-administered questionnaires can under-estimate the prevalence of depression and suicidal ideation when compared with self-administered ones. We report here on differences in prevalence of reporting mental health between four questionnaire delivery modes.

Methods—Mental health was assessed using the Shona Symptom Questionnaire (SSQ), a locally validated 14-item indigenous measure for common mental affective disorders. A representative sample of 1495 rural Zimbabwean adolescents (median age 18) was randomly allocated to one of four questionnaire delivery modes: self-administered questionnaire (SAQ), SAQ with audio (AASI), interviewer-administered questionnaire (IAQ), and audio computer-assisted survey instrument (ACASI).

Results—Prevalence of common affective disorders varied between QDM (52.3%, 48.6%, 41.5%, and 63.6% for SAQ, AASI, IAQ, and ACASI respectively (p<0.001)). Fewer participants failed to complete SSQ using IAQ and ACASI than other methods (1.6% vs. 12.3%; p<0.001). Qualitative data suggested that respondents found it difficult answering questions honestly in front of an interviewer.

Limitations—Direction of accuracy cannot be ascertained due to lack of objective or clinical assessments of affective disorders.

Conclusions—Estimates of prevalence of psychosomatics symptoms and suicidal ideation varied according to mode of interview. As each mode's direction of accuracy remains unresolved evaluations of interventions continue to be hampered.

Keywords

randomised controlled trial; mental health; measurement; adolescents; Zimbabwe

---

Address for correspondence: Lisa Langhaug, 92 Prince Edward Road, Milton Park, Harare, Zimbabwe, Tel: +263 (0)20 66276 or +263 (0)912 551290, Fax: +263 (0)20 63056, Email: lisa.langhaug@gmail.com.

Introduction

Measurement of mental health disorders including anxiety, depression and suicidal behaviour, using a variety of scales and instruments, has been shown to be affected by the method used for delivering that scale or instrument.(Klimes-Dougan, 1998; Moum, 1998) Research suggests that interviewer-administered instruments often detect lower rates of poor mental health than self-administered instruments.(Kendler et al.,1993; Klimes-Dougan, 1998; Moum, 1998) As a result, self-administered questionnaires, predominantly administered using paper and pen, are those most frequently used. The development of computerized questionnaire administration (Audio Computer-Assisted Survey Instrument (ACASI,)), where computer software allows the subject to hear questions and responses through earphones while simultaneously reading the questions written on the computer screen, has been found to significantly increase reporting of drug use, abortion, same-gender sex, and violent behaviour compared with-interview administered and paper-and-pen self-administered questionnaires.(Fu et al.,2004; Tourangeau and Smith 1996; Turner et al.,1998) While ACASI has shown high acceptability and feasibility in assessing quality of life measurements in low literacy populations,(Thumboo et al.,2006) there is little comparative research on the use of ACASI to capture equally sensitive mental health information. Instead, the evaluation of questionnaire administration has focused predominantly on the comparison of self-administered questionnaires, face-to-face interviews and phone interviews.(Aziz and Kenford 2004; Cheung et al., 2006; Hermens et al.,2006; Holbrook et al.,2003; Klimes-Dougan, 1998; Moum, 1998)

The Regai Dzive Shiri Project is a community randomized trial of a multi-component adolescent reproductive health intervention conducted in rural Zimbabwe. In 2006, we nested an experimental evaluation of four questionnaire delivery modes into the interim survey, in order to compare prevalence of reporting of various stigmatized behaviours. Mental health was assessed using the Shona Symptom Questionnaire (SSQ), a locally validated, 14 item, indigenous measure of common affective disorders.(Patel et al.,1997) The aim of the SSQ is to measure psychiatric morbidity. It was developed and validated among patients attending primary care clinics and traditional medical practitioners in Harare and asks about the presence of various symptoms in the previous week. Using a gold standard defined as diagnosis of a mental disorder by a health care worker *and* scoring 12 or more on the Revised Clinical Interview Schedule, the sensitivity and specificity of the SSQ (using a cut-off of 8 or more out of 14 items) for common affective disorders, were 63% and 83% respectively.(Lewis et al., 1992)

We report here on differences in prevalence of reporting of common affective disorders between four questionnaire delivery modes.

