

TELEHEALTH OUTREACH TO RURAL REGIONS IN A COGNITIVE DEMENTIA AND MEMORY SERVICE (CDAMS)

Dr Olivia Gobbo, Elizabeth Mullaly, Beth Kirkwood.
CDAMS - Caulfield Hospital, Alfred Health, Melbourne Victoria.

Abstract

We performed an extensive literature search to investigate the use of, and evidence for, VC in the assessment and diagnosis of dementia. Databases included Medline, CINAHL, journals @OVID and Google Scholar. We also utilised relevant articles to obtain further references. We reviewed a number of articles which were published in a variety of different journals. We found that whilst there is a body of published work that supports the validity and reliability of dementia assessment via VC, the sample sizes are frequently small and the methods and tools used in the assessments vary widely. A number of tools were found to have been validated for use via VC. These included tests such as the Mini Mental State Examination (MMSE) and Rowland Universal Dementia Assessment Scales (RUDAS). However, investigation of neuropsychological assessment via telehealth is at a rudimentary stage. Further work is needed to identify the most appropriate method for assessment of patients with cognitive impairment and whether telehealth assessment is equivalent to face to face assessment in a memory clinic which is the current gold standard.

Introduction

Dementia diagnosis has historically been an area which has relied upon access to specialist services such as Geriatricians, Neuropsychologists, Psychiatrists and Neurologists to assess patients and recommend treatment. The advent of the acetylcholinesterase inhibitors into Australia made the requirement for access to specialist services more important as prescription of these drugs was limited to specialist groups. Due to workforce issues, people living in regional and rural areas have reduced access to these specialist services. The state wide Cognitive Dementia and Memory Services (CDAMS) were established in Victoria in 1997 and were designed to increase access to multidisciplinary assessment and treatment for patients affected by different forms of cognitive impairment and dementia.

Telehealth refers to the transmission of images, voice and data between two sites to provide health services. In this review we will contrast the method known as videoconferencing with direct face to face (FF) assessments.

The current literature includes descriptions of various assessment models either used in research studies or in clinical practice. The cognitive assessment tools used in the studies varied widely. Some used little more than an MMSE and others included more comprehensive cognitive assessment together with physical examination.

Some studies focussed entirely upon establishing equivalency of VC compared to FF methods for either cognitive assessment or physical examination at the level of individual tests and measures. Other studies have looked more broadly at the full diagnostic process. In these latter studies there has been a range of approaches taken to the various components that inform clinical diagnosis.

Test / Measure Equivalence Studies

Some studies have compared the same person being given two forms of the same test, one via VC and one FF. Many of these studies contained small sample sizes and the authors have recognised problems with determining what constitutes an acceptable difference in test scores.

In some of the earlier studies, equivalence was measured using correlations. Equivalency of the MMSE was established by Ball et al (1993) and later by Grob et al (2001), who found good correlations for the MMSE for nursing home residents with mild to moderate cognitive impairment. However, Montani et al (1996) found small but significant differences in the videoconferencing scores on the MMSE and the clock face test. It was thought that the technical problems associated with poor sound quality and small numbers contributed to these findings, but non-equivalence for clock drawing has since also been reported by Hildebrand et al (2004) and Cullum et al (2006).

A study performed by Ball and Puffett in 1998 found good correlation between FF scores compared with VC on some subtests of the CAM-Cog component of the CAMDEX assessment tool. The limitations of this study were the small numbers of participants (eight) and problems related to the early technology (eg. slow transmission speeds and narrow band widths). A subsequent study by Ball et al (1999) commented on differences that could arise in outcomes depending upon the mode of scoring used for written protocols such as the sentence and pentagon copy tasks, with reduced reliability introduced via fax and VC modes. However, many studies do not mention what mode was used for scoring responses collected remotely via VC.

