Received: 1 January 2025 | Revised: 11 April 2025 | Accepted: 28 April 2025

DOI: 10.1002/alz.70307

Alzheimer's & Dementia®
THE JOURNAL OF THE ALZHEIMER'S ASSOCIATION

RESEARCH ARTICLE

Discrepancies in assessing intellectual disability levels in adults with Down syndrome: Implications for dementia diagnosis

Laura del Hoyo Soriano[1,2] | Katja Sandkühler[3] | Laura Videla[1,2,4] | Bessy Benejam[1,2,4] | Maria Carmona-Iragui[1,2,4] | Elisabeth Wlasich[3] | Julia Kustermann[3] | Isabel Barroeta[1,2] | Lídia Vaqué-Alcázar[1,5,6] | Íñigo Rodríguez-Baz[1,2] | Alexandre Bejanin[1,2] | Susana Fernández[4] | Javier Arranz[1,2] | Jose Enrique Arriola-Infante[1,2] | Lucia Maure-Blesa[1,2] | Aida San Juan[1] | Georg Nübling[3,7] | Olivia Wagemann[3,7] | Anna Stockbauer[3,7] | Jason Hassenstab[8] | Johannes Levin[3,8,9] | Juan Fortea[1,3,4]

[1]Sant Pau Memory Unit, Department of Neurology, Hospital de la Santa Creu i Sant Pau, Biomedical Research Institute Sant Pau, Barcelona, Spain

[2]Center of Biomedical Investigation Network for Neurodegenerative Diseases (CIBERNED), Madrid, Spain

[3]Department of Neurology, University Hospital, Ludwig-Maximilians-University (LMU) Munich, Munich, Germany

[4]Barcelona Down Medical Center, Fundació Catalana de Síndrome de Down, Barcelona, Spain

[5]Department of Medicine, Faculty of Medicine and Health Sciences, Institute of Neurosciences, University of Barcelona, Barcelona, Spain

[6]Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Barcelona, Spain

[7]German Center for Neurodegenerative Disease (DZNE), Munich, Germany

[8]Departments of Neurology and Psychological & Brain Sciences, Washington University in St. Louis, St. Louis, Missouri, USA

[9]Munich Cluster for Systems Neurology (SyNergy), Munich, Germany

Correspondence
Laura del Hoyo Soriano, Sant Pau Memory Unit, Department of Neurology, Hospital de la Santa Creu i Sant Pau, Biomedical Research Institute Sant Pau, C. Sant Quintí 77-79, Pavilion 18, 2nd floor, Office 42, 08041 Barcelona, Spain.
Email: lhoyo@santpau.cat

Laura del Hoyo Soriano and Katja Sandkühler share first authorship.

Johannes Levin and Juan Fortea share last authorship.

Funding information
Centro de Investigación Biomédica en Red sobre Enfermedades Neurodegenerativas, Grant/Award Number: Program1; La Caixa Foundation; Alzheimer's Association, Grant/Award Numbers: AARG-22-923680 to

Abstract

INTRODUCTION: Cut-offs derived from baseline cognitive assessments, stratified by intellectual disability (ID) level, have been proposed to diagnose symptomatic Alzheimer's disease (AD) in Down syndrome (DS). However, discrepancies in ID classification risk misclassification when applying cut-offs across sites.

METHODS: This dual-center cohort study included 673 adults with mild to moderate ID at different AD stages. We assessed ID classification discrepancies across sites and the impact on Cambridge Cognitive Examination for Older Adults with Down's Syndrome (CAMCOG-DS) cut-offs for AD dementia diagnosis derived from receiver operating characteristic analysis.

RESULTS: Inter-rater agreement for ID level classification was 95% within sites but 60% between sites. While CAMCOG-DS score distributions in the whole cohort were similar across sites, ID classification discrepancies caused higher cut-offs in Barcelona

A.B, AARG-22-973966 to M.C-I; Carlos III Health Institute, Grant/Award Numbers: INT21/00073, CM23/00291, PI20/01473, PI23/01786 to J.F, PI22/00785 to M.C-I, PI22/00307 to AB; Carlos III Health Institute and European Union, Miguel Servet Program, Grant/Award Numbers: CP24/00112 to L.D.H.S, CP20/00038 to A.B; Carlos III Health Institute, Sara Borrell, Grant/Award Numbers: CD23/00235 to L.V, CM22/00219 to J.E.A.-I, CM22/00052 to I.R.-B, CM23/00291 to L.M.B; Carlos III Health Institute and European Union Río-Hortega program, Grant/Award Numbers: CM22/00219 to J.E.A.I., CM22/00052 to I.R.B., CM23/00291 to L.M.B; National Institutes of Health, Grant/Award Numbers: U01AG024904, R01AG056850, R21AG056974, R01AG061566, R01AG081394, R61AG066543 to J.F; Department de Salut de la Generalitat de Catalunya Pla Estrategic de Recerca I Innovacio´en Salut, Grant/Award Number: LT006/17/00119 to J.F; Fundación Tatiana Pérez de Guzmán el Bueno, Grant/Award Number: IIBSP-DOW-2020-151 to J.F; Research and Innovation Framework Program, Grant/Award Number: H2020-SC1-BHC-2018-2020 to J.F; Jérôme Lejeune Foundation, Grant/Award Numbers: 2326 - GRT-2024A to L.D.H.S, AARG-22-923680 to A.B, AARG-22-973966 to M.C-I; VERUM Foundation; Deutsche Forschungsgemeinschaft DFG, German Research Foundation, under Germany's Excellence Strategy within the framework of the Munich Cluster for Systems Neurology, Grant/Award Number: EXC2145SyNergy–ID390857198; T21RS Scientific Exchange Grant to K.S; Departament de Salut, Generalitat de Catalunya, Grant/Award Number: SLT006/17/00119

for mild and moderate ID compared to Munich. Applying site-specific cut-offs to another cohort reduced sensitivity and specificity.

DISCUSSION: Standardizing ID classification is critical for generalizable cut-offs to accurately diagnose AD dementia based on neuropsychological assessments in DS.

KEYWORDS

Alzheimer's disease, AD21, Cambridge Cognitive Examination for Older Adults with Down Syndrome, cut-off points, dementia, diagnostic performance, Down Alzheimer Barcelona Neuroimaging Initiative, Down syndrome, Down syndrome–associated Alzheimer's disease, intellectual disability

Highlights

- CAMCOG-DS cut-offs by intellectual disability level classify dementia in Down syndrome.
- ID classification discrepancies between sites impact CAMCOG-DS diagnostic cut-offs.
- Applying site-specific cut-offs to other cohorts reduces sensitivity and specificity.
- Standardized ID classification is essential for generalizable cognitive cut-offs.
- Use site-specific cut-offs until ID classification is standardized.

1 | BACKGROUND

Down syndrome (DS) is the leading genetic cause of intellectual disability (ID), affecting ≈ 0.14% of the global population.[1] Due to the triplication of the amyloid precursor protein (*APP*) gene on chromosome 21, individuals with DS have near full penetrance of symptomatic Alzheimer's disease (AD) by the age of 65.[2,3] Therefore, DS-associated AD (DSAD) is recognized as a genetic form of AD by the Alzheimer's Association.[4]

Identifying the AD symptom onset in individuals with DS is challenging, as early cognitive decline is often misattributed to pre-existing cognitive impairments associated with ID,[5–7] leading to under- or delayed diagnosis. Guidelines emphasize tracking within-person cognitive changes over time,[8] but families commonly seek clinical evaluation only after noticeable changes, with no prior assessments available for comparison. In addition, recent studies indicate that longitudinal assessments in DS are influenced by high intra-individual variability, potentially masking true cognitive decline.[9]

An alternative approach involves classifying AD dementia based on cognitive test performance from a single assessment by comparing scores to established diagnostic cut-offs. In DS, however, defining cut-off scores is challenging due to substantial inter-individual variability in premorbid cognitive abilities (mild to profound ID).[10–12] Nonetheless, recent studies show that stratifying cut-offs by premorbid ID level effectively accounts for this variability. This is demonstrated by the high sensitivity and specificity of ID-specific cut-off scores, calculated from baseline neuropsychological assessments such as the Cambridge Cognitive Examination for Older Adults with Down's Syndrome (CAMCOG-DS[13]), in identifying prodromal AD and AD dementia in DS.[14] Indeed, using those ID-specific CAMCOG-DS cut-offs at baseline outperforms 1-year longitudinal changes in diagnostic accuracy.[9] These findings challenge current recommendations for diagnosing AD dementia in DS,[8] highlighting the importance of ID-level specific cut-offs and the overlooked impact of longitudinal intra-individual cognitive variability—also observed in sporadic AD.[15]

Accurate ID classification is therefore critical when establishing CAMCOG-DS thresholds.[9,14] Historically, ID levels were determined solely by IQ scores, as outlined in the International Classification of Diseases, 10th Revision (ICD-10[16]). However, IQ decreases with AD progression,[17,18] making it an unsuitable proxy for premorbid cognitive functioning in older DS adults.[19] Moreover, IQ assessments do not capture functional abilities, or the level of support required in daily life.