Methods

In 2003, the Regai Dzive Shiri baseline survey was conducted in 30 rural communities in three provinces in eastern rural Zimbabwe(Cowan et al.,2008). All Form 2 pupils (in their ninth year of schooling, median age 15 years)) attending trial secondary schools (n=82) and whose parents/guardians had consented were invited to take part in the baseline survey. Of the 7885 eligible pupils, 6791 (87%) took part. The main reason that young people or their parents declined study participation was due to concerns around blood draw. In 2006, the interim survey, into which the trial reported here was nested, was conducted in 12 of the 30 study communities, selected by restricted randomization to ensure balance between intervention and control arms of the trial and between the three provinces. Young people were eligible to take part in the interim survey if they were cohort members who had previously participated in the baseline survey and were currently residing in these communities. The questionnaire was

carefully designed and piloted prior to administration; cognitive interviews with 65 persons were conducted prior to questionnaire administration to check and ensure comprehension (Mavhu et al.,2008) for this age group. Using a random permuted block design, all interim survey participants were randomly allocated to one of four questionnaire delivery modes: i) self-administered questionnaire (SAQ) using paper and pen, ii) audio-assisted survey instrument (AASI), consisting of SAQ accompanied by an audio soundtrack on a CD player; iii) interviewer-administered questionnaire (IAQ) and; iv) audio computer-assisted survey instrument (ACASI). All participants received detailed training of how to use a method before completing their questionnaire. If the survey assistant felt that they were unable to use a method competently after this training then they were assisted to complete the questionnaire – this happened rarely in only three cases. Analysis was based on mode actually taken. Probable cases of common affective disorders were defined in line with published scale criteria, as participants who answered affirmatively ('always' or 'sometimes') to 8 or more of the 14 statements included in the Shona Symptom Questionnaire (SSQ) (see Table 1 for English phrasing of statements). Probable severe cases were defined as those who scored 11 or more.

Participants with missing values for any of the SSQ items were excluded from analysis. Chi-square tests were used to assess the association of mode of administration with these common affective disorders and with responses to each of the 14 SSQ statements separately. In the event of expected frequencies smaller than 5, an extension of the Fisher's exact test was used.(Mehta and Patel, 1983) When analyzing the 14 statements separately, a nominal P-value <0.05/14=0.003 was considered statistically significant according to the Bonferroni adjustment. Risk ratios (RR) were estimated for probable cases and probable severe cases using SAQ as the reference group.

Participants' perceptions (e.g. ease of completion, sense of privacy during completion, and maintenance of confidentiality) of their questionnaire delivery mode was assessed in three ways. Participants from the last five communities were asked to complete an anonymous post-survey questionnaire that used a five-point Likert Scale to explore their opinions. Qualitative data from two gender-mixed focus group discussions, purposively sampled to reflect all modes and 115 randomly selected study participants described those aspects particular to a mode that enhanced its acceptability. All qualitative data were transcribed electronically and coded for thematic issues using Nvivo 7.0 (QSR, Australia)

Results

Of 1,557 cohort participants still living in the study communities, 1,495 (96%) took part in the interim survey (mean age 18.2 years; range 15–23). Twelve participants were removed from the analysis due to data capture errors. Response rates for completion of the SSQ mental health scale were high; 93.9% of survey participants completed the entire scale. Overall 91 (6.1%) failed to complete one or more SSQ items making it impossible to calculate an SSQ score; this varied by method (SAQ=41 (11.1%); AASI=45 (12.0%); IAQ=5 (1.4%); ACASI=0 (0.0%); p=0.002). Of note this included 14 participants (1.0%) who failed to complete *any* of the SSQ questions (SAQ=6, AASI=8).