Kirkwood et al (2000) administered the National Adult Reading Test, the Quick Test and sections of the Adult Memory and Information Processing Battery to 27 patients diagnosed with alcohol dependency. All subjects were given parallel forms of the tests in VC and FF conditions, counterbalanced in order of presentation. They found it was possible to administer the tests with reliability rates that were similar to the test-retest reliabilities in the published manuals, but found there were some inconsistencies in performance between the two media. They suggested that poorer scores via VC on some tasks were due to slow transmission and poor quality of vision and sound, while scores that were better via VC may have reflected reduced stress. They cautioned that it was not possible to generalize from their findings to other tests and that any other tests would need to be rigorously evaluated on an individual basis.

Saligari et al (2002) examined a group of hospital inpatients twice on the same day, once FF and once via VC. Their scores on the MMSE and Geriatric Depression Scale (GDS) were then compared. This study demonstrated a high correlation with the MMSE and the GDS scores. These authors then conducted a second study in the field,

again assessing equivalence of MMSE and GDS scores and again found satisfactory correlations. Clinicians were present at the remote end when assessments were carried out via VC. McEachern et al (2008) also found equivalence in MMSE scores in a large community sample.

Wong et al (2011) examined the reliability of the RUDAS via videoconferencing. The sensitivity was found to be similar to previous studies and specificity was reduced using both FF assessment and VC. Overall no statistically significant difference in RUDAS score was detected.

Diagnostic Outcome Equivalence Studies

A number of equivalency studies have been performed by different groups attempting to investigate whether assessment via video link will produce the same diagnostic or management advice outcome and is reproducible in both time and place.

Lee et al (2000) found highly correlated scores for the Blessed Dementia Scale (Pearson correlation coefficient 0.980, $p < 0.001$), Short Blessed Test (0.956, $p < 0.001$) and Clinical Dementia Rating Scale (0.928, $p < 0.001$) between the patients assessed remotely and by direct assessment. There was also a one hundred percent agreement on whether dementia was the correct diagnosis. Shores et al (2004) reported similar results in a small sample of patients who screened positive for dementia. They administered the Short Blessed, a three word recall, and clock drawing and obtained 100% concordance in diagnosis, but did not report on the equivalence of the individual tests.

In more recent work correlations and paired t-tests have been replaced by measures such as the Intraclass Correlation Coefficient, which is a measure of reliability agreement, and the Bradley Blackwood Procedure which compares the means and variances of the test-retest measurements. (Bland and Altman, 1986; Bartko 1994). Altman recommended that agreement should be expressed in units directly related to clinical measurements. Using these methods expressing the 'limits of agreement', Loh et al (2004) found clinically significant but not systematic differences, between VC and FF approaches when giving the MMSE and GDS to hospitalised patients. They speculated that this may be due to differences in motivation, willingness to disclose sensitive information and inclusion of patients with delirium in their sample. In a later study using similar tools in a community sample, a more satisfactory level of equivalence was found, although in that study carers reported greater impairment when interviewed in person. Taking FF assessment as the gold standard, a k coefficient of concordance between the two methods was 0.80. (Loh et al, 2007).

In the studies mentioned above, Loh et al (2004, 2007) described using VC for the full assessment process. These authors used VC to administer a test battery including a MMSE, GDS, Katz assessment of Activities of Daily Living, Instrumental ADL assessment and the Informant Questionnaire for Cognitive Decline in the Elderly. There was little assistance required at the remote end during cognitive assessment.

An alternative model that has been described (eg Martin-Khan et al, 2007) is one where a standardized cognitive assessment is carried out FF by a trained health professional at the client/remote end. The results of that assessment are then supplied

to a specialist physician at the hub who continues the assessment, including a clinical interview and a limited form of physical examination, via VC. In a pilot study using weighted kappa scores there was moderate to good agreement using the two approaches (FF and VC) both for diagnosis and management recommendations concerning driving and testamentary capacity. The advantage of this format is reduced time spent by the specialist physician. In 2012 these researchers extended their findings using a larger group of patients with more complex diagnostic issues and random allocation to condition. In this well-powered study they demonstrated satisfactory diagnostic agreement between VC and FF methods. However the protocol again relied on administration of all cognitive measures FF at the remote end, and there was no VC administration of tasks.

