Critical Review:
Are observer-rated dementia screening tools effective for identifying dementia of the Alzheimer's type in individuals with Down syndrome?

Alison Csercsics
M.Cl.Sc SLP Candidate
University of Western Ontario: School of Communication Sciences and Disorders

In this paper, the evidence supporting four observer-rated screening tools used to identify the presence of dementia in individuals with Down syndrome (DS) is critically evaluated. Six papers, two prospective longitudinal designs and four cross-sectional designs, were analyzed. The results revealed that three out of the four tools had appropriate psychometric properties; however, these findings are restricted by consequential methodological and statistical flaws. Directions for future research are indicated to support recommendations for clinical appropriateness and utility.

Introduction

It is largely acknowledged that dementia, specifically of the Alzheimer's type, is more prevalent in adults with Down syndrome (DS) than in the general population (Cipriani et al., 2018; Elliot-King et al., 2016; O'Caoimh et al., 2013). Due to the absence of universally accepted diagnostic criteria and various demographic characteristics that influence cognition, prevalence rates are inconsistently reported across the literature. Despite the incongruency, it is speculated that as many as half of the DS population will show signs of dementia by 60 years of age and as many as 75% of the population by 65 years of age (O'Caoimh et al., 2013).

In addition to the alarming prevalence and earlier onset of cognitive decline in this population, the need for robust screening tools is further emphasized by the crucial demand to accurately detect a differential diagnosis of clinical symptoms to ensure appropriate and time-sensitive intervention is provided (O'Caoimh et al., 2013). Cognitive decline in this population can be related to the neurodegenerative properties of the condition as well as changes due to ageing (Orange & Zanon, 2005). Furthermore, various diseases and disorders such as thyroid disease and depression can present similarly to cognitive decline in individuals with DS (Cipriani et al., 2018; O'Caoimh et al., 2013).

Despite the indisputable importance of robust screening tools for identifying dementia in adults with DS, identification of cognitive decline is complicated by many factors that challenge the accuracy and utility of the screening tools developed for the general population. Floor effects with normative comparisons are observed due to below-average intellectual ability at baseline as well as the diversity of baseline cognitive function across the population (Cipriani et al., 2018). Furthermore, emerging evidence is suggestive that the disease may present differently both clinically and neurologically in individuals with DS compared to the general population (Elliot-King et al., 2016; O'Caoimh et al., 2013).

To address the limitations of the existing dementia screening tools, several scales have been adapted for the sole purpose of identifying dementia in individuals with DS (O'Caoimh et al., 2013). These tools have addressed the inherent factors that influence screening ability in this population through using observer-rated design, which has demonstrated better testing results when compared to direct testing in this population and determining identification on deviation from baseline rather than absolute or normative cutoff scores (O'Caoimh et al., 2013). Despite these rational adaptations, which have led to the acceptance into clinical practice, the research supporting the validity of the existing tests is limited.

Objectives

The primary objective of this review is to critically appraise the existing literature addressing four common observer-rated dementia screening tools that are used to identify the presence of dementia in individuals with DS. This information is intended to inform clinical practice as well as emphasize the limitations in the research.

The four screening tools examined in this paper: Dementia Questionnaire for Persons with Mental Retardation (DMR), Dementia Scale for Down Syndrome (DSDS), Dementia Screening Questionnaire for Individuals with Intellectual Disabilities (DSQIID), and National Task Group – Early Detection Screen for Dementia (NTG-EDSD).

Methods

Search Strategy

Literature addressing the topic of interest was found using the online databases: PubMed, CINAHL, and

Google Scholar. The search terms utilized were: ((observer rated screening) AND (dementia)) AND (Down syndrome). The search was refined to only include papers written in English.

Selection Criteria

Papers that were included were required to evaluate the psychometric properties of one or more observer-rated tools indicated above as well as report these findings specific to a sample of adults with DS. Papers that grouped adults with DS with adults with other intellectual disabilities (ID) were excluded.