To address these limitations, recent frameworks such as the Diagnostic and Statistical Manual of Mental Disorders, 5th Edition (DSM-5[20]) and ICD-11 have shifted away from relying exclusively on IQ scores to classify ID severity. Instead, they now define severity solely (DSM-5) or partially (ICD-11) based on adaptive functioning across conceptual, social, and practical domains. This focus on adaptive functioning, particularly in DSM-5, 84 allows for retrospective estimation of an individual's best-ever functioning, even after AD-related cognitive decline has begun (e.g., "At their best, was the individual able to read or manage money?"). However, this introduces subjectivity, as it depends on caregiver recall and environmental factors, including how much autonomy caregivers have historically allowed (e.g., permission to work or navigate public spaces independently). Additionally, DSM-5 criteria lack scoring rules or quantitative thresholds to classify ID levels, relying on clinician judgment to synthesize caregiver reports and assess functional abilities. As a result, evaluators may weigh domains differently, leading to inconsistencies in ID classification. This subjectivity is compounded by cross-cultural differences in expectations, training, and health-care systems, presenting a challenge to harmonizing ID classification across international research.

The subjectivity in DSM-5–based ID classification limits the use of baseline neuropsychological assessments to diagnose symptomatic AD. Inconsistent classification across studies can influence subgroup composition, as individuals categorized with the same ID level may differ in premorbid functioning. This variability affects the derivation and generalizability of ID-specific cognitive cut-off scores.

To our knowledge, no previous study has examined ID classification discrepancies within and between research sites, or their impact on dementia diagnosis in DS. This study addresses that gap by: (1) assessing inter-rater agreement in DSM-5–based ID classification, (2) CAMCOG-DS score distributions by ID level using DSM-5 and ICD-10 frameworks in asymptomatic DS, and (3) evaluating how ID level classification discrepancies affect the cross-site applicability of CAMCOG-DS cut-offs for AD dementia, using data from two large, well-characterized DS cohorts.

RESEARCH IN CONTEXT

1. **Systematic review**: We conducted a literature review using databases like PubMed and Google Scholar to examine methods for classifying intellectual disability (ID) and their impact on baseline cognitive assessment thresholds for identifying prodromal and dementia stages of Alzheimer's disease (AD) in adults with Down syndrome (DS). Few studies addressed whether ID classification varies across research sites or how it affects cut-offs and diagnostic accuracy.
2. **Interpretation**: Our findings highlight considerable variation in DSM-5-based ID classification across centers. These discrepancies affect the accuracy of baseline neuropsychological assessments in international studies. In particular, the sensitivity of CAMCOG-DS thresholds to sample composition illustrates the risk of using globally derived cut-offs, which may compromise dementia diagnosis reliability in DS.
3. **Future directions**: There is a clear need for standardized ID classification across research settings. Until such frameworks exist, site-specific thresholds should be used to ensure consistent and accurate diagnoses.

2 | METHODS

2.1 | Study design, setting, and participants

This cross-sectional study includes data from 673 adults with DS from two cohorts: the Down Alzheimer Barcelona Neuroimaging Initiative (DABNI) cohort in Barcelona, Spain[3,21] and the AD21 cohort in Munich, Germany. The DABNI cohort was recruited from the Alzheimer-Down Unit at the Catalan Down Syndrome Foundation and the Hospital of Sant Pau. The AD21 cohort was recruited from the outpatient clinic for adults with DS, embedded in the Department of Neurology at the University Hospital, Ludwig-Maximilians University (LMU) Munich. Both units implement population-based health plans focused on neurological conditions, particularly AD, in individuals with DS, with patients usually undergoing semi-annual or annual structured neurological and neuropsychological assessments by experienced clinicians. For the purposes of this study, only data from the initial visit were included.

We included individuals with DS (≥ 18 years old, of both sexes) with mild or moderate ID. Those with severe or profound ID were excluded due to low completion rates of the CAMCOG-DS.[14] We also excluded individuals with unstable medical, pharmacological, or psychiatric conditions interfering with cognition. Figure S1 in supporting information shows the study flowchart.

The study was approved by the Sant Pau Ethics Committee and by the local ethics committee of the LMU medical faculty, following the standards for medical research in humans recommended by the Declaration of Helsinki. All participants or their legally authorized representatives gave written informed consent before enrolment. All data were anonymized according to good clinical practice guidelines and general data protection regulations prior to analysis.

TABLE 1 Summary of DSM-5 ID classification criteria.

ID severity levels	Conceptual domain	Social domain	Practical domain
Mild	- Academic skills up to 6th grade level - May need support with complex problem solving, money management, or time concepts	- Capable of meaningful relationships - Difficulties with interpreting subtle social cues - May show immature social judgment	- Independent in basic self-care - May need support for complex tasks (e.g., shopping, cooking, transportation) - May live semi-independently with minimal support
Moderate	- Academic skills typically reach 2nd grade level - Learns through concrete instruction - Limited understanding of money and time	- Capable of basic conversation and social interactions - Limited understanding of social norms - May be vulnerable to social manipulation	- Performs personal care with moderate supervision - Requires training and support for household tasks and community navigation - May live in supported settings
Severe	- Very limited academic understanding (e.g., basic counting, simple words) - Requires extensive support for all learning	- Limited verbal communication; may use gestures or AAC - Difficulties with social reciprocity - Often needs support to avoid isolation	- Requires assistance for all daily living activities (e.g., dressing, eating) - Cannot manage routines independently - Constant supervision needed
Profound	- Minimal symbolic understanding - May respond primarily to visual or tactile cues	- Very limited or no verbal communication - May express needs through non-verbal behavior - Dependent on others for all social interaction	- Fully dependent on all activities of daily living - Requires constant care - Safety and health supervision essential

Note: ID severity levels: Functional expectations by domain.
Abbreviations: AAC, argumentative and alternative communication; DSM-5, Diagnostic and Statistical Manual of Mental Disorders, 5th Edition; ID, intellectual disability.

2.2 | Study outcomes

2.2.1 | ID level classification according to DSM-5

At both sites, ID level was categorized as mild, moderate, severe, or profound based on the caregivers' reports of the participant's best-ever functional skills across conceptual, practical, and social domains according to DSM-5 criteria.[20] As mentioned in the previous section, those with severe to profound ID were excluded from the current study. ID level classification was carried out by a neuropsychologist via semi-structured interviews conducted as part of the clinical routine practice at both sites, aiming to include all aspects from the DSM-5 (see Table 1 for details). While the structure of the interview was not formally standardized, similar guiding questions were typically used within sites (e.g., "Does the participant use public transportation independently?"). As the DSM-5 does not provide specific scoring rules or thresholds to distinguish between ID levels, no standardized operational procedure (SOP) or predefined coding framework was applied. Instead, classification relied on the neuropsychologist's clinical judgment to classify a patient into an ID level, drawing on the information gathered during the interviews and their overall impression of the participant's behavior and functioning.

As the DSM-5 does not provide explicit scoring criteria, we evaluated the percent agreement in DSM-5–based ID classification within and between sites to assess inter-rater agreement. For within-sites rating, co-first authors, both experienced neuropsychologists in assessing individuals with DS, independently reviewed the clinical reports used to determine ID level by other neuropsychologists of their corresponding site. For between-sites rating, co-first authors (both fluent in English) translated a subset of randomly selected participant records from their respective sites (Spanish or German) into English, ensuring that all functional descriptions based on DSM-5 criteria were preserved. Both raters were blinded to the original ID classification but were aware of the participant's age, IQ, and AD diagnosis. A random selection of reports from 29 asymptomatic individuals and 23 patients with AD dementia of both sites was analyzed. Percent agreement was calculated for the following categories: (1) inter-rater agreement within Barcelona, (2) inter-rater agreement within Munich, and (3) inter-rater agreement across sites.

2.2.2 | ID level classification according to ICD-10

Only for the purpose of the current study and to enable comparison between classification frameworks, participants were retrospectively classified into ID levels based on ICD-10 criteria, which define severity using IQ score ranges: mild ID (IQ 50–69) and moderate ID (IQ 35–49). IQ testing was conducted at the initial visit in young, asymptomatic individuals only. IQ tests are not suitable for individuals with dementia, as scores reflect current cognitive decline rather than best-ever cognitive ability. Additionally, due to cognitive impairment, individuals with dementia are often no longer able to complete the complex tasks required by these assessments (nor those with severe to profound ID). Consequently, IQ data were not available for the AD dementia groups.