A similar approach (whereby cognitive testing was done FF by a skilled clinician at the remote site and followed by a VC diagnostic interview) is also described in a paper by Weiner et al (2011). This paper describes a VC based dementia diagnosis and management service focussing on clinical as opposed to research outcomes, given the reluctance of their client group to consent to research procedures.

Role of physical examination in cognitive assessment process

The role of physical examination in VC assessment of patients with cognitive impairment remains controversial. Some studies found that a limited physical examination was able to be performed with the assistance of a trained clinician at the remote end. This included assessment of basic functions such as gait, co-ordination and reflexes. In a study by Craig et al (1999) it was argued that neurological examination performed by Neurologists via video link was equivalent to a FF assessment using a junior doctor. Loh et al (2007) used a model which relied on physical examination performed by the General Practitioner prior to the patient attending for cognitive assessment. Shores et al (2004) found that essential elements of the physical examination were able to be performed via video link. These included gait, eye movements and observation of hand tremor. Primitive reflexes were tested with the assistance of the remote clinician. Martin-Khan et al (2008) commented that physical examination was not necessary to make a diagnosis of dementia but was required to define the dementia type.

Neuropsychological Assessment

Our literature search, uncovered four relevant studies that dealt with VC neuropsychological assessment, as opposed to battery approach cognitive assessment.

Jacobsen et al (2003) found a total of 25 neuropsychological tests that were able to be given remotely between two psychologists without a testing assistant present at the remote end. They chose twelve measures to sample a range of domains and administered this battery to a sample (n=32) of normal volunteers. In this study the VC approach enhanced performance on auditory memory (Logical Memory I) and auditory attentional (Seashore Rhythm Test) measures, but the remainder of the measures obtained equivalent results regardless of method of delivery. These measures included the Grooved pegboard test of visuo-motor speed, the Seashore Rhythm Test of auditory attention, Logical Memory II, the Benton Visual Retention test of nonverbal memory, the Silhouette subtest of visual perception, the vocabulary

subtest (WAIS), Digit Span (WAIS) and the Symbol Digit Motor test of information processing. Visuospatial, visuoconstructional and executive measures were not included. In a similar study, Hildebrand et al (2004) showed reasonable equivalence for the RAVLT, the WAIS-III Matrix Reasoning, the WAIS-III Vocabulary, the Brief Test of Attention and letter fluency, but not for clock drawing. They cautioned, however, that they did not rule out small differences which could potentially make a difference diagnostically for patients with subtle impairment.

For clinical populations a study by Cullum et al (2006) established the equivalence of various neuropsychological tests when administered to 14 patients with mild cognitive impairment and 19 patients with mild to moderate Alzheimer's disease. Diagnosis had been made prior to entry into the study using measures other than those that were the target of the investigation and the authors noted that this meant the subjects were familiar with neuropsychological assessment in all cases. The tests that were included in the study (the MMSE, the Hopkins Verbal Learning Test-R, Clock Drawing, Digit Span, Category Fluency, letter fluency, and the 15 item version of the Boston Naming Test) were chosen on the basis of having parallel forms and of being able to be administered successfully via VC. The findings were generally positive for the tests that were included, with equivalence established for many, although there was a failure to establish equivalence on the delayed recall of the Hopkins Verbal Learning Test-R, a measure which is of primary importance to neuropsychologists in detection of clinically relevant memory disturbance. Patients had higher mean retention scores on this test in the FF situation and the authors speculated that the FF condition provided a better incentive for effort on this test. Clock drawing results also failed to reach equivalence.