Data Collection

The search revealed six articles relevant to the topic of interest. Two papers utilized a prospective longitudinal design to evaluate the DMR tool. The remaining four papers employed a non-experimental cross-sectional design with one paper comparing the DMR and DSDS, two papers assessing the DSQIID tool, and one paper evaluating the NTG-EDSD tool.

Results

Dementia Questionnaire for Persons with Mental Retardation (DMR)

The DMR, currently known as the Dementia Questionnaire for People with Learning Disabilities (DLD) (Silverman et al., 2020), is an observer-rated screening tool that was developed in the 1980s to identify the presence of dementia in individuals with ID with and without concomitant DS (Evenhuis, 1996; Prasher, 1997). The tool is comprised of eight categories that contribute to two scores: the sum of cognitive scores and the sum of social scores. The scale can be administered and interpreted at a single point in time (absolute score) or over multiple points in time (longitudinal score changes) (Prasher, 1997).

Evenhuis (1996) utilized a prospective longitudinal design to investigate the proposed diagnostic criteria of the DMR by reporting the specificity and sensitivity of the absolute and longitudinal scores. The study included two groups: 33 participants over the age of 70 without DS and 45 participants over the age of 35 with DS. The results revealed that, with modification of cut-off criteria, using the DMR over time provided the most appropriate diagnostic accuracy across both groups, resulting in a sensitivity value of 100% and a specificity value of 75% in the DS group.

Participants were recruited from an assisted living facility; however, no selection criteria were reported. Important group differences existed between the experimental and control group: the average age of the senior group was significantly greater than those in the DS group and the majority of participants in the seniors' group had mild to high-moderate ID while the majority of participants in the DS group had low-moderate to severe ID. These group differences have implications on the validity of the statistical comparisons. Furthermore, the small sample size contributes to low statistical power.

The longitudinal design of the study permitted the investigation of the validity of the repeated use of the DMR screening tool in comparison to the use at a single point in time; however, a limitation of this design resulted in high participant attrition, primarily due to death. The reference standard utilized in this study, DSM-III-R, was a reasonable comparison for the time of research, yet the author did not report blinding protocols used during assessments, resulting in potential biases in the interpretation of the diagnoses.

The statistical procedures employed in this study were limited as only the specificity and sensitivity values for each criterion were reported. The use of more advanced statistical tests such as a receiver operating characteristic (ROC) analysis and precision values would increase the validity and reliability of the findings.

Succinctly, due to limited statistical rigor and sample biases, the study offers equivocal evidence of the validity of the DMR screening tool in identifying dementia in the DS population. Despite the low-quality evidence, the insufficient diagnostic accuracy and inconsistent cut-off criteria present further limitations for clinical appropriateness and utility.

Prasher (1997), in using a prospective longitudinal design, replicated Evenhuis' study (1996) to provide an objective evaluation of the DMR scale and address the previous limitations. One hundred individuals with DS were evaluated twice annually with both the DMR and a multifactorial clinical assessment. The results indicated that modifications to the existing diagnostic criteria are necessary to maximize precision and such criteria should be dependent on comorbid medical diagnoses. Based on the proposed criteria, the DMR had overall sensitivity and specificity values of 82%.

The sample was divided into four groups based on medical diagnosis: no dementia or disorder, dementia, depression, and hypothyroidism. This assignment accounted for implications of differential diagnoses on clinical presentation and controlled for confounding variables impacting the screening tool accuracy. However, the groups were inadequately represented with the majority of the participants (n=81) comprising the no disorder group. Further limitations of the sample were evident as insufficient representation across levels of ID

was noted. The small participant size in the remaining groups, particularly the dementia group, reduces the statistical power, ultimately decreasing the validity of the results.

The statistical analyses employed were limited to sensitivity and specificity calculations, lessening the statistical rigor of the findings. However, these measures of diagnostic accuracy were provided for each ID severity (mild, moderate, and severe) as well as for both the existing and the proposed criteria, which accounts for differences that may exist across groups.