At the Barcelona site, the Spanish version of the Kaufman Brief Intelligence Test (KBIT-1),[22] was used, assessing verbal (vocabulary) and non-verbal (matrices) cognition across ages 4 to 90. The total score reflects overall cognitive ability, and the site regularly uses it to

determine IQ. At the Munich site, the Coloured Progressive Matrices (CPM[23]) was initially used. However, as the CPM does not provide IQ norms for adults, it was later replaced by the more recent Raven's Progressive Matrices 2 (Raven's 2[24]), which uses sets A–C (12 items each; max 36 points) for adults with ID. The Raven's 2 is a matrix-based, non-verbal IQ test with a structure and theoretical foundation similar to the Matrices subtest of the KBIT-1. For comparability across sites, only non-verbal IQ scores were used for ID classification.

2.2.3 | Clinical stage of AD

At both sites, people with DS were classified into the following diagnostic categories along the AD continuum: (1) asymptomatic (aDS): no clinical or neuropsychological suspicion of symptomatic AD (absence of cognitive impairment beyond the ID or functional decline compared to previous functioning); (2) prodromal AD (pDS): suspicion of AD-related cognitive impairment, but symptoms did not fulfil criteria for dementia (i.e., no additional functional impairment in activities of daily living); (3) AD dementia (dDS): evidence of overt cognitive impairment that interferes with everyday activities beyond baseline functioning. Importantly, for the current study we used the diagnosis ascertained by one or two neurologists who conducted and/or reviewed the clinical exam data but were blinded to the neuropsychological assessment to avoid circular analyses.

2.2.4 | The CAMCOG-DS

The CAMCOG-DS was administered by experienced neuropsychologists using the corresponding validated language/cultural version at each site: The Spanish version in Barcelona[25] and the German version[26] in Munich. This comprehensive cognitive battery assesses abilities across multiple domains, including orientation, language, memory, attention, praxis, abstract thinking, and perception. Specifically designed for adults with DS, the CAMCOG-DS aims to detect AD-related cognitive decline, aiding in the identification of prodromal and dementia stages of AD. The test has a maximum score of 109, with higher scores indicating better cognitive performance.

2.3 | Statistical analysis

Statistical analyses were conducted in R (version 4.2.2[27]). Significance was set at $P < 0.05$. We defined outliers as CAMCOG-DS values beyond 3 standard deviations (SDs) from the mean. Three asymptomatic participants were excluded according to this criterion.

We stratified the analyses by diagnostic group (aDS, dDS). Within diagnostic group, we compared demographic characteristics, such as age and sex, between the Barcelona and Munich cohorts. Categorical variables were analyzed using χ^2 tests, while independent samples *t* tests were applied to continuous variables. We compared cognitive performance between sites based on CAMCOG-DS total scores. To quantify the effect, we calculated Cohen *d*, with values of $|d| > 0.2$ indicating a small, $|d| > 0.5$ a medium, $|d| > 0.8$ a large, and $|d| > 1.2$ a very large effect.[28] Moreover, we applied *F* tests to compare the variance (SD) as well as Kolmogorov–Smirnov (K–S) two-sample tests to compare the score distributions of the CAMCOG-DS total scores between sites.

2.3.1 | ID level classification based on DSM-5

In a second step, we conducted the same analyses after stratifying participants by ID level, as classified at each site according to DSM-5 criteria. To assess whether site differences could be explained by variations in ID classification practices, we implemented a two-step matching strategy to enhance comparability between the Barcelona and Munich cohorts on other factors. First, we matched aDS participants between Munich and Barcelona based on cognitive abilities (i.e., CAMCOG-DS scores) via the MatchIt R package[29] with optimal matching based on Mahalanobis distance. Second, we accounted for potential sociodemographic confounders by matching the aDS samples from Munich and Barcelona on age (optimal matching) and sex (exact matching). This approach aimed to reduce systematic differences between cohorts, thereby enhancing the validity of cross-site comparisons.

In a next step, we defined cut-offs to classify AD-related cognitive impairment for the CAMCOG-DS with two different approaches. (1) Using the receiver operating characteristic (ROC) curve analyses, the index of union (IU) method (defined as the value at which the difference between the sensitivity and specificity is minimized) to define the optimal cut-off,[30] and the DeLong test to compare area under the curve (AUC). (2) Defining the scores at percentile ranks of 1st, 5th, and 10th in the young aDS participants (age ≤ 35) in mild and moderate ID (classified by DSM-5) separately.[19] The decision to include participants ≤ 35 years was based on the fact that individuals with DS develop AD neuropathology by their fourth decade, increasing the likelihood of cognitive decline in those > 35 years. Although we included only asymptomatic cases, individuals > 35 years were excluded to ensure the analysis focused on participants without AD-related cognitive decline.[3,21] Due to the limited number of pDS cases ($n = 7$) in Munich, all pDS participants were excluded from the ROC analyses.

Finally, we applied the cut-off points that we determined with the ROC analysis to the data from the other study cohort (i.e., cut-offs from Barcelona applied to the Munich cohort and vice versa) and calculated sensitivity and specificity. This approach allowed us to evaluate the generalizability of the cut-offs and to assess whether the diagnostic accuracy changes when applied to another cohort.

2.3.2 | ID level classification based on ICD-10

In an exploratory analysis, we also compared CAMCOG-DS scores between sites in the aDS groups, stratified by ID level based on ICD-10 criteria (using non-verbal IQ cut-offs, mild and moderate ID only), with independent samples *t* tests. However, due to the small sample size for IQ test scores in the Munich cohort ($N = 36$), the results should be

considered preliminary. Importantly, as IQ data were not available for the dDS groups, all analyses involving dementia-related cut-off scores could only be conducted using ID levels classified according to DSM-5 criteria.

Furthermore, to assess the potential impact of IQ test modality (verbal vs. non-verbal) on ID classification under ICD-10, we used the McNemar test to compare the proportion of individuals classified as mild versus moderate ID based on non-verbal IQ (KBIT Matrices) versus verbal IQ (KBIT Vocabulary) within the Barcelona cohort, including only those with scores on both tests.

3 | RESULTS

3.1 | Demographic characteristics of the study population

Our sample consisted of 498 aDS ($N_{Barcelona}$ = 420, N_{Munich} = 78), 81 pDS ($N_{Barcelona}$ = 74, N_{Munich} = 7), and 110 dDS adults ($N_{Barcelona}$ = 84, N_{Munich} = 26). Table 2 shows the demographics and cognitive scores in both cohorts.

There were no significant differences in sex distribution between sites for either aDS ($P = 0.88$) or for dDS ($P = 0.36$). The mean age of the aDS sample was significantly higher in Barcelona ($M_{Barcelona}$ = 37.03) compared to Munich (M_{Munich} = 30.37), $t(496) = 5.66$, $P < 0.001$, $d = 0.70$. No significant age differences were observed between the dDS samples from Barcelona and Munich ($P = 0.20$). Importantly, Barcelona had a significantly lower proportion of aDS individuals with mild ID (37.4%) compared to Munich (64.1%), χ^2 (1) = 18.30, $P < 0.001$. There were no significant differences in ID levels in the dDS samples ($P = 0.30$). See Tables S1 and S2 in supporting information for additional sociodemographic information on the two cohorts. While some variables were not directly comparable due to site-specific data collection, both cohorts showed similar trends: most participants were single, approximately half lived in rural areas, and the majority attended special-needs schools.

3.2 | ID level classification according to DSM-5

3.2.1 | Impact of ID level classification on cognitive performance in asymptomatic participants

Figure 1A shows the distribution of the CAMCOG-DS scores for both sites. We found no significant differences in the CAMCOG-DS scores between Barcelona and Munich in the total sample of aDS individuals, for mean scores ($M_{Barcelona} = 75.45$, $M_{Munich} = 72.14$, $t[496] = 1.89$, $P = 0.06$), variance ($F[419, 77] = 1.06$, $P = 0.76$), and score distribution (K–S test, $D = 0.14$, $P = 0.15$).

When stratifying samples by ID level based on DSM-5 criteria, the distribution of the CAMCOG-DS total scores differed notably between the Barcelona and the Munich cohort. The K–S test revealed a significant difference in the CAMCOG-DS score distribution between sites

TABLE 2 Demographic and clinical descriptive data.