While this study was important in advancing our understanding of the effects of VC on neuropsychological assessment, it does not provide any information about the possibility of neuropsychological methodology being applied using VC approaches. As the authors pointed out, a number of commonly used neuropsychological tests were unable to be administered by VC and some of these are frequently included in neuropsychological evaluations. Thus while the study showed equivalence at the test level there was no attempt to establish equivalence at the level of drawing inferences or making diagnostic conclusions.

Barton et al (2011) described VC neuropsychological assessment in a clinical setting, with reliance on a skilled technician at the remote end to assist with administration of certain tests requiring more "hands on" assistance (e.g. the Trail-Making Test which is not easily administered via VC). Using this approach, the authors claimed that they were able to make diagnostic decisions (distinguishing dementia, mild cognitive impairment and normal for age). However, the study did not compare the findings derived from the VC methodology with the gold standard of FF assessment.

In summary, there is no evidence in the literature to reassure neuropsychologists that assessments conducted via VC would come up with the same conclusions as assessments of the same patient conducted FF. Since many commonly used neuropsychological tests cannot be administered via VC, a VC approach would involve restriction of the available tests that could be utilized. The most obvious repercussions would be the restriction of assessment of nonverbal, visuo-perceptual, visuoconstructional, and executive functions. Adequate assessment of these domains

is essential to providing the fine-grained, sophisticated type of assessment for which neuropsychological testing is valued (Clark et al, 2013). In addition, the studies of both Jacobsen et al (2003) and Cullum et al (2006) demonstrated inconsistent results on certain memory tests (Logical Memory I and HVLT-R delayed recall) that are vital in neuropsychological diagnosis of dementia. Use of these tasks in VC assessment will remain unsubstantiated until relevant normative data sets can be developed.

Acceptability to patients and clinicians

A number of studies investigated patient acceptability of assessment via videoconferencing. Kirkwood et al (2000) found that participants rated the use of VC with high satisfaction scores (mean 8.2). A ten point scale was used to express satisfaction with the video consultation. Twenty-two of the twenty-seven respondents said they would use the link again, three were unsure and one would not like to use it again. Participants reported that they liked various aspects of the video consultation. Five reported liking everything, eight reported feeling relaxed and at ease, a further eight reported no effort needed to use the equipment and three liked having access to a health professional. Criticisms included issues with sound quality (6) and delays in communication (3).

Shores et al (2004) also found that a group of veterans subjected to cognitive assessment via VC after a positive screening test, rated the assessment with a high degree of satisfaction. The respondents completed a questionnaire on the interaction which consisted of six items. They were asked to rate the level of agreement on a five-point scale with 5 corresponding to highest level of agreement and 1 to the lowest level. The questionnaires examined usefulness of the interaction, efficacy compared to in person or telephone and willingness to try VC again. All sixteen subjects agreed that they would prefer VC to travelling for an appointment. Most (86.7%) strongly agreed with this statement. Most (93.7%) felt they understood the clinician as well as if the examination had been in person. The clinicians reported that the technology worked well 71.4% of the time and the audio and visual quality was adequate 78.6% of the time.

A study of 140 residential care patients in Korea over two years found good acceptance of a VC assessment system by both staff and patients. In addition to acceptability, this study also demonstrated that 46% of patients showed clinical improvements through the telehealth intervention and a decrease in carer burden as measured by the Zarit burden interview (Lee et al, 2000). Hildebrand et al (2004) found that all participants in their study felt comfortable with the communication and found the sound quality to be acceptable. Only 10% felt their performance would have been better in person

Saligari et al (2002) found a high rate of clinician and patient satisfaction utilising a Likert scale questionnaire based on the work of Field (1996).

Jacobsen et al (2003) commented on the fact that VC is relatively demanding from the clinician's point of view. There is increased reliance on verbal communication and continuous eye contact is required to ensure patient engagement. The clinician needs to look at the camera and this makes reading instructions difficult. The test instructions need to be clear and the test material within easy reach.