Although this study addressed the limitations of the prior literature regarding the validity of the DMR, the evidence provided by this paper is equivocal due to limitations in statistical rigor and sample representation. Clinical applicability is further questioned as a lack of consistent criteria as well as poor specificity limits the effectiveness as a screening tool for dementia in the DS population.

Dementia Scale for Down Syndrome (DSDS)

Developed by Gedye in 1995, the DSDS is a screening tool used to identify dementia specifically in individuals with DS through interviews with caregivers (Deb & Braganza, 1999). The tool has a particular focus on behavioural changes from baseline in daily activities. The tool is comprised of 60 questions that correspond to three stages of dementia: early, middle, and late-stage (Deb & Braganza, 1999). The scale is intended to be administered solely by trained psychologists (Takenoshita et al., 2020).

Deb and Braganza (1999) utilized a non-experimental cross-sectional design to explore the diagnostic accuracy, specificity and sensitivity, of the DSDS and the DMR in comparison to a clinical diagnosis. Sixty-two individuals with DS were recruited, twenty-six with a confirmed diagnosis of dementia and thirty-six with no known history of cognitive decline. The authors found both observer-rated scales to be accurate tools in identifying dementia in the DS population, suggesting the use in clinical practice alongside neuropsychological tools.

Participants were recruited from community-based databases; however, selection criteria were insufficiently described. Further, details regarding participant demographics were limited and no distinctions between experimental and control groups were identified, leading to uncertainty about homogeneity between groups as well as concerns regarding generalizability. Biased participant selection has implications on the study's validity as well as replicability.

All participants received a comprehensive psychiatric evaluation based on the ICD-10 framework for diagnosing dementia in DS to act as a reference point for the observer-rated tools. Although no gold standard exists, the ICD-10 is a suitable guide for making a diagnosis in this population. Despite valid comparisons, the study's validity is guarded on the lack of blinding of clinical diagnosis results during the administration of the screening tools. This can contribute to the over-interpretation of the findings.

The statistical methods utilized in this study were limited as no precision values or ROC calculations were employed. Despite the limitations, the DMR yielded a specificity and sensitivity of 0.92, indicative of good diagnostic accuracy, and the DSDS yielded a sensitivity of 0.85 and specificity of 0.89, indicative of fair diagnostic accuracy. Further, the authors reported a positive correlation between both the DSDS and the DMR, suggesting that both tools identify dementia in the same individuals.

This study provides somewhat suggestive evidence that the DMR and the DSDS have adequate psychometric properties due to the valid study design and appropriate sample size. The strength of evidence is guarded due to the unrepresentative sample and the absence of blinding during the assessment process.

Dementia Screening Questionnaire for Individuals with Intellectual Disabilities (DSQIID)

The observer-rated tool, the DSQIID, developed by Deb and colleagues in 2007, assesses cognitive decline in individuals with ID (Deb et al., 2007). The tool is comprised of three categories that aim to document function prior to decline and record emerging dementia-related behaviours (Deb et al., 2007). This tool, originally developed in English, has been adapted and validated for other languages including Japanese (Takenoshita et al., 2020).

Deb and colleagues (2007) reported the psychometric properties of the DSQIID using a non-experimental cross-sectional design. Data were obtained from 193 caregivers of people with DS to calculate the feasibility, content validity, construct validity, internal consistency, and reliability. Criterion validity was calculated based on the comparison of the DSQIID to a clinical diagnosis comprising of the ICD-10 framework for 117 of the participants. Findings of the study suggested that the DSQIID has robust psychometric properties, indicating the observer-rated scale is an appropriate tool for screening dementia in individuals with DS.

Participant selection was achieved through community-based contacts; however, the specific inclusion and

exclusion criteria for the sample were not sufficiently described, resulting in concerns regarding biases in sample characteristics and generalizability of results. The authors did not measure nor report the intellectual abilities of the sample, although they speculated that based on diversity in language abilities and living situations among participants, the sample was representative of a range of intellectual abilities.