	Barcelona (n = 562)			Munich (n = 111)		
	aDS	dDS	pDS	aDS	dDS	pDS
***n* Total**	420	84	58	78	26	7
Age years (M, [SD]), range	37.03 (9.67), 18–61	52.91 (5.70), 39–68	51.64 (4.74), 40–62	30.37 (8.78), 18–54	54.58 (5.67), 43–63	45.86 (9.58), 33–59
Sex female[a]**, *n* (%)**	203 (48%)	43 (51%)	31 (53%)	39 (50%)	10 (38%)	4 (57%)
ID level (DSM-5) mild/moderate, *n* (%)	157 (37%)/263 (63%)	13 (15%)/71 (85%)	12 (21%)/46 (79%)	50 (64%)/28 (36%)	7 (27%)/19 (73%)	4 (57%)/3 (43%)
CAMCOG-DS total (M, [SD]), range	75.45 (14.29), 33–103	46.10 (15.75), 11–85	63 (19.5), 33–92	72.14 (13.86), 30–94	43.81 (16.84), 15–78	65.43 (16.09), 38–93
Non-verbal IQ[b] **(M, [SD]), *n***	57.03 (7.61), 352			58.56 (8.00), 36		

Abbreviations: aDS, asymptomatic Down syndrome; CAMCOG-DS, Cambridge Cognitive Examination for Older Adults with Down Syndrome; dDS, dementia stage Down syndrome; DSM-5, Diagnostic and Statistical Manual of Mental Disorders, 5th Edition; ID, intellectual disability; M, mean; pDS, prodromal stage Down syndrome; SD, standard deviation.

[a]Assigned at birth.

[b]Non-verbal IQ was assessed in Barcelona with the Matrices subtest of the Kaufman Brief Intelligence Test (KBIT-1) and in Munich with the Raven's Progressive Matrices 2.

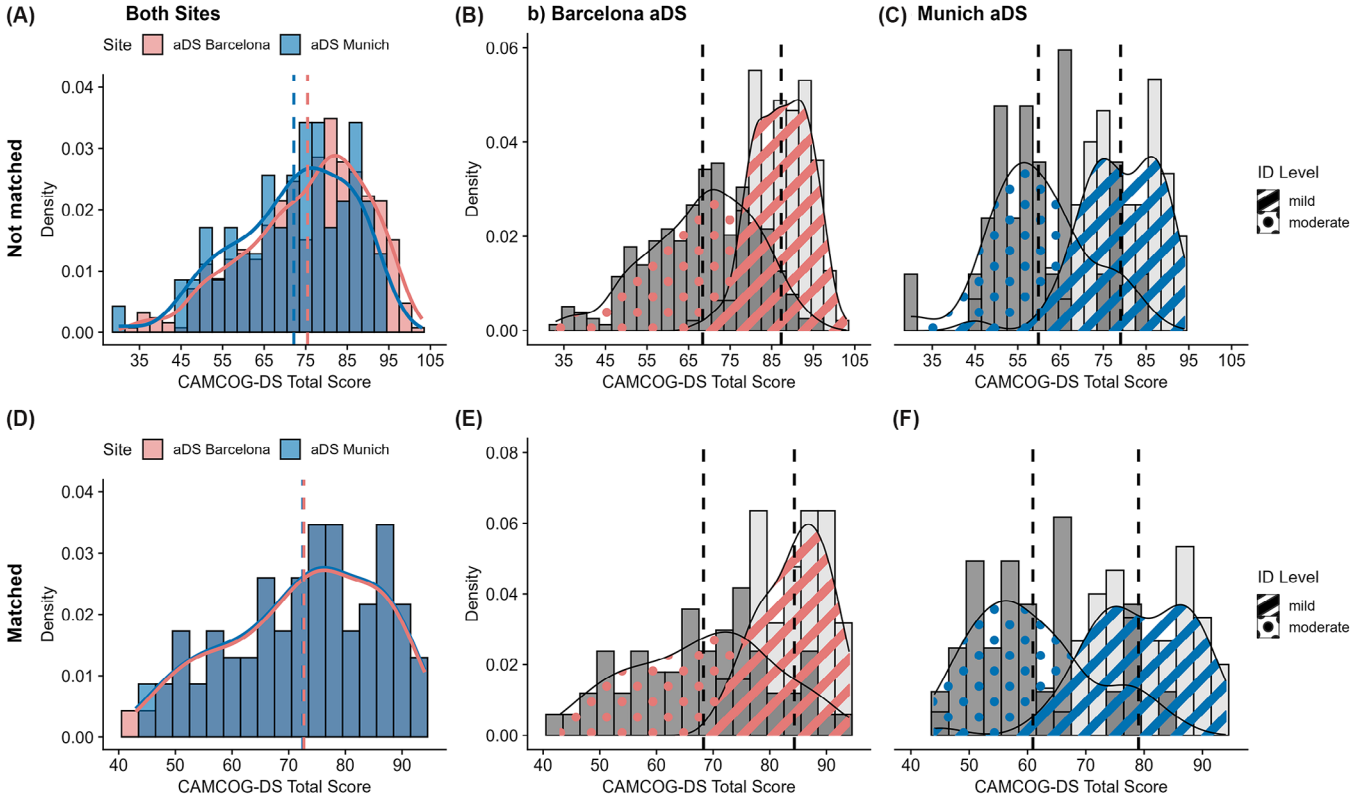

FIGURE 1 A, Density plot of the CAMCOG-DS total scores for the total asymptomatic (aDS) sample (not matched) in both sites ($N = 420$ Barcelona, $N = 78$ Munich) and density plots split by ID level with density curves for (B) Barcelona and (C) Munich. Dashed lines represent the means. D, Density plot of the CAMCOG-DS for the aDS samples matched on CAMCOG-DS scores ($N = 77$ Barcelona, $N = 77$ Munich) and density plots split by ID level with density curves for (E) Barcelona and (F) Munich. aDS, asymptomatic Down syndrome; CAMCOG-DS, Cambridge Cognitive Examination for Older Adults with Down Syndrome; ID, intellectual disability.

for both mild ID ($D = 0.40$, $P < 0.001$) and the moderate ID groups ($D = 0.39$, $P < 0.001$). In the mild ID group, the Barcelona cohort had significantly higher CAMCOG-DS scores compared to the Munich cohort ($M_{Barcelona} = 87.29$, $M_{Munich} = 79.04$, $t[205] = 6.77$, $P < 0.001$, $d = 1.1$, large effect size). A similar pattern was observed in moderate ID, where Barcelona again had higher scores ($M_{Barcelona} = 68.39$, $M_{Munich} = 59.82$, $t[289] = 3.38$, $P < 0.001$, $d = 0.67$, medium effect size). Moreover, the Barcelona cohort showed a smaller variance ($SD_{Barcelona} = 6.74$) than the Munich cohort ($SD_{Munich} = 9.52$, $F[156] = 0.50$, $P = 0.002$), indicating a narrower range of CAMCOG-DS scores within the mild ID group in Barcelona. We found no significant variance difference for the moderate ID subgroups ($SD_{Barcelona} = 12.87$, $SD_{Munich} = 11.75$, $F[262] = 1.2$, $P = 0.59$). In short, in the mild ID group, the distribution of the CAMCOG-DS scores in the Barcelona cohort was narrower and shifted toward higher scores (Figure 1B, shown in red stripes) compared to the Munich sample (Figure 1C, shown in blue stripes).

We repeated the analyses after matching the aDS samples based on CAMCOG-DS total score (Figure 1, lower row). Despite being identical in CAMCOG-DS scores at a cohort level (Figure 1D, $M_{Barcelona} = 72.68$, $M_{Munich} = 72.68$), the mean scores still differed between sites when stratified by ID level for both mild (Figure 1E, $M_{Barcelona} = 84.33$, $M_{Munich} = 79.04$, $t[69] = 2.36$, $P = 0.02$, $d = 0.61$) and moderate ID (Figure 1F, $M_{Barcelona} = 68.30$, $M_{Munich} = 60.93$, $t[81] = 2.68$, $P = 0.009$, $d = 0.63$). The same pattern of results was observed when matching the cohorts by age and sex (Figure S2, S3 and Table S3 in supporting information).

3.2.2 | Impact of ID level classification on cognitive performance in symptomatic participants

We found no differences in the CAMCOG-DS total scores in the dDS group between Munich and Barcelona regarding the mean ($t[108] = 0.64$, $P = 0.53$), the distributional shape ($D = 0.12$, $P = 0.83$) or the variance ($F[83, 25] = 0.88$, $P = 0.64$). When split by ID level according to DSM-5 (see Figure 2), the analyses revealed no significant differences in CAMCOG-DS total scores between Barcelona and Munich, despite again a trend toward higher scores in Barcelona. For individuals with mild ID, the mean scores were $M_{Barcelona} = 58.39$ and $M_{Munich} = 53.43$ ($t[18] = 0.50$, $P = 0.62$), and for those with moderate ID, mean scores were $M_{Barcelona} = 43.85$, $M_{Munich} = 40.26$ ($t[88] = 1.02$, $P = 0.31$). However, the small sample sizes when stratifying by ID level—particularly in the mild ID subgroup ($N_{Barcelona} = 13$, $N_{Munich} = 7$)—warrant cautious interpretation of the results. Note that for this same reason a matching procedure was not applied.