In Cullum's (2006) study, which involved administration of a wide range of tasks, no participant expressed concerns about VC administration but several suggested they would prefer FF contact.

Mair et al (2000) performed a systematic review of patient satisfaction with VC. They reviewed 32 studies which used various different study methods including simple survey instruments (26 studies), exact methods not specified (5) and qualitative methods (1). All studies reported good levels of patient satisfaction. Unfortunately methodological deficiencies of the published data limited generalisability of the findings. The authors commented that communication issues, the quality of interpersonal relationships with this medium and subsequent effects, if any, on the outcome of consultations have yet to be fully explored.

Factors impacting on successful implementation of a telehealth program

Impediments to introducing a telehealth program were discussed in a number of papers. There is a high prevalence of visual and hearing impairment in the memory clinic population which can restrict suitability for VC (Ball et al, 1998). More severe forms of cognitive disorder can also make the process unreliable (Loh et al, 2004). Loh et al (2004) also described medico-legal issues, risks related to malpractice and medical defence, duty of care issues, cost concerns, lack of reimbursement, lack of clinical standards, scheduling difficulties and time limitations all as factors potentially impacting on take-up. The factors needed for success included academic and clinical champions, a dedicated co-ordinator, development of protocols and reliance on previously established referral practices. (Saligari et al, 2002).

Summary

Studies relevant to VC diagnosis of dementia have been carried out since the early 1990s. While equivalence of tools and measures was difficult to establish in the early years, due to technological limitations, a body of evidence has emerged over recent times to support the use of screening tools such as the MMSE, the RUDAS, and the Geriatric Depression Scale (GDS) by VC with clinical populations based in remote locations. Studies have also demonstrated the feasibility of administering more extensive cognitive batteries and have established equivalence of many neuropsychological assessment tools.

Caution has been expressed, however, about the need to evaluate equivalence for each specific test individually and there have been some findings of non-equivalence that would suggest separate normative data will need to be developed for VC administration. In addition, it has been noted that options for assessing some cognitive domains remain restricted, thus providing a barrier to using neuropsychology via VC as a means of investigating complex and atypical patients.

In recent large-scale studies of VC diagnostic outcomes, cognitive assessments were carried out by trained staff at the remote location with the VC component consisting of the medical specialist's clinical interview and a limited physical examination. Using this methodology VC has been found to be equivalent to FF for the diagnosis of dementia and promises to provide a suitable option for remotely located patient groups. Patient and staff satisfaction has proved positive on most measures.

Practitioners need to pay attention to medico-legal issues, insurance, costing concerns, scheduling difficulties, and time limitations when setting up arrangements.