Of the 193 caregivers recruited, only 117 respective individuals with DS were assessed to confirm the absence or presence of dementia, resulting in 49 individuals with and 68 without a clinical diagnosis of dementia. The two groups were not matched although to address this limitation, an analysis of intergroup differences was completed. Individuals with dementia were significantly older and had significantly higher prevalence of hearing and visual problems; however, all other health-related variables were not significantly different between the groups.

The statistical measures utilized in this study were appropriate given the methodological design. A cut-off for positive diagnosis was established at a score of 20, yielding a sensitivity of 0.92 and a specificity of 0.97, which is suggestive of good diagnostic accuracy. Furthermore, the positive likelihood ratio was 31 and the negative likelihood ratio was 0.08. The statistical analysis also revealed adequate reliability including an intraclass correlation of 0.95 (n=52, $P<0.01$) for test-retest reliability and intraclass correlation of 0.9 (n=42, $P<0.01$) for interrater reliability.

This study offers highly suggestive evidence that the DSQIID tool has robust psychometric properties in screening dementia for individuals with DS. Despite robust methodological rigor and compelling findings, bias in participant selection limit clinical applicability and generalizability to clinical populations.

Takenoshita and Colleagues (2020) examined the validity of the Japanese version of the DSQIID to determine the clinical application of the scale in identifying dementia in individuals with ID. The sample included 493 caregivers of individuals with ID with or without a diagnosis of DS. The results indicated that the screening tool had credible reliability (interrater, test-retest, and internal consistency) and validity across all participants, suggesting the scale is an appropriate tool for screening dementia in individuals with ID and specifically, those with DS.

Participant selection criteria and participant demographic information were appropriately defined; however, the authors acknowledged the limitations to the sample representation as the participants were recruited solely from community facilities. This sample bias has implications on the representation of intellectual abilities of the participants in comparison to the population and thus limits the generalizability of the results. Furthermore, despite the large sample size, participants with DS were largely underrepresented with only 34 of the participants (n=27 without dementia, n=7 with dementia) having a diagnosis of DS. Although the data from the DS group were computed separately, the small sample size reduces the power, limiting the validity of the findings.

The methodological procedures utilized in this cross-sectional design study strengthened the validity of the results. The diagnostic accuracy of the DSQIID version was compared to a comprehensive dementia evaluation based on the DSM-5 criteria which served as a valid reference measure. The three physicians who completed the assessment were blinded to the DSQIID results, which reduced the likelihood of biased interpretations.

The findings of the study are supported by robust statistical analyses. An ROC computation was used to report optimum cut-off criteria and associated specificity and sensitivity. Specifically for the DS group, good diagnostic accuracy of both values was reported (specificity of 96.3% and sensitivity of 100%).

Due to the appropriate study design and robust statistical analyses, this study provides highly suggestive validity of the DSQIID-J that translates into clinical appropriateness for screening dementia in individuals with ID. However, due to the limitations in the DS sample, interpretation to this population is cautioned.

National Task Group – Early Detection Screen for Dementia (NTG-EDSD)

Developed from the DSQIID, the NTG-EDSD is a screening tool for identifying the presence of dementia, specifically at the prodromal cognitive decline stage. The tool is comprised of questions that target 6 domains including activities of daily living, language, circadian rhythms, mobility, behaviour, and memory (Silverman et al., 2020).

Silverman and colleagues (2020) investigated the validity of the NTG-EDSD in identifying dementia during the prodromal stage in individuals with DS. In using a cross-sectional design, the NTG-EDSD results were compared to a comprehensive diagnostic assessment that denoted three main groups of participants: cognitively stable, mild cognitive impairment, and dementia. The results of the study indicated that the NTG-EDSD scale is an insufficient stand-alone screening tool due to inadequate diagnostic

accuracy of both mild cognitive impairment and dementia in individuals with DS.

Participant selection criteria were sufficiently defined, and relevant demographic information of the sample was provided for each group. Intellectual impairment severity was reported for each group; however, participants with severe ID were underrepresented. Further, the two groups were not matched, and intergroup differences were not examined. This has potential implications on appropriate sample representation as well as the influence of confounding variables.

