

Use of Video Conferencing for Psychiatric and Forensic Evaluations

Frances J. Lexcen, Ph.D.

Gary L. Hawk, Ph.D.

Steve Herrick, Ph.D.

Michael B. Blank, Ph.D.

Objective: This study investigated whether the quality of results from video interviews is comparable with that of in-person interviews. **Methods:** Interrater reliabilities for two video conference interview conditions were compared with those for in-person interviews with the Brief Psychiatric Rating Scale–Anchored Version and the MacArthur Competence Assessment Tool–Criminal Adjudication, given to 72 forensic inpatients. The video conditions included in-person and remote interviewers. In the first condition, an in-person interviewer administered the instruments, with remote observation and scoring. The second condition entailed remote administration and an in-person observer. The third condition used an in-person interviewer and observer. **Results:** Good to excellent reliabilities resulted from all conditions with intraclass correlations of .69 to .82. **Conclusions:** Results suggest that pro-

viders can expect remote interviews to provide clinical information similar to that obtained by in-person interviews. (*Psychiatric Services* 57:713–715, 2006)

Video conferencing offers the possibility of improved services to patients with limited access to care. This mode of communication has become feasible as the prices of hardware, software, and video equipment have declined. Forensic evaluations of secured populations demonstrate the financial benefits of video assessments because these patients often require evaluations to protect their due-process rights, but the number of clinicians trained to do forensic evaluations is limited. Video conferencing offers a way to reduce the risks of transporting defendants to evaluations and the costs associated with security escorts.

One question in a reliability study of telemedicine is whether results obtained from video conferencing are similar to results from face-to-face interviews. This study produced intraclass correlations (ICCs) for interviews conducted by video and in person.

Using standardized instruments for video studies is one method of minimizing differences between in-person and remote interviews. Several studies of video conferencing have used the Brief Psychiatric Rating Scale–Anchored Version (BPRS-A), alone or with other measures (1). Zarate and colleagues (2) compared the BPRS-A using 15 live simultaneous interviews and remote administrations of single interviews. ICCs for most BPRS-A items ranged from .01 to .96, although the authors noted that correlations

could not be calculated for five comparisons because of the lack of variance between raters.

Using a similar experimental design, Chae and colleagues (3) compared single scoring of 15 live interviews with results from 15 interviews conducted via video with an observer located on site. Agreement was similar in both conditions for three items, greater in the video interviews for eight items, and greater in the in-person interviews for seven items. The agreement correlations for BPRS total scores were higher for the video interviews (.87) than for the in-person interviews (.67).

Using a different design, Jones and colleagues (4) found that BPRS items scored by the patients' verbal report of symptoms obtained better reliabilities than items scored by clinicians' observations. The design included a clinician who administered the BPRS by video while a clinician observing on site scored the items. For a subsample, a second clinician observed by video. The study compared simultaneously scored video administration with 30 in-person observation interviews and with a subset of 18 simultaneously scored live and video observation interviews. ICCs were .83 for the remote interviewer and in-person observer condition and .89 for the remote observer and in-person observer condition.

Grob and colleagues (5) used yet another design, interviewing 27 nursing home residents twice weekly in two conditions. Residents received either an in-person interview and a remote interview or two in-person interviews. ICCs were .81 for the remote group and .49 for the in-person group.

Salzman and colleagues (6) used

Dr. Lexcen is affiliated with the Child Study and Treatment Center, W27-25, 8805 Steilacoom Boulevard, S.W., Lakewood, Washington 98498 (e-mail, lexceff@dshs.wa.gov). The late Dr. Hawk was affiliated with the Institute of Law, Psychiatry, and Public Policy at the University of Virginia in Charlottesville, Virginia. Dr. Herrick is affiliated with the Piedmont Geriatric Hospital in Burkeville, Virginia. Dr. Blank is affiliated with the Center for Mental Health Policy and Services Research and the Leonard Davis Institute of Health Economics at the University of Pennsylvania in Philadelphia.

desktop video equipment to assess six adult inpatients with psychotic disorders. In this study, the researchers used one in-person and one remote interviewer and two in-person interviewers. Interrater reliability was .92 when one rater was remote and .93 when both raters were on site.

No previously published studies reporting interrater reliabilities for video applications of the MacArthur Competence Assessment Tool—Criminal Adjudication (MacCAT-CA) were identified for this review. However, two traditional studies have reported interrater reliabilities for the understanding, reasoning, and appreciation subscales of the MacCAT-CA. Otto and colleagues (7) obtained ICCs of .90 for the understanding subscale, .85 for the reasoning subscale, and .75 for the appreciation subscale. Grisso and colleagues (8) used the MacCAT-CA and reported ICCs for the beginning and end stages of a study of juvenile competence to stand trial. Initial interrater reliability was generally acceptable: understanding subscale, .64; reasoning subscale, .60; appreciation subscale, .41. Later reliability with adolescent participants was improved: understanding subscale, .91; reasoning subscale, .81; appreciation subscale, .86. Interrater reliability with adult participants at the conclusion of the study produced ICCs for the understanding and reasoning subscales—.88 and .71, respectively—that were consistent with those of the original MacCAT-CA study (7). The result for the appreciation subscale was lower, with an ICC of .18, apparently because of a lack of variance in the response scores (8).