References

- Ball C, Puffett A. The assessment of cognitive function in the elderly using videoconferencing. *J VC and Telecare*. **1998**; 4(suppl 1): 36-38
- Ball C, Scott N, McLaren PM, & Watson JP. A preliminary evaluation of a low cost videoconferencing system for remote cognitive testing of adult psychiatric patients. *British Journal of Clinical Psychology*, **1993**;32:203-7
- Ball C, Tyrrell J, & Long C. Scoring written material from the MMSE: a comparison of face-to-face, fax and video-linked scoring. *J VC and Telecare*, **1995**; 5: 253-256
- Barton C, Morris R, Rothlind J, & Yaffe K. Video-VC in a memory disorders clinic: Evaluation and management of rural elders with cognitive impairment. *VC and e-Health*, **2011**;17(10):789-793
- Clark L, Delano Wood, L, Libon, D, McDonald, C, Nation, D, Bangen, K, Jak, A, Au, R, Salmon, D, & Bondi, M. Are empirically-derived subtypes of mild cognitive impairment consistent with conventional subtypes? *Journal of the International Neuropsychological Society*, **2013**; 19, 1-11
- Craig J, McConville J, Patterson V, & Wootton R. Neurological examination is possible using VC. *J VC and Telecare*, **1999**; 5:177-181
- Cullum C, Weiner M, Gehrman H, & Hynan L. Feasibility of telecognitive assessment in dementia. *Assessment*, **2006**; 13: 385-390
- Grob P, Weintraub D, Sayles D, Raskin A & Ruskin P. Psychiatric assessment in a nursing home population using audiovisual telecommunication. *J.Geriatr Psych Neurol*, **2001**; 14:63-65
- Hildebrand R, Chow H, Williams C, Nelson M, & Wass P. Feasibility of neuropsychological testing of older adults via videoconference: implications for assessing the capacity for independent living. *J VC and Telecare*, **2004**;10:130-134
- Jacobsen S, Sprenger T, Andersson S, & Krogstad J. Neuropsychological assessment and VC: A preliminary study examining the reliability of neuropsychology services performed via telecommunication, *J International Neuropsychological Society*, **2003**; 9:472-478
- Kirkwood K, Peck D, Bennie L. The consistency of neuropsychological assessments performed via telecommunication and face to face. *J VC and Telecare*, **2000**; 6:147-151
- Lee J, Kim J, Jhoo J, Lee K, Kim K, Lee D, Woo JA. VC system as a care modality for dementia patients in Korea. *Alzheimer Disease and Associated Disorders* **2000**; 14: 94-101
- Loh P, Donaldson M, Flicker L, Maher s, & Goldswain P. Development of a VC protocol for the diagnosis of Alzheimer's disease. *J VC and Telecare*, **2007**;13:90-94

Loh PK, Ramesh P, Maher S, Saligari J, Flicker L, & Goldswain P. Can patients with dementia be assessed at a distance? The use of Telehealth and standardised assessments. *Int Medicine Journal*, **2004**;34:239-242

Mair F, Whitten P. Systematic review of studies of patient satisfaction with VC. *BMJ*, **2000**; 320(7248):1517-20

Martin-Khan M, Flicker L, Wootton R, Loh, PK, Edwards, H, Varghese, P, Byrne, G, Klein, K, & Grey, L. The diagnostic accuracy of telegeriatrics for the diagnosis of dementia via videoconferencing. *JAMDA*, **2012**;13: 487.e19-487.e24

Martin-Khan M, Varghese P, Wootton R, & Gray L. Successes and failures in assessing cognitive function in older adults using video consultation. *J VC and Telecare*, **2007**;13(suppl 3):60-62

Martin-Khan M, Varghese P, Wootton R, & Gray L. Dementia examination and diagnosis of dementia for video consultation. *JAGS*, **2008**; 56(5): 947-949

McEachern W, Kirk A, Morgan DG, Crossley M, & Henry C. Reliability of the MMSE administered in-person and by telehealth. *Canadian Journal of Neurol Sciences*, **2008**; 35(5): 643-6

Montani C, Billaud N, & Couturier P, et al. "Telepsychometry": a remote psychometry consultation in clinical gerontology: Preliminary study. *VC Journal*, **1996**;2:145-150

Ramos-Rios R, Mateos R, Lojo D, Conn D, & Patterson T. Telepsychogeriatrics: a new horizon in the care of mental health problems in the elderly. *Int Psychogeriatrics*, **2012**;10:1-17

Saligari J, Flicker L, Loh PK, Maher, Ramesh P, & Goldswain P. The clinical achievements of a geriatric telehealth project in its first year. *J VC and Telecare*, **2002**; 8(suppl. 3):S3:53-55

Shores M, Ryan-dykes P, Williams R et al. Identifying undiagnosed dementia in residential care veterans: comparing VC to in-person clinical examination. *Int J Geriatric Psychiatry*, **2004**;19:101-108

Wong L, Martin-Khan M, Rowland J, Varghese P, & Gray L. Reliability of the RUDAS via videoconferencing. *Int J Geriatric Psychiatry*, **2011**; 25:988-989