

Are computerized interviews equivalent to human interviewers? CIDI-Auto *versus* CIDI in anxiety and depressive disorders

L. PETERS¹, D. CLARK AND F. CARROLL

*From the Clinical Research Unit for Anxiety Disorders, St Vincent's Hospital,
Darlinghurst, NSW, Australia*

ABSTRACT

Background. The equivalency of the Composite International Diagnostic Interview delivered by human interviewers (CIDI) and its computerized version (CIDI-Auto) was examined for anxiety and depressive disorders.

Methods. Subjects were 40 patients at an Anxiety Disorders Clinic and 40 general medical practice attenders. The CIDI-Auto and CIDI were administered in counterbalanced order on the same day and measures of computer attitudes and the acceptability of the two interview formats were also taken.

Results. The CIDI-Auto and the CIDI were found to be equally acceptable to subjects on the dimensions of comfort and preference, while the CIDI-Auto was rated as less embarrassing but too long in comparison with the CIDI. The agreement between the two formats was acceptable with kappa values for ICD-10 diagnoses being above 0.65 and for DSM-III-R diagnoses above 0.5 except for two diagnoses (generalized anxiety disorder and dysthymia). Discrepancies between the two formats were predicted by computer attitudes and not by computer experiences or the tendency to respond in a socially desirable fashion.

Conclusions. It is concluded that the CIDI-Auto in its self-administered form is an acceptable substitute for the CIDI for suitable subjects.

INTRODUCTION

The Composite International Diagnostic Interview (World Health Organization, 1993*a*, 1997*a*) is a fully structured interview that is amenable to administration by people who are not trained to make psychiatric diagnoses. It has been used, in its various forms, in research studies and in particular it has been used in large epidemiological studies (Robins & Regier, 1991; Kessler *et al.* 1994). Its reliability and validity have been established (see Wittchen, 1994 and Andrews & Peters, 1997 for reviews). Given its high level of structure, the CIDI is amenable to computerization and the WHO computerized version of the CIDI, known as CIDI-Auto

(World Health Organization, 1993*b*, 1997*b*) has been in use since 1993. The CIDI-Auto can be administered by an interviewer who reads the questions from the computer screen and codes the answers directly into the computer or it can be self-administered by the respondent.

The computerized version of the CIDI has a number of supposed advantages over the standard administration of the CIDI: interviewer error is minimized by the presentation of questions in a standard manner, direct entry of data, and the self-administration of the CIDI-Auto may allow subjects more privacy and anonymity to reveal sensitive material (see Lautenschlager & Flaherty, 1990, for a discussion of this last issue). The CIDI-Auto has been shown to be test-retest reliable when administered by an interviewer (Peters, 1997). In addition, the self-administered version has been

¹ Address for correspondence: Dr Lorna Peters, Clinical Research Unit for Anxiety Disorders, 299 Forbes Street, Darlinghurst, NSW 2010, Australia.

compared to the best possible clinical diagnosis and, although CIDI-Auto makes more diagnoses per subject than the clinicians, the discriminations it makes are similar, attesting to its procedural validity (Peters & Andrews, 1995). Despite its supposed advantages, its reliability and validity, the question remains whether the self-administered version of the CIDI can be substituted for a human interviewer delivering the CIDI as it was originally developed.

Evidence from studies of other diagnostic interviews (Greist *et al.* 1987; Lewis *et al.* 1988; Erdman *et al.* 1992; Lewis, 1994) shows that the agreement between computerized versions and standard administrations is acceptable. Administration of the Diagnostic Interview Schedule (DIS) by computer and by human interviewer an average of 3.2 days apart resulted in an average kappa of 0.51 for 15 diagnoses (Greist *et al.* 1987). A similar average kappa of 0.57 for agreement between the computerized and the standard DIS was found when the interviews were administered within a 6-week period (Erdman *et al.* 1992). These values are similar to those found for the DIS given in its standard human-interviewer format by lay interviewers *versus* the DIS given by psychiatrists (Robins *et al.* 1981). Another psychiatric interview, the Clinical Interview Schedule, has been found to have acceptable agreement with its computerized version in two studies ($\kappa > 0.50$; Lewis *et al.* 1988; $\kappa = 0.70$; Lewis, 1994) when the two versions of the interview were administered on the same day. Similarly, computerized assessments have been found to provide similar results to their standard paper-and-pencil versions in the domain of attitude testing (Finegan & Allen, 1994) and in the domain of alcohol and drug history interviews (Skinner & Allen, 1983).