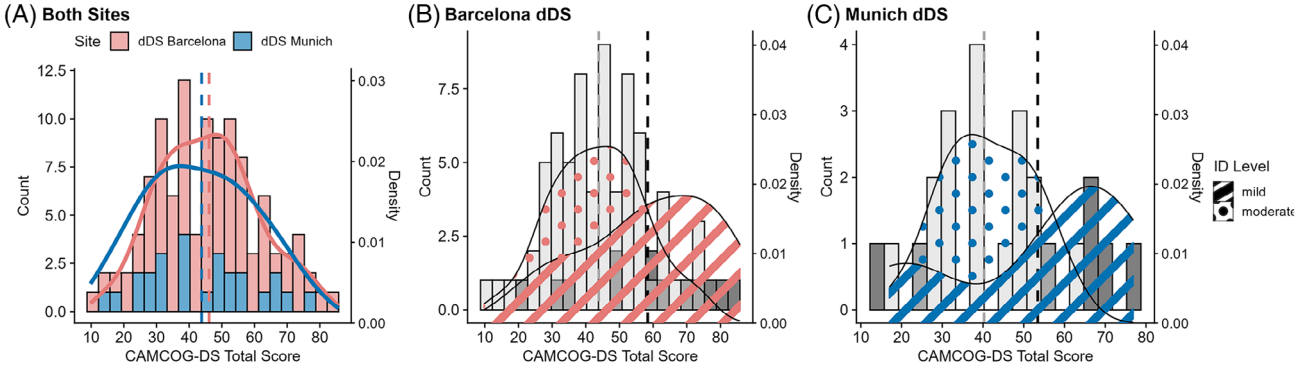

FIGURE 2 A, Histogram of the CAMCOG-DS total scores only for individuals with a dementia diagnosis (dDS) at first visit for both sites ($N = 84$ Barcelona, $N = 26$ Munich) and split by ID level for (B) Barcelona and (C) Munich. Dashed lines represent the means. CAMCOG-DS, Cambridge Cognitive Examination for Older Adults with Down Syndrome; dDS, dementia stage Down syndrome; ID, intellectual disability.

TABLE 3 Inter-rater agreement for DSM-5-based ID classification.

	aDS	dDS	Total agreement
Inter-rater agreement between sites			
Barcelona rates Munich cases	7/10 (70%)	6/10 (60%)	65%
Munich rates Barcelona cases	16/19 (84.2%)	10/13 (76.9%)	81.3%
Inter-rater agreement within sites			
Barcelona rates Barcelona cases	18/19 (94.7%)	11/13 (84.6%)	90.6%
Munich rates Munich cases	9/10 (90%)	8/10 (80%)	85%

Abbreviations: aDS, asymptomatic Down syndrome; dDS, dementia stage Down syndrome.

3.2.3 | Inter-rater agreement

The results of the inter-rater agreement for DSM-5–based ID classification are summarized in Table 3. As observed, the inter-rater agreement within sites (ranging from 80% to 94.7%) was higher compared to the agreement of the examiners between sites (ranging from 60% to 84.2%).

3.2.4 | Diagnostic performance for the CAMCOG-DS by ID level across sites

Two-way analyses of variance revealed significant effects of diagnosis and ID level on CAMCOG-DS scores in both Barcelona ($F[2, 556] = 234.67$, $P < 0.001$; $F[1, 556] = 306.55$, $P < 0.001$) and Munich ($F[2, 105] = 53.17$, $P < 0.001$; $F[1, 105] = 53.22$, $P < 0.001$). Figure 3 shows cognitive scores in the CAMCOG-DS along the AD continuum by level of ID in both sites. This figure shows a clear gradient in the scores both between ID levels and between the clinical stages of the AD continuum, clearly emphasizing the importance of stratifying by ID when defining cut-off points to classify normal versus impaired cognition for the CAMCOG-DS.

Figure 4 shows the ROC curves for the CAMCOG-DS in distinguishing between normal versus AD-related cognitive impairment in mild ID and moderate ID in each cohort separately. The ROC analysis demonstrated similarly high AUC values for accurately classifying AD-related cognitive impairment in both sites. For the subgroup with mild ID, the AUC was 0.95 with a sensitivity of 85% and a specificity of 88% for Barcelona, and 0.89 with a sensitivity of 86% and a specificity of 86% for Munich ($D = -0.81$, $P = 0.42$). However, in accordance with the discrepancies in ID level classification across sites, the optimal cut-off scores were different (80 for Barcelona, 70 for Munich). Similar results were obtained in the moderate ID subgroup: the AUC were 0.90 for Barcelona with a sensitivity of 82%, and a specificity of 83%, and 0.89 (with a sensitivity of 79%, and a specificity of 86%) for Munich ($D = -0.16$, $P = 0.87$). The cut-off score was higher in Barcelona than in Munich (56 vs. 50).

Applying Barcelona's cut-offs (mild ID: 80, moderate ID: 56) to the Munich cohort resulted in a substantial drop in specificity, decreasing to 48% for mild ID and 64% for moderate ID, while sensitivity remained relatively high (100% and 89%, respectively). Conversely, when Munich's cut-offs (mild ID: 70, moderate ID: 50) were applied to the Barcelona cohort, sensitivity decreased, reaching 62% for mild ID and 63% for moderate ID, despite high specificity values of 99% and 92%, respectively.

3.2.5 | Normative data for the CAMCOG-DS by ID level across sites

We finally derived normative data for the CAMCOG-DS in the younger subjects (≤ 35 years: $N = 245$ subjects). Normative data was generated

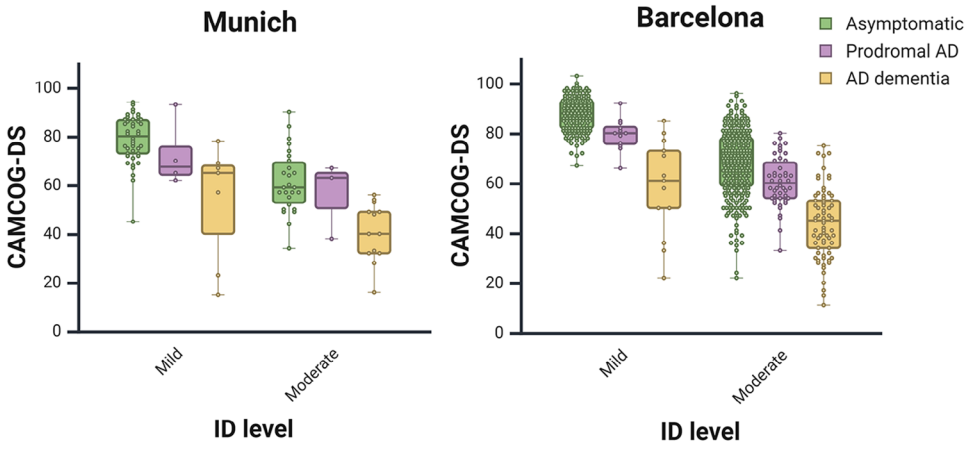

FIGURE 3 CAMCOG-DS total scores by ID level and AD dDiagnosis across sites. Boxplots illustrate the CAMCOG-DS performance by ID level (mild and moderate) and AD diagnosis (asymptomatic, prodromal AD, AD dementia) for Barcelona and Munich. The figure shows that ID level exerts a stronger influence on CAMCOG-DS scores than AD diagnosis, with a clear decline in scores from mild to moderate ID regardless of AD status. This highlights the significant role of ID classification in interpreting cognitive decline at a cross-sectional level. AD, Alzheimer's disease; CAMCOG-DS, Cambridge Cognitive Examination for Older Adults with Down Syndrome; ID, intellectual disability.