As shown in Table 1, the prevalence of common affective disorders as estimated by SAQ, AASI, IAQ, and ACASI were 52.3%, 48.6%, 41.5%, and 63.6% respectively (p<0.001). There was no significant difference in prevalence between AASI and SAQ (RR=0.93; 95% CI: 0.80 to 1.08), whereas IAQ was associated with a lower prevalence (RR=0.79; 95% CI: 0.68 to 0.93) and ACASI with a higher prevalence (RR=1.21; 95% CI: 1.07 to 1.38).

Estimates of prevalence of probable severe cases were 19.6%, 23.6%, 14.2%, and 36.8% for SAQ, AASI, IAQ, and ACASI respectively (p<0.001). Again there was no significant

difference between AASI and SAQ (RR=1.20; 95% CI: 0.90 to 1.61), and as before IAQ gave a lower prevalence (RR=0.73; 95% CI: 0.52 to 1.02) and ACASI a significantly higher reported prevalence than SAQ (RR=1.88; 95% CI: 1.45 to 2.43).

Table 1 shows the percentages of affirmative responses for each of the 14 SSQ statements. All but two statements, "I sometimes saw or heard things which others could not see or hear" (p=0.014), and "I felt I had problems in deciding what to do" (p=0.075) showed a significant difference across modes of administration (each p-value being less than the Bonferroni-adjusted P-value cut-off of 0.003), although the non significant direction of effect was the same as for other questions. Out of the 12 statements that were significantly associated with questionnaire delivery mode, nine showed lowest prevalence when assessed by IAQ and 11 showed highest prevalence when assessed by ACASI.

Post survey, quantitative (post survey questionnaire) and qualitative data (exit interviews and focus group discussion) collected information on mode acceptability and feasibility. Of 697 participants from the last five communities, 650 (93%) completed anonymous post-survey questionnaires with equal completion rates between modes. Of 115 qualitative interviews, 61% were with males. Detailed results are presented elsewhere(Langhaug et al., submitted). Emerging themes focussed on the importance of privacy, ease of method use, and interviewer presence, especially in relation to sensitive questions and questionnaire comprehension. Overall participants stressed the importance of being able to complete the questionnaire 'on their own'. IAQ users expressed difficulty answering sensitive questions. While a few highlighted the benefits of seeking instant clarification from the interviewer, IAQ users predominantly reported feeling embarrassed having to respond to sensitive questions in front of someone as illustrated in the following quotes. *'With [AASI], there is [no question] that you can't answer [truthfully] because no one else ….knows what you said. When you do not see any other person there is nothing to be afraid of, as you would be with [an interviewer]'.* Another participant explained *'but with [an interviewer]. I was just thinking that somebody was watching me.'* By contrast, participants reported feeling more comfortable using the self-completion methods: *'I was telling the truth [with ACASI] because no one was ever going to know what I had written or said.'*.

Discussion

This is one of the first studies to compare the effect of random assignment of questionnaire delivery modes on reporting of common affective disorders conducted amongst young people in developing countries. We found significant differences in prevalence of reporting of this mental ill-health between modes. The interviewer-administered questionnaire was associated with the lowest prevalence of reporting these common affective disorders and ACASI with the highest prevalence. Non-item response rate, often considered a limitation of self-report, also differed significantly between modes of administration with ACASI and IAQ showing an appreciably lower rate of incomplete response than the other two methods.

The post survey data collected to explore the relative acceptability of the various QDM clearly illustrate that participants who completed their questionnaire through an interviewer felt inhibited by their presence, despite the advantage of feeling able to ask questions for clarification. ACASI users stressed the ease and increased sense of confidentiality they felt completing their questionnaire using a computer. Of note, recent improvement in computer programming now allows ACASI to provide additional clarification (albeit standardized) for those who need it(Macalino et al.,2002). Although initially more expensive than other methods, ACASI does not rely on highly skilled interview staff for successful implementation and data-entry time and errors are reduced. While keeping laptops powered in this non-electrified rural setting proved challenging, we overcame this by using solar panels connected to truck batteries.