Despite good agreement between computerized and standard administrations of diagnostic interviews and psychological tests, there are concerns that the results given by the two modes of administration are not identical. George *et al.* (1992) found that scores on the Beck Depression Inventory and on the State-Trait Anxiety Inventory were higher when these tests were administered by computer than when administered as a standard paper-and-pencil test. Similarly, Lautenschlager & Flaherty (1990) found that 'impression management' scores were higher when the test was administered as a

computerized interview rather than as a paper-and-pencil test. It has been suggested that a number of factors may affect responses to computerized interviews including attitudes towards computers, particularly anxiety about using computers, the amount of computer experience the subject has and the tendency for the subject to respond in a socially desirable fashion. In particular, it has been found that subjects with high levels of computer anxiety have altered scores on tests administered as a computerized test as opposed to a pencil-and-paper test, whereas low levels of computer anxiety do not have such an impact on scores (Lankford *et al.* 1994). Mathisen *et al.* (1985) found that education and previous computer experience was related to the ease with which a computerized interview (the DIS) was completed: those with more experience were more able to complete the interview. The tendency to respond in a socially desirable fashion has been found to relate to the number of discrepancies found between a traditionally administered DIS and the computerized version of the DIS (Erdman *et al.* 1992).

The aim of this study was to compare the self-administered CIDI-Auto with the CIDI delivered by a human interviewer. The effects of computer attitudes, computer experience and social desirability on the answers given to the computerized and the standard administration of the CIDI were examined.

METHOD

Subjects

Subjects were 40 consecutive (30 female, 10 male) patients accepted for treatment at the Clinical Research Unit for Anxiety Disorders and 40 volunteers (30 female, 10 male) who had attended the office of one of four general medical practices. The decision to include subjects recruited at general medical practices was made so that a broad range of subjects was included, both clinical and non-clinical, to enhance the generalizability of the findings. The average age of the whole sample was 32.9 (range = 18–82; s.d. = 14.2), of the clinical sample was 33.5 (range = 19–56; s.d. = 9.7) and of the medical practice attenders was 32.3 (range = 18–82; s.d. = 17.7). There was no difference in the age ($t(1, 78) = 0.39$; $P = 0.70$) of the two samples.

Interviewers

Two interviewers (1 male, 1 female) administered the standard paper-and-pencil CIDI. Both had 4 years of undergraduate training in psychology and both had been trained in CIDI administration at the WHO CIDI Training and Reference Centre in Sydney. Before gathering data for the study, the interviewers were observed by one of the trainers of the CIDI Training and Reference Centre (L.P.) until there was 100% agreement between the interviewer and the trainer in the coding of the interviews.

Materials

The Composite International Diagnostic Interview

Version 1.1 of the CIDI (World Health Organization, 1993a) was used in this study. The anxiety and depressive disorders section of the paper-and-pencil interview were used. The computerized data entry and scoring program for CIDI Version 1.1 (World Health Organization, 1993c) was used to provide DSM-III-R and ICD-10 diagnoses from the paper-and-pencil interview. Version 1.1 of the CIDI-Auto (World Health Organization, 1993b) was also used. The CIDI-Auto is the WHO approved computerized version of the paper-and-pencil CIDI. The questions in the CIDI-Auto faithfully represent the questions in the paper-and-pencil version of the CIDI, with the skip patterns and probe flow chart questions being implemented by the program. The self-administered form of the CIDI-Auto was used: the subject sits at the computer terminal and moves through the interview at his/her own pace. A number of tutorial screens, prior to the interview, educate the subjects about how to move through the interview and about the types of questions they will be asked. The tutorial provides interactive examples so that the interview administrator can ensure that the subjects understand the instructions on the screen. The data generated by CIDI-Auto was scored by the same program, used for scoring the paper-and-pencil version of the CIDI for DSM-III-R and ICD-10 diagnoses (World Health Organization, 1993c). For both the paper-and-pencil and the computerized version of the CIDI, hierarchical rules were not applied in scoring the CIDI-Auto data for DSM-III-R diagnoses.