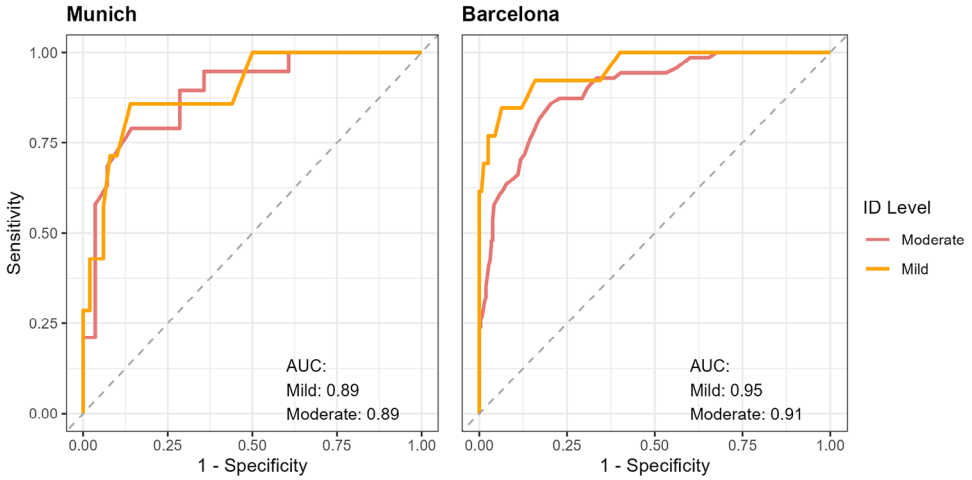

FIGURE 4 ROC CAMCOG-DS diagnostic performance. ROC curves illustrating the CAMCOG-DS diagnostic performance for AD dementia in individuals with mild and moderate ID across sites (Munich and Barcelona). The AUC values, ranging from 0.89 to 0.95, indicate high diagnostic accuracy, with confidence intervals provided for each threshold. Barcelona's data show higher CAMCOG-DS cut-off points for mild and moderate ID compared to Munich's cohort. AD, Alzheimer's disease; AUC, area under the curve; CAMCOG-DS, Cambridge Cognitive Examination for Older Adults with Down Syndrome; ID, intellectual disability; ROC, receiver operating characteristic.

in aDS individuals for the overall group and for mild and moderate ID separately for each site. Scores corresponding to the 1st, 5th, and 10th percentiles were used to establish thresholds for impaired cognitive performance (see Table S4 in supporting information). In the overall young (≤ 35 years) asymptomatic population, the CAMCOG-DS cut-points corresponding to the 1st, 5th, and 10th percentiles were similar between the two sites. For the overall population, cut-points were 39, 52, and 58 for Barcelona ($N_{Barcelona}$ = 188), and 38, 48, and 52 for Munich (N_{Munich} = 57). However, the results were substantially different across sites when splitting by level of ID. For subjects with mild ID, the cut-off points were 70, 75, and 78 for Barcelona ($N_{Barcelona}$ = 81) and 51, 63, and 68 for Munich (N_{Munich} = 36). In subjects with moderate ID, cut-points corresponding to these percentiles were as follows: 36, 48, and 53 for Barcelona ($N_{Barcelona}$ = 107) and 33, 44, and 49 for Munich (N_{Munich} = 21), respectively.

3.3 | ID level classification according to ICD-10 criteria

IQ test scores were only available for a subset of aDS participants for the Munich site ($N_{Barcelona}$ = 352, N_{Munich} = 36). CAMCOG-DS

scores of aDS participants were not significantly different between sites at the group level ($M_{Barcelona}$ = 75.03, M_{Munich} = 74.22, t[386] = 0.34, P = 0.74). When stratifying by ID level based on ICD-10 criteria, no significant differences in the CAMCOG-DS total score were found between sites for individuals with mild ($M_{Barcelona}$ = 77.20, M_{Munich} = 75.03, t[318] = 0.93, P = 0.36) or moderate ID ($M_{Barcelona}$ = 65.02, M_{Munich} = 67.75, t[66] = 0.36, P = 0.92).

In addition, verbal versus non-verbal IQ significantly impacted ID level classification based on ICD-10 at the Barcelona site, χ^2 (1, N = 332) = 127, P < 0.001. In more detail, based on non-verbal IQ, 82% of individuals were classified as having mild ID and 18% as moderate ID, while based on verbal IQ, 39% were classified as mild and 61% as moderate ID (see Figure S4 in supporting information).

4 | DISCUSSION

This study is the first to demonstrate that discrepancies in DSM-5–based ID classification across sites affect cut-off scores from baseline neuropsychological assessments (e.g., CAMCOG-DS) for classifying AD dementia. While ID-specific CAMCOG-DS cut-offs showed excellent diagnostic accuracy within site, their performance dropped substantially when applied across sites. Crucially, these differences between cohorts were not driven by cognitive variations, but rather by inconsistencies in how DSM-5 criteria are interpreted to assign ID levels.

This is best illustrated by a pattern resembling Simpson's paradox:[31] While CAMCOG-DS scores in aDS were comparable across cohorts at group level, substantial differences emerged when data were stratified by ID level. Specifically, participants classified as mild ID in the DABNI cohort (Barcelona) exhibited higher mean CAMCOG-DS scores and lower score variability than those in the AD21 cohort (Munich). This persisted even after matching cohorts by the CAMCOG-DS total score, age, or sex. A similar but non-significant trend was observed in the dDS group, likely due to low statistical power. Notably, clinicians in Munich classified a much larger proportion of aDS as having mild ID (64.1%) than clinicians in Barcelona (37.4%). These findings suggest that the Barcelona site applied DSM-5 criteria more conservatively, resulting in fewer individuals being assigned to the mild ID group, but with higher CAMCOG-DS scores and lower variability, compared to the wider interpretation of mild ID observed in Munich. As a result, individuals who might have been classified as mild ID in Munich were instead grouped as moderate ID in Barcelona, leading to higher CAMCOG-DS scores for individuals with moderate ID in the DABNI cohort. Importantly, these differences do not imply one approach is more accurate than the other but rather reflect the inherent subjectivity of DSM-5 criteria.

Inter-rater analysis of ID classification revealed agreement rates as low as 60% between sites but as high as 95% within sites, supporting our hypothesis that these discrepancies likely arise from differences in how DSM-5 criteria are interpreted across sites, rather than from actual cohort differences. Without a quantifiable framework for evaluating and mapping the DSM-5 domains (conceptual, social, practical) to ID levels, it remains unclear whether certain aspects—such as verbal skills, use of public transportation, or financial autonomy—are weighted differentially. Differences in clinical training across sites may contribute to inconsistent interpretations. These discrepancies may be further influenced by cultural and systemic differences between countries, including expectations of independence, availability of support services, and family involvement. However, lower inter-rater agreement between sites, compared to within-site ratings, may also have been influenced by methodological limitations, as clinical reports were retrospectively reviewed and translated into English.

As anticipated, differences in ID level classification impacted CAMCOG-DS cut-off scores for AD dementia. Cut-off scores were higher in the DABNI cohort for both mild and moderate ID compared to the AD21 cohort. In fact, applying site-specific cut-offs to the other cohort led to a substantial drop in sensitivity and specificity, demonstrating that these cut-offs are not generalizable due to differences in ID level classification. However, these classification differences did not impact the diagnostic performance within sites (i.e., high AUC values for cut-offs at both sites), highlighting both their value at site level but also their sensitivity to sample composition, consistent with previous findings.[9] Notably, only the 10th percentile normative values obtained in young asymptomatic individuals were comparable to those ROC-derived cut-offs, suggesting that, despite high AUCs, ROC thresholds should not be used in isolation, as they may misclassify a substantial proportion of asymptomatic individuals as symptomatic.

As expected, applying ICD-10 criteria based solely on IQ cut-offs to classify ID levels eliminated the discrepancies observed with DSM-5 classification. While this result needs to be interpreted with caution due to the small sample size, it suggests that using premorbid IQ scores to derive CAMCOG-DS cut-offs may reduce between-site variability by providing a more objective reference. This outcome is also internally consistent, as both measurements are targeting cognition—making IQ-based classification naturally result in more comparable cognitive performance across sites than classification based on adaptive functioning. While this supports the value of IQ as a proxy for premorbid cognitive functioning and for setting cognitive cut-offs, it is important to note that ID is a broader concept than just cognitive ability—it also includes adaptive functioning, which IQ alone cannot capture.

Moreover, IQ scores are only valid proxies for premorbid cognition if obtained before cognitive decline onset, limiting their use in symptomatic individuals. In fact, this limitation prevented us from calculating ROC-based cut-offs. Additionally, most IQ tests are not validated for DS populations, leading to misclassification due to the lack of norms for IQ scores < 40, which limits their use to distinguish between moderate (IQ 35–49) and severe ID (IQ 20–34) under ICD-10. While alternative methods such as deviation scores, mental age equivalents, or discrepancy indices (e.g., Stanford–Binet Intelligence Scale Fifth Edition profiles) may offer greater granularity, they remain limited by a lack of linguistic adaptation and applicability in adults with DS. Extrapolating IQ scores < 40 via regression models relies on a narrow normative range and may produce misleading estimates when applied beyond it. A more viable alternative is the use of DS-specific cognitive

TABLE 4 Comparison between DSM-5 versus ICD-10 criteria for ID classification in the context of DSAD.