The methodology employed in this study was scientifically sound. The NTG-EDSD was evaluated in comparison to an appropriate reference standard, the AAMR-IASSID, which supports the validity of the findings. Methods were limited due to the lack of blinding, although 92% of informants were not aware of cognitive changes in those diagnosed with mild cognitive impairments. Additionally, the authors did not investigate other psychometric properties of the screening tool.

The statistical analyses used in this study were valid and appropriate. Differences between the three participant groups in terms of single test items, domain sums, and total concerns were investigated using multi-level analyses. In using Kruskal-Wallis U and Mann-Whitney U tests, the three groups demonstrated statistically significant differences in the total number of concerns reported, with the dementia group having the highest number of concerns and the cognitive stable group with the lowest number of concerns reported. However, based on Deb and colleagues' (2007) proposed cutoff (20 concerns), the NTG-EDSD had very poor sensitivity for the dementia group (0.421) and the mild cognitive impairment group (0.056). Further, an ROC analysis was conducted and revealed a sensitivity value of 0.833 and specificity of 0.640 (cut off ≥ 2) when cognitive stable and mild cognitive impairment groups were compared and a sensitivity value of 0.868 and a specificity value of 0.802 (cut off ≥ 5) when cognitive stable and dementia groups were compared. These findings are indicative of inadequate diagnostic accuracy, specifically in identifying prodromal cognitive decline in the DS population.

Overall, this study employed robust methodological design and statistical procedures indicating highly suggestive validity of the findings. The results, however, have equivocal clinical importance as the NTG-EDSD has insufficient specificity and sensitivity to be used as an independent screening tool for identifying dementia in individuals with DS.

Discussion

The purpose of this paper was to critically evaluate the validity and reliability of four common observer-rated screening tools used to identify dementia in individuals with DS. Three of the four scales, the DMR, DSDS, and DSQIID, demonstrated to be effective tools in screening dementia in adults with DS; however, crucial limitations across the studies impacted the validity of the evidence, resulting in appraisals ranging from equivocal to highly suggestive. Due to these inherent limitations in the research, these screening tools do not have sufficient evidence supporting robust psychometric properties, which is essential to endorse clinical implementation.

Small and unrepresentative sample sizes were a main deficit across the studies, limiting the statistical power as well as the generalizability of results. Due to the heterogeneity of several domains across the DS population, employing a sample that is illustrative of the varying intellect abilities, subtypes of DS, common comorbid or differential diagnoses, and various living situations is imperative to ensure screening of dementia is accurate and not confounded by other variables. Future research should aim to address participant sampling limitations to ensure results provide effective screening as well as are applicable to adults with DS that are seen in clinical practice.

The evidence supporting the screening tools is further complicated by the inconsistent diagnostic cut-off criteria used to report the diagnostic accuracy. The lack of robust cut-off scores associated with the tools limits the applicability and utility in clinical settings. In order to be appropriate screening tools, further in-depth investigation into robust diagnostic criteria of the scales is imperative to ensure screening is truly sensitive and specific in identifying dementia in adults with DS.

Investigation into the effectiveness and clinical utility of dementia screening tools and assessments are also hindered by the absence of a standard reference point for diagnosis. The lack of a standard comparison raises concerns when comparing the reported psychometric properties of the tests across different studies. Although the papers included in this review utilized reputable guidelines for clinical diagnosis (i.e., physician decision based on ICD-10, DSM-III-R, DSM-V, or AAMR-IASSID criteria), differences can have impacts on the validity of psychometric computations as well as comparisons across studies. Future research establishing a widely accepted reference standard is necessary not only to support screening tool development and use but also to ensure and enhance the diagnostic accuracy of dementia in this population.