Computer attitudes and computer experience

A 20-item measure of computer attitudes was constructed using items from the Computer Attitudes Scale (Blouin *et al.* 1988) and from the Computer Anxiety Index (CAIN; Simonson *et al.* 1987). The items addressed general attitudes towards computers (e.g. 'I feel a computer is too impersonal to give information to', 'I look forward to a time when computers are more widely used'), perceived computing ability (e.g. 'I am good at using computers', 'The operation of computers is easy') and anxiety about using computers (e.g. 'Using a computer makes me tense', 'I avoid using computers when I can'). Items were scored on a 5-point scale from 'strongly agree' to 'strongly disagree'.

Computer experience was measured using four questions that addressed the frequency of use of computers ('almost every day', 'more than once a week', 'more than once a month', 'at least once a year', 'not at all'), the duration of use of computers ('today will be my first time', 'less than 6 months', '6 months to 1 year', 'more than one year'), a description of the level of computer ability ('no computer training or experience', 'limited training or experience', 'considerable training and experience (some programming skills)', 'proficient computer programming skills (passed a computer programming course)') and whether the subjects had a computer at home.

Social desirability

Social desirability was measured by the 10-item Marlowe-Crowne Scale (Greenwald & Satow, 1970).

Interview acceptability

Acceptability of the computerized interview was measured by asking subjects to rate their experience of the computerized and paper-and-pencil interview. Four questions asked which interview was preferred, which was more embarrassing, which seemed to take longer, and with which interview the subject felt more comfortable.

Design and procedure

Subjects in the clinical sample (at the anxiety disorders clinic) completed the CIDI-Auto as part of the routine assessment procedure.

Subjects were invited to participate in a study to investigate the comparability of computerized and standard (paper-and-pencil) interviews. None of the subjects who were approached to participate refused. Subjects in the non-clinical sample were similarly invited to participate in the study being given the same instructions as subjects in the clinical sample. Non-clinical subjects were paid a nominal fee for their participation. One in two subjects who were approached to participate refused.

Subjects completed the CIDI as a self-administered computerized interview (CIDI-Auto) and as delivered by a trained CIDI interviewer using the standard paper-and-pencil interview (CIDI) in counterbalanced order on the same day. The anxiety disorders and depressive disorders sections of the paper-and-pencil interview were administered to all subjects. All subjects completed the anxiety disorders and depressive disorders sections of the computerized interview and the clinical sample completed, in addition, the substance use disorders section (these data are not reported in this study). Subjects were told before the second interview that answers to the previous interview were not known and that they should answer the questions in the second interview honestly, regardless of whether answers were the same or different to those given in the previous interview. One of the interviewers was available at all times during the computerized interview to assist the subjects with difficulties that they experienced with the interview, but did not sit with the subjects while they completed the interview. Thus, interviewers were blind to the subjects' responses to the computerized interview. The measures of computer attitudes and computer experiences were given before the first interview and measures of acceptability of the interview and social desirability were administered after the second interview.

RESULTS

Number of diagnoses

Among the 80 subjects, there was a total of 167 DSM-III-R and 169 ICD-10 diagnoses generated by the paper-and-pencil CIDI and a total of 172 DSM-III-R and 159 ICD-10 diagnoses generated by the computerized CIDI-Auto. Twenty sub-

jects did not receive a DSM-III-R diagnosis after the computerized interview and 19 subjects after the paper-and-pencil interview; 19 subjects did not receive an ICD-10 diagnosis after the computerized interview and 21 subjects after the paper-and-pencil interview. For the paper-and-pencil interview, there was an average of 2.2 (s.d. = 1.9; 95% CI 1.73, 2.57) DSM-III-R and 2.0 (s.d. = 1.9; 95% CI 1.57, 2.41) ICD-10 diagnoses per subject. For the computerized interview there was an average of 2.1 (s.d. = 1.7; 95% CI 1.70, 2.48) DSM-III-R and 2.1 (s.d. = 1.8; 95% CI 1.71, 2.52) ICD-10 diagnoses per subject. Thus, there were the same number of DSM-III-R ($t(1, 79) = -0.6$; $P = 0.56$) and ICD-10 ($t(1, 79) = 1.2$; $P = 0.25$) diagnoses generated by the paper-and-pencil and computerized versions of the CIDI. This pattern was observed whether the subject was from the clinical sample or from the non-clinical sample.