Feature	DSM-5	ICD-10
Primary focus	Functioning in daily life and support needs	Cognitive performance via direct testing relative to population norms
ID level classification	Based on deficits in adaptive functioning across 3 domains (Conceptual, Social, Practical): Mild, Moderate, Severe, Profound	Based on IQ thresholds: Mild (IQ 50–69), Moderate (IQ 35–49), Severe (IQ 20–34), Profound (IQ < 20)
Assessment method	Clinical evaluation, caregiver reports, tools assessing adaptive behavior (e.g., ABAS)	IQ testing
Use of IQ	Optional; IQ can aid but does not define severity level	IQ defines severity level
Applicability in DS	Uses informant ratings, enabling assessment across all ability levels	Limited applicability due to floor effects of IQ scores
Use in DSAD	Can be retrospectively estimated via caregivers report of best-ever functioning	IQ does not reflect best-ever functioning in DSAD; Retrospective use limited by availability and timing of IQ testing.
Strengths	Captures real-world functioning; adaptable to clinical context; available in several languages.	Objective, norm-based classification with clear cut-offs
Limitations	Lacks operational scoring rules; lack cut-offs; inter-rater variability	Problematic in DSAD due cognitive decline; many tests lack norms < IQ 40; results vary by test modality

Abbreviations: ABAS, adaptive behavior assessment system; DS, Down syndrome; DSAD, Down syndrome–associated Alzheimer's disease; DSM-5, Diagnostic and Statistical Manual of Mental Disorders, 5th Edition; ICD-10, International Classification of Diseases, 10th Revision; IQ; intellectual quotient.

norms for standardized instruments, such as the KBIT-1 (e.g., DABNI cohort norms available upon request), which can improve classification accuracy, especially for individuals in the lower cognitive range. As our results showed, multi-center studies must also be consistent with the IQ test modality—verbal, non-verbal, or composite—to ensure valid cross-site comparisons. See Table 4 for a comparison between ICD-10 and DSM-5 criteria for ID level categorization in the context of DSAD.

Taken together, our findings underscore two key points. (1) Whenever possible, obtaining IQ scores before age 35 in individuals with DS is highly valuable for estimating premorbid cognition, ideally integrated as part of a population-based health plan that proactively tracks cognitive development from young adulthood.[32] Future research should investigate AD biomarkers in asymptomatic individuals > 35 years old to determine whether IQ testing remains appropriate. For example, a 43-year-old with normal plasma phosphorylated tau 217 levels might still undergo IQ testing, while biomarker positivity may preclude its use. (2) For individuals already experiencing cognitive decline without prior IQ data, DSM-5–based ID level classification must be used. However, robust, cross-site ID level classification requires the operationalization of DSM-5 criteria through standardized approaches. In the meantime, when using DSM-5–based ID classification, site-specific cognitive cut-offs should be applied to account for cross-site variability, akin to the approach used for cerebrospinal fluid (CSF) core AD biomarkers prior to the adoption of automated platforms.[33]

As an interim solution, future studies may incorporate structured adaptive behavior assessments such as the Adaptive Behavior Assessment System, Third Edition (ABAS-3[34]), which aligns with DSM-5 domains and provides quantifiable, norm-referenced scores to reduce subjectivity. While this improves consistency across sites, it does not resolve how to translate adaptive functioning scores into specific ID levels (e.g., mild vs. moderate) and is not designed for retrospective use. However, it offers a practical and psychometrically robust alternative until a harmonized framework is developed. For multi-site studies, joint case classification or consensus meetings during the preparatory phase are recommended to improve consistency—established practices in AD dementia diagnosis (e.g., in Handen et al.[35]).

Finally, while our study highlights important findings, some limitations should be addressed. First, the demographic composition of our cohort was predominantly White, reflecting the population from which participants were recruited. Further research in racially and ethnically diverse populations is urgently needed to investigate how ID classification may be influenced by these demographic factors. Second, although we made considerable efforts to match cohorts on key variables (e.g., age, sex, cognitive scores), unmeasured differences—such as educational or social factors—may still have influenced the results, given cross-country system variations. Nevertheless, available sociodemographic data (Table S1 and S2) suggest broadly similar cohort compositions.

Although the asymptomatic sample was relatively large, the sample size for individuals with dementia was much lower (N = 110) and Munich participants with IQ scores (N = 36) were small, limiting result interpretation. Additionally, we were unable to include prodromal cases, limiting comparisons across disease stages. This group is particularly important due to the need for early detection of AD symptoms. Individuals with severe and profound ID were also not included due to high CAMCOG-DS non-completion rates, raising the question of whether ID classification discrepancies might also affect these groups. Finally, although this study focused on overall CAMCOG-DS cut-offs, future research should examine domain-level performance per ID level to identify specific constructs that may vary in difficulty or discriminative value within the DS population.

Overall, our research illustrates the caveats surrounding ID level classification and the impact on using international cut-off scores to

classify AD-related cognitive impairment in DS, particularly with a view to multi-site studies. Future research should focus on operationalizing DSM-5 criteria for ID classification through the development of a structured scoring system for conceptual, social, and practical domains with clear thresholds for ID levels, adapted across languages and cultures, and suitable for retrospective use, particularly relevant in the context of DSAD. Importantly, its development should be guided by a thorough examination of cultural variations in adaptive functioning and expectations. In parallel, the field should work toward validating a hybrid ID classification framework that integrates (retrospective) functional assessment with cognitive testing (e.g., premorbid IQ when available), allowing for the generation of a weighted, composite ID severity score. These efforts are essential for consistent, reliable, and globally applicable ID level classification in DSAD research.

ACKNOWLEDGMENTS

The authors thank all the participants with Down syndrome, their families, and caregivers for their support and dedication in this research. The authors also acknowledge Fundació Catalana Síndrome de Down for global support and the members of the Alzheimer Down Unit and the Memory Unit from Hospital de la Santa Creu i Sant Pau and the members of the outpatient clinic for adults with DS, embedded in the Department of Neurology at the University Hospital, LMU Munich for their daily work and dedication. Moreover, we would like to thank Cecilia Mota, Cristina Pastor, and Nina Smrzka for the administrative support. We thank the Center for Advanced Studies (CAS) at Ludwig-Maximilians University Munich (CAS LMU) for supporting the CAS Research Group "Tools for Transnational Neuropsychiatric Research." L.D.H.S., K.S., J.H., and J.F. were part of the group, of which J.L. was the spokesperson. We thank the institutions that funded this study, the Fondo de Investigaciones Sanitario, Carlos III Health Institute, the Centro de Investigación Biomédica en Red sobre Enfermedades Neurodegenerativas and the Generalitat de Catalunya and La Caixa Foundation, the Jérôme Lejeune Foundation as well as the NIH, Horizon 2020 and the Alzheimer's Association. This study was supported by the National Institute on Aging, by the Fondo de Investigaciones Sanitario, Carlos III Health Institute (INT21/00073, PI20/01473 and PI23/01786 to J.F, PI22/00785 to M.C-I, PI22/00307 to AB, through the Miguel Servet program co-funded by the European Union CP24/00112 to L.D.H.S., and CP20/00038 to A.B., through Sara Borrell Postdoctoral Fellowship CD23/00235 to L.V.-A., and through Río-Hortega programm co-funded by the European Union -CM22/00219 to J.E.A.-I., CM22/00052 to I.R.-B., and CM23/00291 to L.M.B) and the Centro de Investigación Biomédica en Red sobre Enfermedades Neurodegenerativas Program 1, partially jointly funded by Fondo Europeo de Desarrollo Regional, Unión Europea Una Manera de Hacer Europa. This work was also supported by the National Institutes of Health grants (R01 AG056850; R21 AG056974, R01 AG061566, R01 AG081394 and R61AG066543 to J.F) and ADNI (U01 AG024904), the Department de Salut de la Generalitat de Catalunya, Pla Estrategic de Recerca I Innovació en Salut (SLT006/17/00119 to J.F). It was also supported by Fundación Tatiana Pérez de Guzmán el Bueno (IIBSP-DOW-2020-151 to J.F.) and Horizon 2020–Research and Innovation Framework Program from the European Union (H2020-SC1-BHC-2018-2020 to J.F.), the Jérôme Lejeune Foundation (2326 - GRT-2024A to L.D.H.S.), and the Alzheimer's Association (AARG-22-923680 to A.B, AARG-22-973966 to M.C-I). Moreover, it was supported by the VERUM Foundation, the Deutsche Forschungsgemeinschaft DFG, German Research Foundation, under Germany's Excellence Strategy within the framework of the Munich Cluster for Systems Neurology (Grant number EXC 2145 SyNergy – ID 390857198) and the T21RS Scientific Exchange Grant (to K.S).