The literature review resulted in a limited number of published papers addressing the validity of the four dementia screening tools: DMR, DSDS, DSQIID, and NTG-EDSD. Fewer papers investigated the psychometric properties of the tests with the DS population separate from participants with other ID. Investigating the validity and reliability of the scales within the DS population independently ensures results are valid and applicable for adults with DS as differences in cognitive decline are recognized (O'Caoimh et al., 2013). Future research pertaining to the development and investigation of these screening tools should obtain appropriate and representative samples to ensure evaluation is reflective of the DS population.

Clinical Implications

Due to the limited evidence that is further hindered by the insufficient methodological and statistical rigor, all four screening tools, DMR, DSDS, DSQIID, and NTG-EDSD, should be evaluated and utilized with caution to prevent under- or over-identification of dementia in adults with DS. With the currently available evidence, despite the evident limitations, the DSQIID screening scale has demonstrated adequate reliability and validity in identifying dementia in adults with DS; however, use is recommended alongside other tools. Incorporation into clinical use is cautioned pending further research to confirm diagnostic accuracy and appropriateness for clinical implementation.

References

Cipriani, G., Danti, S., Carlesi, C., & Di Fiorino, M. (2018). Aging with Down syndrome: The dual diagnosis: Alzheimer's disease and Down syndrome. *American Journal of Alzheimer's Disease & Other Dementias, 33*(4), 253-262. https://doi:10.1177/1533317518761093

Deb, B. & Braganza, J. (1999). Comparison of rating scales for the diagnosis of dementia in adults with Down's syndrome. *Journal of Intellectual Disability Research, 43*(5), 400–407. https://doi.org/10.1046/j.1365-2788.1999.043005400.x

Deb, S., Hare, M., Prior, L., & Bhaumik, S. (2007). Dementia Screening Questionnaire for Individuals with Intellectual Disabilities. *British Journal of Psychiatry, 190*(5), 440-444. https://doi:10.1192/bjp.bp.106.024984

Elliott-King, J., Shaw, S., Bandelow, S., Devshi, R., Kassam, S., & Hogervorst, E. (2016). A critical literature review of the effectiveness of various instruments in the diagnosis of dementia in adults with intellectual disabilities. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring, 4*(1), 126–148. https://doi.org/10.1016/j.dadm.2016.06.002

Evenhuis H. M. (1996). Further evaluation of the Dementia Questionnaire for Persons with Mental Retardation (DMR). *Journal of intellectual disability research, 40*(4), 369–373. https://doi.org/10.1046/j.1365-2788.1996.786786.x

O'Caoimh, R., Clune, Y., Molloy, D. W. (2013). Screening for Alzheimer's disease in Downs syndrome. *Journal of Alzheimer's Disease & Parkinsonism, 03*(03). https://doi.org/10.4172/2161-0460.s7-001

Orange, J. B., & Zanon, M. V. (2005). Language and communication in adults with Down syndrome and dementia of the Alzheimer type: A review. *Journal on Developmental Disabilities, 12*(1), 53-62.

Prasher, V.P. (1997). Dementia Questionnaire for Persons with Mental Retardation (DMR): Modified criteria for adults with Down's syndrome. *Journal of Applied Research in Intellectual Disabilities, 10*(1), 54-60. https://doi.org/10.1111/j.1468-3148.1997.tb00006.x

Silverman, W., Krinsky-McHale, S.J., Lai, F., Diana Rosas H., Hom, C., Doran, E., Pulsifer, M., Lott, I., Schupf, N. & Alzheimer's Disease in Down Syndrome (ADDS) Consortium. (2020). Evaluation of the National Task Group-Early Detection Screen for Dementia: Sensitivity to 'mild cognitive impairment' in adults with Down syndrome. *Journal of Applied Research in Intellectual Disabilities*. https://doi.org/10.1111/jar.12849

Takenoshita, S., Terada, S., Kuwano, R., Inoue, T., Kurozumi, T., Choju, A., Suemitsu, S., & Yamada, N. (2020). Validation of the Japanese version of the Dementia Screening Questionnaire for Individuals with Intellectual Disabilities. *Journal of Intellectual Disability Research, 64*(12), 970–979. https://doi.org/10.1111/jir.12788.