Number of symptoms

Given the skip patterns in the CIDI, not all items are asked of all subjects. There are, however, in the anxiety and depressive disorders sections of CIDI 1.1, 61 (23 anxiety, 38 depression) primary items i.e. symptom questions which are asked of all subjects. For the paper-and-pencil interview, there was an average of 19.95 (s.d. = 11.24; 95% CI 17.44, 22.46) primary items endorsed and for the computerized interview, there was an average of 21.16 (s.d. = 11.95; 95% CI 18.49, 23.84) primary items endorsed. This difference was significant ($t(1, 79) = 3.68$; $P < 0.0001$). As expected given the source of the subjects, subjects from the clinical sample endorsed more primary items on both the paper-and-pencil (mean (M) = 24.73; s.d. = 10.79; 95% CI 21.31, 28.14) and the computerized interview (M = 26.48; s.d. = 11.52; 95% CI 22.83, 30.12) than did the non-clinical sample (paper-and-pencil M = 15.18, s.d. = 9.62, 95% CI 12.13, 18.22, $t(1, 78) = 4.18$; $P < 0.0001$; computerized M = 15.85, s.d. = 9.95; 95% CI 12.71, 18.99, $t(1, 78) = 4.42$; $P < 0.0001$).

Agreement

Diagnoses

The agreement between the CIDI-Auto and the CIDI at the diagnostic level is shown in Table 1. As measured by percentage agreement, the CIDI

Table 1. Agreement between the CIDI-Auto and CIDI for DSM-III-R and ICD-10 diagnoses

Diagnosis	DSM-III-R		ICD-10	
	Agreement (%)	Kappa	Agreement (%)	Kappa
Panic disorder	88.8	0.54	92.5	0.83
Panic/agoraphobia	93.8	0.84	—	—*
Agoraphobia	—	—†	91.3	0.80
Social phobia	96.3	0.92	87.5	0.65
Specific phobia	83.8	0.64	96.3	0.71
OCD	88.8	0.58	—	—†
GAD	88.8	0.46	87.5	0.67
Dysthymia	82.5	0.43	93.8	0.75
Depression	82.5	0.65	90.0	0.80

* Diagnosis not made in ICD-10.

† Base rate of the disorder in this sample is less than 5%.

and CIDI-Auto agreed well: for ICD-10 diagnoses the agreement between CIDI-Auto and CIDI was over 90% except for Social Phobia (87.5%) and GAD (87.5%); for DSM-III-R diagnoses, the agreement was over 90% for three diagnoses and over 80% for the remaining diagnoses. The chance-corrected agreement (kappa) values show that the agreement between CIDI and CIDI-Auto was, for the majority of diagnoses, good to excellent: for ICD-10 diagnoses all diagnoses had kappa values showing good to excellent agreement between CIDI-Auto and CIDI ($\kappa > 0.60$) and for DSM-III-R diagnoses four diagnoses had kappa values showing good to excellent agreement (Panic Disorder with or without Agoraphobia, Social Phobia, Specific Phobia, and Major Depressive Episode) while the remaining diagnoses had kappa values indicating acceptable levels of agreement. Note that kappas were not calculated for diagnoses which had a base rate in this sample of less than 5% because kappa is known to be unstable at low base rates (Spitznagel & Helzer, 1985).

Number of discrepancies on primary items

To examine the agreement between the CIDI-Auto and CIDI further, the number of discrepancies between the two interviews on the primary items of the CIDI was calculated. Primary items are, in general, answered by 'yes' or 'no'. A discrepancy on a primary item means that for one interview the answer was 'yes' and for the other it was 'no'. For the whole sample,

the mean number of discrepancies was 4.94 (s.d. = 3.44; range 0–14). While the number of discrepancies was the same for the clinical (M = 5.25; s.d. = 3.40; 95% CI 4.18, 5.64) and non-clinical (M = 4.63; s.d. = 3.51; 95% CI 3.52, 5.73; $t(1, 78) = 0.81$; $P = 0.42$) subjects, there was a non-significant tendency for subjects who completed the computerized interview before the interview with a person (M = 5.61; s.d. = 3.57; 95% CI 4.45, 6.76) to have more discrepancies than those who completed the interview with the person before the computerized interview (M = 4.33; s.d. = 3.25; 95% CI 3.33, 5.34; $t(1, 78) = 1.67$; $P = 0.09$).