Other methodological research examining measurement of the prevalence of psychological morbidity has highlighted the importance of questionnaire delivery modes and has emphasized the benefits of self-administered as opposed to interview-administered questionnaires.(Klimes-Dougan, 1998; Moum, 1998) Three of the four methods compared here were self-administered (SAQ, AASI, and ACASI). AASI and ACASI offer the additional benefit over SAQ of allowing subjects to hear the questions through head phones in addition to reading them, which improves understanding. Traditionally, comparative research has explored the difference between in-person and telephone interviews.(Aziz and Kenford 2004; Hermens et al., 2006; Holbrook et al., 2003) However, while the interviewer is not physically present when using the telephone, a real person is distantly 'in attendance.' Here, the voice for AASI and ACASI, an identical recording for both modes, was even more distant in that it was unable to judge the response that was given.

As shown here, estimates of the prevalence of common affective disorders varies substantially according to mode of data collection. For example, the prevalence of self-reported suicidal ideation more than doubled when it was assessed by ACASI (12.0%) compared to when it was interviewer-administered (5.3%). Such uncertainty hampers informed estimation of the potential impact of clinical and preventive health services. In many studies it is assumed that a higher prevalence of self-reported sensitive data reflects more accurate reporting. While this hypothesis makes intuitive sense, evidence using an objective or clinical assessment continues to be lacking. Further research on clarifying the direction of accuracy or relative bias of these modes is urgently needed.

Acknowledgments

Bibliography

Aziz MA, Kenford S. Comparability of Telephone and Face-to-Face Interviews in Assessing Patients with Posttraumatic Stress Disorder. Journal of Psychiatric Practice 2004;10:307–313. [PubMed: 15361745]

Cheung YB, Goh C, Thumboo J, Khoo KS, Wee J. Quality of life scores differed according to mode of administration in a review of three major oncology questionnaires. J Clin Epidemiol 2006;59:185–191. [PubMed: 16426954]

Cowan FM, Pascoe SJS, Langhaug LF, Dirawo J, Chidiya S, Jaffar S, Mbizvo M, Stephenson JM, Johnson AM, Power R, Woelk G, Hayes RJ. The Regai Dzive Shiri Project: a cluster randomised controlled trial to determine the effectiveness of a multi-component community based HIV prevention intervention for rural youth in: study design and baseline results. Trop Med Int Health 2008;13(10): 1235–1244. [PubMed: 18778329]

Fu H, Darroch JE, Henshaw SK, Kolb E. Measuring the extent of abortion underreporting in the 1995 national survey of family growth. Fam Plann Perspect 2004;30:128–133. [PubMed: 9635261]

Hermens MLM, Ader HJ, van Hout HPJ, Terluin B, van Dyck R, de Haan M. Administering the MADRA by telephone or face-to-face: a validity study. Annals of General Psychiatry 2006;5:3–3. [PubMed: 16553958]

Holbrook AL, Green MC, Krosnick JA. Telephone versus Face-to-Face Interviewing of National Probability Samples with Long Questionnaires: Comparisons of Respondent Satisficing and Social Desirability Bias. Public Opin Q 2003:79–125.

Kendler KS, Neale MC, Kessler RC, Heath AC, Eaves LJ. The lifetime history of major depression in women: reliability of diagnosis and heritability. Am J Psychiatry 1993;50:863–870.