We hypothesized that the three conditions of this study would obtain ICCs in the acceptable range for the BPRS-A total score and the MacCAT-CA subscale scores. Scores in the acceptable range would suggest that video evaluations produce ratings comparable with those obtained by in-person interviews.

Methods

The study was conducted between 1997 and 1998 with 72 inpatients from the maximum security forensic unit of Central State Hospital in Petersburg, Virginia. Most participants were de-

tained after being found not guilty by reason of insanity. The remaining participants were detained pretrial. All participants had *DSM-IV* axis I diagnoses of severe mental illnesses. Many also had axis II diagnoses.

Approval for the study was obtained from the institutional review board of the University of Virginia Medical School and from the human subjects committee of Central State Hospital in Petersburg. An individual not associated with the design or data collection obtained consent. Participants received a \$10 payment to their canteen accounts. Completed consent forms were kept separate from data protocols and remained at the hospital to preserve the participants' anonymity.

Two standardized instruments were administered. The BPRS-A was used as a reliable and valid estimate of the presence and severity of psychopathology as measured by a 7-point rating of symptoms on a Likert scale (1). The MacCAT-CA, which has been standardized with large samples of pretrial adult defendants, was used to assess competence to stand trial (9). The MacCAT-CA comprises three subscales: the understanding subscale assesses the defendant's knowledge of courtroom procedures and roles of trial participants, the reasoning subscale assesses the individual's recognition of the relevance of factual information for a defense, and the appreciation subscale assesses whether a defendant's perceptions of his or her own legal situation are unduly influenced by symptoms of thought disorder. Items receive one of three scores: 0, no credit; 1, partial credit; or 2, full credit.

Three administration and observation conditions were used for this study. In each condition, the MacCAT-CA and BPRS-A were administered by one person. Both the interviewer and observer scored the instruments independently. The first condition (local administration, remote observation) entailed in-person administration, with observation via video conferencing. The second condition (remote administration, local observation) entailed administration by video conferencing and observation by the in-person rater. The third condition (local administration, local observation) entailed in-person administration and observation.

Four researchers participated. Three were experienced, university-affiliated clinical psychologists, and one was a senior graduate student in clinical psychology. All were trained and experienced in performing clinical evaluations with the BPRS-A and forensic evaluations with the MacCAT-CA. Pairs of raters were counterbalanced for administration and observation status within and across conditions.

The transmission sites were the maximum security forensic unit at Central State Hospital in Petersburg; the Institute for Law, Psychiatry, and Public Policy in Charlottesville; and the Southeastern Rural Mental Health Research Center at University of Virginia in Charlottesville. Each site had established a two-way interactive video system consisting of the Intel Proshare desktop video conferencing system, Intel Pentium computers, and a transmission rate of 128 kilobits per second.

We calculated interrater reliabilities with ICCs for total scale scores of the BPRS-A and subscales of the MacCAT-CA with the Statistical Package for Social Sciences, version 11.0. Reliabilities greater than .75 indicated excellent reproducibility, greater than .40 and less than .75 indicated good reproducibility, and less than .40 indicated poor reproducibility (10).

Results

ICCs for the BPRS-A total scores in the local administration-remote observation and local administration-local observation conditions were .82 (95% confidence interval [CI]=.63 to .92) and .78 (CI=.55 to .90), respectively (Table 1). The ICC for the remote administration-local observation condition was slightly lower, at .69 (CI=.41 to .85), but still in the range of good reproducibility. All ICCs were significantly different from zero; the confidence intervals did not include zero. The differences between the three conditions were statistically nonsignificant as indicated by the overlap among the three confidence intervals.

ICCs for the subscales of the MacCAT-CA (Table 1) were excellent in all conditions except for the appreciation subscale in the local administration-local observation condition, which was in the range of good repro-

Table 1

Intraclass correlations (ICCs) between scores of two raters on instruments administered to 72 forensic inpatients, by three administration conditions^a

Instrument ^b	Local-remote ^c (N=24)		Remote-local ^d (N=24)		Local-local ^e (N=24)	
	ICC	95% CI	ICC	95% CI	ICC	95% CI
BPRS-A total score	.82	.63–.92	.69	.41–.85	.78	.55–.90
MacCAT-CA subscale score						
Understanding	.86	.71–.94	.81	.62–.92	.95	.89–.98
Reasoning	.82	.63–.92	.89	.76–.95	.90	.78–.96
Appreciation	.79	.58–.90	.87	.73–.94	.69	.40–.85

^a In each condition, one researcher administered both instruments and another researcher observed; both researchers rated the instruments independently.