Acceptability

The acceptability of the different formats for presentation of the CIDI is shown in Fig. 1. While equal numbers of subjects said that they felt more comfortable with the computerized interview (26%) and the interview with a person (25%), more subjects said they were equally comfortable with both formats (49%; $\chi^2(2) = 8.58$; $P = 0.01$). Very few subjects said that they found the computerized interview more embarrassing (3.8%) compared with the number who said they found the interview with the person more embarrassing (32.5%). Most subjects (63.8%) said they found neither interview more embarrassing ($\chi^2(2) = 43.23$, $P = 0.001$). While there was a tendency for more subjects to say they preferred the interview with the person (45%) compared with the number that said they preferred the computerized interview (25%) or the number who had no preference (30%), this difference was not significant ($\chi^2(2) = 5.20$, $P = 0.07$). The majority of the subjects said that the computerized interview (56.3%) seemed to take longer, while 31.3% said the interview with the person seemed to take longer and only 12.5% said that both seemed to take the same amount of time ($\chi^2(2) = 23.13$, $P = 0.001$). Although this may be an artefact of the fact that the subjects from the Anxiety Disorders Unit did do more sections of the CIDI-Auto than of the CIDI, the pattern of ratings for the 40 subjects from the primary care setting who completed the same sections of the CIDI and CIDI-Auto was similar: 50% said the computerized interview seemed to take longer, 35% said the interview with the person seemed to take longer, and 15% said the two

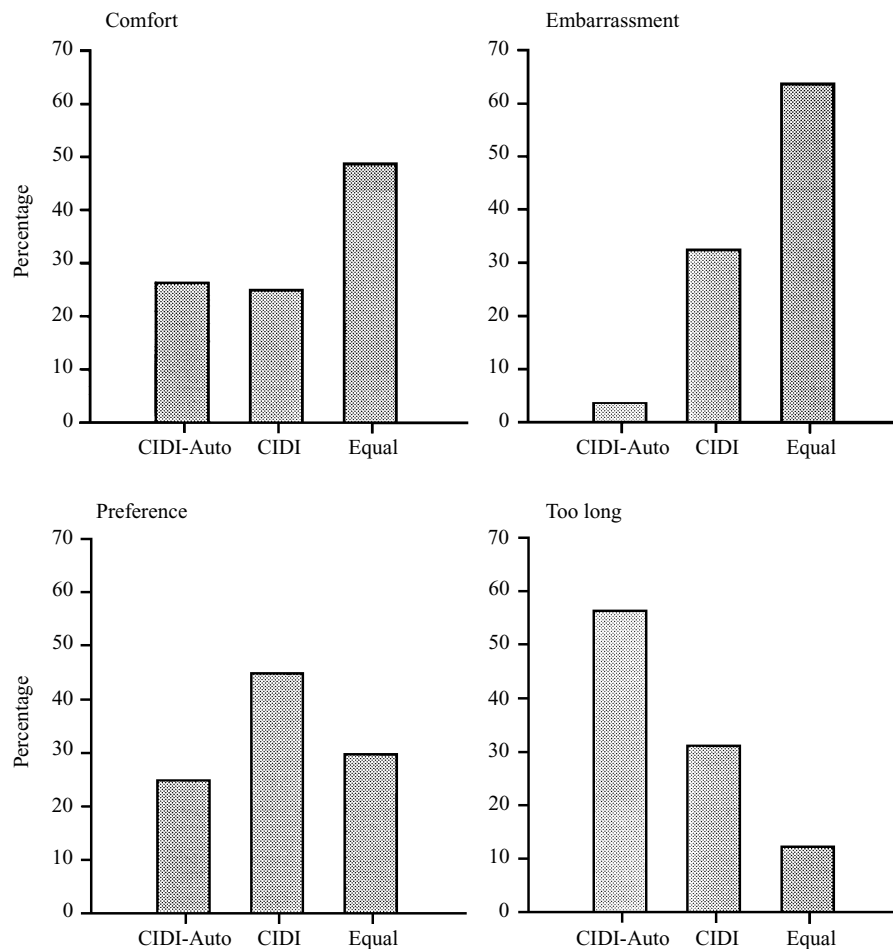


FIG. 1. Acceptability of the CIDI-Auto versus CIDI (percentage endorsed).

formats seemed to take the same amount of time. The ratings of subjects regarding the length of interview were reflected in slightly longer average administration times for the anxiety and depression sections of the CIDI-Auto compared with the CIDI (CIDI-Auto $M = 32.2$ min, $s.d. = 17.6$, 95% CI 28.19, 36.21; CIDI $M = 28.9$ min, $s.d. = 13.2$, 95% CI 25.94, 31.96; $t(1, 76) = 2.10$, $P = 0.04$).