Klimes-Dougan B. Screening for suicidal ideation in children and adolescents: methodological considerations. J Adolesc 1998;21:435–444. [PubMed: 9757408]

Langhaug LF, Cheung YB, Pascoe SJS, Chirawu P, Woelk G, Hayes RJ, Cowan FM. How you ask the question really matters: a randomized comparison of four questionnaire delivery modes to assess validity and reliability of self-reported data on sexual behaviour in young people in rural Zimbabwe. 2009(manuscript submitted)

Lewis G, Pelosi A, Araya R. Measuring psychiatric disorder in the community: a standard assessment for use by lay interviewers. Psychol Med 1992;22:465–486. [PubMed: 1615114]

Macalino G, Celentano D, Latkin C, Stathdee C, Vlahov D. Risk behaviours by audio-computer-asisted self interviews among HIV-seropositive and HIV-seronegative injection drug users. AIDS Educ Prev 2002:367–378. [PubMed: 12413183]

Mavhu W, Langhaug LF, Manyonga B, Power R, Cowan FM. What is 'sex' exactly? Using cognitive interviewing to improve validity of sexual behaviour reporting among young people in rural Zimbabwe. Culture, Health, and Sexuality 2008;10:563–572.

Mehta CR, Patel NR. A network algorithm for performing Fisher's exact test in r x c contingency tables. Journal of American Statistics 1983;78:427–434.

Moum T. Mode of administration and interviewer effects in self-reported symptoms of anxiety and depression. Social Indicators Research 1998;45:279–318.

Patel V, Simunyu E, Gwanzura F, Lewis G, Mann A. The Shona Symptom Questionnaire: the development of an indigenous measure of common mental disorders in Harare. Acta Psychiatrica Scandinavia 1997;95:469–475.

Thumboo J, Wee HL, Cheung YB, Machin D, Luo N, Fong KY. Development of a Smiling Touchscreen multimedia program for HRQoL assessment in subjects with varying levels of literacy. Value Health 2006;9:312–319. [PubMed: 16961549]

Tourangeau R, Smith TW. The impact of data collection mode, question format, and question context. Public Opin Q 1996;60:275–304.

Turner CF, Ku L, Rogers SM, Lindberg LD, Pleck JH, Sonenstein FL. Adolescent sexual behavior, drug use, and violence: Increased reporting with computer survey technology. Science 1998;280:867–873. [PubMed: 9572724]

Table 1

Percentages of respondents giving affirmative responses to SSQ (boldface type indicates highest percentage for each statement)

SSQ Scores	SAQ %	AASI %	IAQ %	ACASI %	p value
total number of participants by method	*327*	*331*	*359*	*375*	
overall response rate.	*88.8*	*88.0*	*98.6*	***100.0***	*=0.002*
At risk for common affective disorder (scored 8 or more affirmatively)	*52.3*	*48.6*	*41.5*	***63.5***	*<0.001*
Severely at risk for common affective disorder (scored 11 or more affirmatively)	*19.6*	*23.6*	*14.2*	***36.8***	*<0.001*
SSQ Statements					
There were times in which I was thinking deeply or thinking about many things (thinking too much).	77.7	74.3	70.8	**87.2**	<0.001
I found myself sometimes failing to concentrate	61.2	68.6	54.6	**81.3**	<0.001
I lost my temper or got annoyed over trivial matters	54.1	56.5	59.9	**73.6**	<0.001
I had nightmares or bad dreams	68.2	66.2	66.9	**78.4**	<0.001
I sometimes saw or heard things which others could not see or hear	20.8	26.6	25.1	**31.5**	0.014
My stomach was aching	60.9	56.2	52.9	**69.1**	<0.001
I was frightened by trivial things	44.0	38.1	35.9	**52.0**	<0.001
I sometimes failed to sleep or lost sleep	50.8	51.4	41.8	**59.7**	<0.001
There were moments when I felt life was so tough that I cried or wanted to cry	58.7	61.6	52.4	**72.5**	<0.001
I felt run down (tired)	68.2	61.3	59.1	**77.1**	<0.001
At times I felt like committing suicide	8.3	**13.0**	5.3	12.0	0.001
I was generally unhappy with things that I would be doing each day	48.3	44.1	43.7	**57.3**	<0.001
My work was lagging behind (impairment of functioning)	40.4	47.4	42.3	**53.6**	0.002
I felt I had problems in deciding what to do	52.9	54.1	50.4	**59.7**	0.075