^b BPRS-A, Brief Psychiatric Rating Scale–Anchored Version; MacCAT-CA, MacArthur Competence Assessment Tool–Criminal Adjudication

^c Instruments were administered in person, and observation was by video conference.

^d Instruments were administered by video conference, and the observer was on site.

^e Instruments were administered in person, and the observer was on site.

ducibility, with an ICC of .69 (95% CI=.40 to .85). The confidence intervals for all three conditions overlapped for all subscales, and none of the confidence intervals included zero.

Discussion

This study confirms previous findings of the use of the BPRS-A for evaluations conducted via video conferencing. Correlations for total scale scores for the BPRS-A were in the good to excellent range. The results for the items of the BPRS-A were consistent with previous studies that found good to excellent reproducibility in experimental conditions using video conferencing (2–6). Our results for the video conditions (.82 and .69) were similar but slightly lower than those from previous studies using remote conditions, which ranged from .81 to .92. Our result for the in-person condition (ICC=.78) was also similar to previous results for in-person conditions, which ranged from .49 to .93.

This study provides the first available data to demonstrate that structured forensic interviews conducted with video conferencing equipment are generally comparable with interviews conducted in person. Subscale scores for the MacCAT-CA were in the good to excellent range. For the understanding subscale, our remote (.86 and .81) and in-person (.95) results were similar to the results of studies conducted by Otto and colleagues (7) and Grisso and colleagues

(8), which ranged from .64 to .91 for nonvideo conditions. Our results for the reasoning subscale—.82 and .89 for remote and .90 for in-person—were similar to the range of reasoning subscale results (.60 to .85) for those studies. Our appreciation subscale results—.79 and .87 for remote and .69 for in-person interviews—were similar or slightly better than the appreciation subscale results (.18 to .86) of the previous studies.

There are several limitations to the present study. Interviews with standardized instruments constitute only one aspect of mental health and forensic assessments and cannot be substituted for complete evaluations. Furthermore, these results may not represent the quality of interviews that can be conducted with newer and less expensive equipment that is now available. In addition, the raters in the study were trained separately on the BPRS-A. Concurrent training as a group might have minimized variability due to rater differences.

Conclusions

Video conferencing offers a method of delivering mental health services to patients who are located far from providers. The expense and risks associated with transporting forensic psychiatric patients securely make remote evaluations especially attractive. Perhaps most promising is the potential to improve the procedural justice of examinations by increasing access

to interviewers with forensic expertise. The results of this study provide support for the reliability of structured interviews administered via video conferencing, suggesting that providers can expect remote interviews to provide clinical information similar to that obtained by interviews conducted in person.

Acknowledgments

This research was sponsored in part by the Southeastern Rural Mental Health Research Center at the University of Virginia, with support from grant P50-MH-49173 from the National Institute of Mental Health. The authors thank James Montgomery, Ph.D., for technical assistance with video equipment and Richard Bonnie, LL.B., John Monahan, Ph.D., and Paul Lombardo, Ph.D., J.D., for consultation on legal and ethical issues.

References

1. Overall J, Gorham D: The Brief Psychiatric Rating Scale. *Psychological Reports* 10: 799–812, 1962
2. Zarate CA, Weinstock L, Cukor P, et al: Applicability of telemedicine for assessing patients with schizophrenia: acceptance and reliability. *Journal of Clinical Psychiatry* 58:22–25, 1997
3. Chae YM, Park HJ, Cho J, et al: The reliability and acceptability of telemedicine for patients with schizophrenia in Korea. *Journal of Telemedicine and Telecare* 6:83–90, 2002
4. Jones BN, Johnston D, Reboussin B, et al: Reliability of telepsychiatry assessments: subjective versus observational ratings. *Journal of Geriatric Psychiatry and Neurology* 14:66–71, 2001
5. Grob P, Weintraub D, Sayles D, et al: Psychiatric assessment of a nursing home population using audiovisual telecommunication. *Journal of Geriatric Psychiatry and Neurology* 14:63–65, 2001
6. Salzman C, Orvin D, Hanson A, et al: Patient evaluation through live video transmission (ltr). *American Journal of Psychiatry* 153:968, 1996
7. Otto RK, Poythress NG, Nicholson RA, et al: Psychometric properties of the MacArthur Competence Assessment Tool–Criminal Adjudication. *Psychological Assessment* 10:435–443, 1998
8. Grisso T, Steinberg L, Woolard J, et al: Juveniles' competence to stand trial: a comparison of adolescents' and adults' capacities as trial defendants. *Law and Human Behavior* 27:333–363, 2003
9. Poythress N, Nicholson R, Otto R, et al: The MacArthur Competence Assessment Tool–Criminal Adjudication: Professional Manual. Odessa, Fla, Psychological Assessment Resources, 1999
10. Landis JR, Koch GG: The measurement of observer agreement for categorical data. *Biometrics* 33:159–174, 1977