Influence of computer attitudes, computer experiences and social desirability

To determine whether computer anxiety, computer experiences or social desirability had an impact on the responses to the CIDI given in the different ways, a standard multiple regression

was performed. The independent variables (IVs) were the scores on the computer attitudes, computer experience and social desirability measures, and the dependent variable was the number of discrepancies between the two formats for the primary items in the CIDI (see above for a description of the calculation of this variable). Analysis was performed using SPSS REGRESSION. R for regression ($R = 0.43$) was significantly different from zero ($F(3, 74) = 5.77$; $P = 0.0013$). The three IVs accounted for 19% of the variance ($R^2 = 0.19$). Only computer attitudes ($sr^2 = 0.09$; $P = 0.0057$) had a significant unique contribution to R^2 (computer experience: $sr^2 = 0.0001$, $P = 0.9145$; social desirability: $sr^2 = 0.04$, $P = 0.0611$).

DISCUSSION

The results of this study suggest that the CIDI-Auto, in its self-administered form, is an acceptable substitute for a human interviewer delivering the CIDI in its traditional paper-and-pencil format when the anxiety and depressive disorders are being assessed. The number of diagnoses produced by the CIDI-Auto is similar to the number of diagnoses produced by the CIDI, despite the number of symptoms endorsed being higher for the CIDI-Auto than for CIDI. In addition, the agreement between the computerized version of the CIDI and the CIDI administered in its standard fashion is acceptable: the diagnoses produced by the two modes of administration are similar suggesting that the responses given to the questions that make up those diagnoses were similar in most cases. In fact, the number of discrepancies on the primary items is very low (an average of 4.94 items, out of a possible 61, were discrepant). The agreement at a diagnostic level is not, however, perfect. This is particularly surprising given that the two types of CIDI interview were conducted on the same day. The kappa values in this study were similar to those found in a study of the test-retest reliability of the interviewer-administered CIDI when the administrations were one week apart ($\kappa = 0.39$ (specific phobia) to $\kappa = 0.75$ (agoraphobia); Peters, 1997) and to those found in test-retest reliability studies of the CIDI when conducted 3 days apart ($\kappa = 0.52$ (dysthymia) to $\kappa = 0.84$ (panic disorder); reported in Wittchen, 1994). This suggests that the changes in answers between time 1 and time 2 in these test-retest studies are not likely to be due to changes in symptoms, but rather may reflect a change in the understanding of the question having heard it before.

The computerized version of the CIDI was found to be an acceptable form of administration. Subjects felt equally comfortable with the two forms of administration. Comparable to the findings of Greist *et al.* (1987) for the DIS, fewer subjects in the present study said they felt embarrassed in revealing information about their symptoms to the computerized interview when compared with the human interviewer. This finding is particularly important given the context in which the study was conducted. Half

of the subjects had reported to an anxiety disorders unit for treatment of an anxiety disorder, and thus, subjects knew that the interviewer was aware of their clinical status. Nevertheless, subjects still reported that the computerized interview was not as embarrassing as the standard administration of the CIDI. It may be argued that the characteristics of the particular interviewers caused this difference. The number of subjects reporting that the computerized interview was less embarrassing than the interview with the person, did not vary, however, across the two interviewers ($\chi^2 = 1.11$; $P = 0.58$). In addition, the same number of diagnoses were elicited by the human interviewers as by the computerized interview.

Although the anxiety and depressive disorders section of the CIDI took longer to complete in the self-administered computerized version, about half of the subjects said the interview with the person was too long or that the two types of interview took the same amount of time. The longer administration time for the computerized CIDI may well reflect the fact that subjects felt more able to take their time to understand the questions (cf. Greist *et al.* 1987). In addition, the longer time taken to complete the self-administered CIDI-Auto must be balanced against the advantages of not requiring a trained interviewer to administer the CIDI.

Consistent with previous studies (e.g. Lankford *et al.* 1994), the results of this study suggest that attitudes towards computers may have an impact on the responses made to a computerized interview. The number of discrepancies between the computerized CIDI and the standard administration of the CIDI was predicted by the scores on the computer attitudes measure, but not by computer experience or by the tendency to respond in a socially desirable manner. Since the data for acceptability of the interview show that more subjects found the interview with the person more embarrassing than the computerized interview, it is surprising that the tendency to respond in a socially desirable fashion did not predict discrepancies. Perhaps subjects responded in a socially desirable fashion to both interviews, or more likely, the setting in which the research was conducted (a clinic for anxiety disorders and medical practitioners' offices) mitigated against

socially desirable responding. This finding is different from that reported by Erdman *et al.* (1992) who found that scores on the Marlowe–Crowne Scale were significantly related to discrepancies between the computerized and traditional DIS.

Where there are discrepancies between CIDI-Auto and CIDI, the present study cannot answer the question of which is the correct diagnosis, that produced by CIDI-Auto or that produced by CIDI. Neither interview can be seen as the gold standard against which to compare the other. In fact, gold standards for validity do not exist in psychiatry and have been the subject of discussion for many years (e.g. Spitzer, 1983). The absence of gold standards has meant that the validity of structured diagnostic interviews in psychiatry is often examined by comparison with good quality clinical diagnoses, such as a LEAD standard (Spitzer, 1983). A procedural validity study of the self-administered version of CIDI-Auto, comparing the CIDI-Auto diagnoses to a LEAD standard clinical diagnosis (Peters & Andrews, 1995), conducted in the same environment as the present study, has shown that the CIDI-Auto has acceptable validity.

One of the tentative findings of this study was that the order of administration of the two types of interview had a tendency to affect the number of discrepancies on primary items. Subjects who completed the CIDI-Auto before the CIDI tended to have more discrepancies than those who completed the interviews in the opposite order. Although this difference did not reach statistical significance, it may well point to the fact that subjects understood the same questions in a different way when presented on a computer screen and when presented by a human interviewer who was able to emphasize the important features of the questions and interpret whether the answers met the intent of the question. This may well explain the kappa values for diagnoses being less than perfect although the interviews were administered on the same day. Future research may examine, using cognitive interviewing techniques (cf. Schwarz & Sudman, 1996; Sudman *et al.* 1996), the way in which CIDI questions are understood.

How the interviewer-administered form of the CIDI-Auto would perform is unknown in this study. Erdman *et al.* (1992) compared the

interviewer-administered version of the computerized DIS to both the self-administered computerized DIS and the traditional human interviewer DIS and found few differences between the three modes of presentation. Given the similarities between the DIS and the CIDI, it can be assumed that the interviewer-administered version of the CIDI-Auto is likely to produce similar results to those produced by the CIDI, although this could be the subject of future research.

The findings from this study will need to be replicated in other samples, especially those with more severe psychiatric disorders. Two types of subject were included in the present study in order to enhance generalizability of the findings: a clinical sample reporting for treatment at an anxiety disorders clinic and a 'non-clinical' sample from primary-care settings. The finding that the mean number of discrepancies on primary symptom items is the same for the clinical and non-clinical samples in this study suggests that the CIDI-Auto and CIDI are comparable at least in these two settings.

In conclusion, the self-administration of CIDI-Auto was found to be equivalent to the traditional delivery of the CIDI by a human interviewer. In addition, the computerized version of the CIDI was acceptable to respondents. Given that large epidemiological surveys are being conducted using the CIDI and instruments like it, and increasingly these surveys are using the computerized version of the instrument, the results of the present study are reassuring in that the prevalence rates found with the standard instrument are likely to be similar to those found using the computerized interview, all other things being equal.

This study was supported by a grant from the New South Wales Department of Health.

REFERENCES

- Andrews, G. & Peters, L. (1997). Psychometric properties of the CIDI. *Social Psychiatry* (in the press).
- Blouin, A. G., Perez, E. L. & Blouin, J. H. (1988). Computerized administration of the Diagnostic Interview Schedule. *Psychiatry Research* **23**, 335–344.
- Erdman, H. P., Klein, M. J., Greist, J. H., Skare, S. S., Husted, J. J., Robins, L. N., Helzer, J. E., Goldring, E., Hamburger, M. & Miller, J. P. (1992). A comparison of two computer-administered

- versions of the NIMH Diagnostic Interview Schedule. *Journal of Psychiatric Research* **26**, 85–95.
- Finegan, J. E. & Allen, N. J. (1994). Computerized and written questionnaires: are they equivalent? *Computers in Human Behavior* **10**, 483–496.
- Greenwald, H. J. & Satow, Y. (1970). A short social desirability scale. *Psychological Reports* **27**, 131–135.
- Greist, J. H., Klein, M. H., Erdman, H. P., Bires, J. K., Bass, S. M., Machtinger, P. E. & Kresge, D. G. (1987). Comparison of computer- and interviewer-administered versions of the Diagnostic Interview Schedule. *Hospital and Community Psychiatry* **38**, 1304–1310.
- Kessler, R. C., McGonagle, K. A., Zhao, S., Nelson, C. B., Hughes, M., Eshleman, S., Wittchen, H.-U. & Kendler, K. (1994). Lifetime and 12-month prevalence of DSM-III-R psychiatric disorders in the United States: results of the National Comorbidity Survey. *Archives of General Psychiatry* **51**, 8–19.
- Lankford, J. S., Bell, R. W. & Elias, J. W. (1994). Computerized versus standard personality measures: equivalency, computer anxiety, and gender differences. *Computers in Human Behavior* **10**, 497–510.
- Lautenschlager, G. J. & Flaherty, V. L. (1990). Computer administration of questions: more desirable or more social desirability? *Journal of Applied Psychology* **75**, 310–314.
- Lewis, G. (1994). Assessing psychiatric disorder with a human interviewer or a computer. *Journal of Epidemiology and Community Health* **48**, 207–210.
- Lewis, G., Pelosi, A. J., Glover, E., Wilkinson, G., Stansfeld, S. A., Williams, P. & Shepherd, M. (1988). The development of a computerized assessment for minor psychiatric disorder. *Psychological Medicine* **18**, 737–745.
- Mathisen, K. S., Evans, F. J., Meyers, K. & Kogan, L. (1985). Human factors influencing patient-computer interaction. *Computers in Human Behavior* **1**, 163–170.
- Peters, L. (1997). Test-retest reliability of the computerised Composite International Diagnostic Interview (CIDI-Auto): anxiety and depressive disorders. (In preparation.)
- Peters, L. & Andrews, G. (1995). Procedural validity of the computerized version of the Composite International Diagnostic Interview (CIDI-Auto) in the anxiety disorders. *Psychological Medicine* **25**, 1269–1280.
- Robins, L. N. & Regier, D. A. (1991). *Psychiatric Disorders in America*. Free Press: New York.
- Robins, L. N., Helzer, J., Croughan, J. & Ratcliff, K. (1981). National Institute of Mental Health Diagnostic Interview Schedule: its history, characteristics, and validity. *Archives of General Psychiatry* **38**, 381–389.
- Schwarz, N. & Sudman, S. (1996). *Answering Questions: Methodology for Determining Cognitive and Communicative Processes in Survey Research*. Jossey-Bass: San Francisco.
- Simonson, M. R., Maurer, M., Montag-Torardi, M. & Whitaker, M. (1987). Development of a standardised test of computer literacy and a computer anxiety index. *Journal of Educational Computing Research* **3**, 231–247.
- Skinner, H. A. & Allen, B. A. (1983). Does the computer make a difference? Computerized versus face-to-face versus self-report assessment of alcohol, drug, and tobacco use. *Journal of Consulting and Clinical Psychology* **51**, 267–275.
- Spitzer, R. L. (1983). Psychiatric diagnosis: are clinicians still necessary? *Comprehensive Psychiatry* **24**, 399–411.
- Spitznagel, E. L. & Helzer, J. E. (1985). A proposed solution to the base rate problem in the kappa statistic. *Archives of General Psychiatry* **42**, 725–728.
- Sudman, S., Bradburn, N. M. & Schwarz, N. (1996). *Thinking about Answers: The Application of Cognitive Processes to Survey Methodology*. Jossey-Bass: San Francisco.
- Wittchen, H.-U. (1994). Reliability and validity studies of the WHO Composite International Diagnostic Interview: a critical review. *Journal of Psychiatric Research* **28**, 57–84.
- World Health Organization (1993a). *Composite International Diagnostic Interview (Core Version 1.1)*. APPI: Washington, DC.
- World Health Organization (1993b). *CIDI-Auto Version 1.0: Administrator's Guide and Reference*. Training and Reference Centre for WHO CIDI: Sydney.
- World Health Organization (1993c). *CIDI Core Computer Manual for Data Entry and Diagnostic Programs for the CIDI (Version 1.1)*. WHO: Geneva.
- World Health Organization (1997a). *Composite International Diagnostic Interview (Core Version 2.1)*. WHO: Geneva.
- World Health Organization (1997b). *CIDI-Auto Version 2.1: Administrator's Guide and Reference*. Training and Reference Centre for WHO CIDI: Sydney.