CONFLICT OF INTEREST STATEMENT

J.F. has received personal fees for service on advisory boards, adjudication committees, or speaker honoraria from AC Immune, Adamed, Alzheon, Biogen, Eisai, Esteve, Fujirebio, Ionis, Laboratorios Carnot, Life Molecular Imaging, Lilly, Lundbeck, Perha, and Roche, outside the submitted work. J.F. also holds a patent for markers of synaptopathy in neurodegenerative disease (licensed to Adx, EPI8382175.0). J.L. reports speaker fees from Bayer Vital, Biogen, EISAI, TEVA, Zambon, Esteve, Merck, and Roche; consulting fees from Axon Neuroscience, EISAI, and Biogen; author fees from Thieme Medical Publishers and W. Kohlhammer GmbH medical publishers; and is inventor in a patent "Oral Phenylbutyrate for Treatment of Human 4-Repeat Tauopathies" (PCT/EP2024/053388) filed by LMU Munich. In addition, J.L. reports compensation for serving as chief medical officer for MODAG GmbH, is beneficiary of the phantom share program of MODAG GmbH, and is inventor in a patent "Pharmaceutical Composition and Methods of Use" (EP 22 159 408.8) filed by MODAG GmbH, all activities outside the submitted work. The other authors report no relevant disclosures. Author disclosures are available in the supporting information.

DATA AVAILABILITY STATEMENT

The authors may share de-identified data that underlie the results reported in this article. Data will be available upon receipt of a request detailing the study hypothesis and statistical analysis plan. All requests should be sent to the corresponding authors. The steering committee of this study will discuss all requests and decide, based on the novelty and scientific rigor of the proposal, whether data sharing is appropriate. All applicants will be asked to sign a data access agreement.

CONSENT STATEMENT

The study was approved by the Sant Pau Ethics Committee and by the local ethics committee of the LMU medical faculty, following the standards for medical research in humans recommended by the Declaration of Helsinki. All participants or their legally authorized representatives gave written informed consent before enrolment. All data were anonymized according to good clinical practice guidelines and general data protection regulations prior to analysis.

ORCID

Laura del Hoyo Soriano https://orcid.org/0000-0003-4372-1599
Katja Sandkühler https://orcid.org/0009-0000-7117-2270

REFERENCES

1. Antonarakis SE, Skotko BG, Rafii MS, et al. Down syndrome. *Nat Rev Dis Primers*. 2020;6(1):1-20. doi:10.1038/s41572-019-0143-7
2. Fortea J, Zaman SH, Hartley S, et al. Alzheimer's disease associated with Down syndrome: a genetic form of dementia. *Lancet Neurol*. 2021;20(11):930-942. doi:10.1016/S1474-4422(21)00245-3
3. Fortea J, Vilaplana E, Carmona-Iragui M, et al. Clinical and biomarker changes of Alzheimer's disease in adults with Down syndrome: a cross-sectional study. *The Lancet*. 2020;395(10242):1988-1997. doi:10.1016/S0140-6736(20)30689-9
4. Jack CR, Andrews JS, Beach TG, et al. Revised criteria for diagnosis and staging of Alzheimer's disease: Alzheimer's Association Workgroup. *Alzheimers Dement*. 2024;20(8):5143-5169. doi:10.1002/ALZ.13859
5. Hartley SL, Handen BL, Devenny D, et al. Cognitive indicators of transition to preclinical and prodromal stages of Alzheimer's disease in Down syndrome. *Alzheimers Dement*. 2020;12(1):e12096. doi:10.1002/dad2.12096
6. Startin CM, Hamburg S, Hithersay R, et al. Cognitive markers of preclinical and prodromal Alzheimer's disease in Down syndrome. *Alzheimers Dement*. 2019;15(2):245-257. doi:10.1016/J.JALZ.2018.08.009
7. Del Hoyo L, Xicota L, Sánchez-Benavides G, et al. Semantic verbal fluency pattern, dementia rating scores and adaptive behavior correlate with plasma Aβ42 concentrations in Down syndrome young adults. *Front Behav Neurosci*. 2015;9:301. doi:10.3389/fnbeh.2015.00301
8. Moran JA, Rafii MS, Keller SM, Singh BK, Janicki MP. The National Task Group on intellectual disabilities and dementia practices consensus recommendations for the evaluation and management of dementia in adults with intellectual disabilities. *Mayo Clin Proc*. 2013;88(8):831-840. doi:10.1016/J.MAYOCP.2013.04.024
9. Videla L, Benejam B, Carmona-Iragui M, et al. Cross-sectional versus longitudinal cognitive assessments for the diagnosis of symptomatic Alzheimer's disease in adults with Down syndrome. *Alzheimers Dement*. 2023;19(9):3916-3925. doi:10.1002/ALZ.13073
10. Fleming V, Hom CL, Clare IC, et al. Cognitive outcome measures for tracking Alzheimer's disease in Down syndrome. *Int Rev Res Dev Disabil*. 2022;62:227-263. doi:10.1016/BS.IRRDD.2022.05.006
11. Firth NC, Startin CM, Hithersay R, et al. Aging related cognitive changes associated with Alzheimer's disease in Down syndrome. *Ann Clin Transl Neurol*. 2018;5(6):741-751. doi:10.1002/ACN3.571
12. Nadeau PA, Jobin B, Boller B. Diagnostic sensitivity and specificity of cognitive tests for mild cognitive impairment and Alzheimer's disease in patients with Down syndrome: a systematic review and meta-analysis. *J Alzheimers Dis*. 2023;95(1):13-51. doi:10.3233/JAD-220991
13. Hon J, Huppert FA, Holland AJ, Watson P. Neuropsychological assessment of older adults with Down's syndrome: an epidemiological study using the Cambridge Cognitive Examination (CAMCOG). *Br J Clinl Psychol /Br Psychological Soc*. 1999;38(Pt 2):155-165. http://www.ncbi.nlm.nih.gov/pubmed/10389597
14. Benejam B, Videla L, Vilaplana E, et al. Diagnosis of prodromal and Alzheimer's disease dementia in adults with Down syndrome using neuropsychological tests. *Alzheimers Dement*. 2020;12(1):e12047. doi:10.1002/dad2.12047
15. Pudumjee SB, Lundt ES, Albertson SM, et al. A comparison of cross-sectional and longitudinal methods of defining objective subtle cognitive decline in preclinical Alzheimer's disease based on cogstate one card learning accuracy performance. *J Alzheimers Dis*. 2021;83(2):861-877. doi:10.3233/JAD-210251
16. World Health Organization. *ICD-10 : International Statistical Classification Of Diseases and Related Health Problems : Tenth Revision*. 2nd ed. World Health Organization; 2004. https://iris.who.int/handle/10665/42980
17. Ghezzo A, Salvioli S, Solimando MC, et al. Age-related changes of adaptive and neuropsychological features in persons with Down syndrome. *PLoS One*. 2014;9(11):e113111. doi:10.1371/JOURNAL.PONE.0113111
18. Das JP, Divis B, Alexander J, Parrila RK, Naglieri JA. Cognitive decline due to aging among persons with Down syndrome. *Res Dev Disabil*. 1995;16(6):461-478. doi:10.1016/0891-4222(95)00030-5
19. Hamburg S, Lowe B, Startin CM, et al. Assessing general cognitive and adaptive abilities in adults with Down syndrome: a systematic review. *J Neurodev Disord*. 2019;11(1):20. doi:10.1186/S11689-019-9279-8
20. First MB. *DSM-5® Handbook of Differential Diagnosis*. American Psychiatric Publishing; 2013. doi:10.1176/appi.books.9781585629992
21. Videla L, Benejam B, Pegueroles J, et al. Longitudinal clinical and cognitive changes along the Alzheimer Disease Continuum in Down syndrome. *JAMA Netw Open*. 2022;5(8):e2225573. doi:10.1001/JAMANETWORKOPEN.2022.25573
22. Kaufman AS, Kaufman AL. *KBIT: Test Breve de Inteligencia de Kaufman. Manual de interpretación. [KBIT: Kaufman's Brief Intelligence Test. Interpretation Manual]*. TEA; 1994.
23. Raven JC. *CPM—Coloured Progressive Matrices: Testhandbuch zur deutschen Version (Deutsche Bearbeitung: A. Bulheller & H. Häcker)*. Pearson Assessment; 2009.
24. Raven J, Rust J, Chan F, and Zhou X. *Raven's 2 Progressive matrices, clinical edition (raven's 2)* Pearson; 2018.
25. Esteba Castillo S, Dalmau Bueno A, Ribas Vidal N, Vilà Alsina M, Novell Alsina R, García Alba J. Adaptation and validation of CAMDEX-DS (Cambridge Examination for Mental Disorders of Older People with Down's Syndrome and Others with Intellectual Disabilities) in Spanish population with intellectual disabilities. *Rev Neurol*. 2013;57(8):337-346. 10.33588/RN.5708.2013259
26. Nübling G, Loosli SV, Wlasich E, et al. A German version of the Cambridge examination for mental disorders of older people with Down's syndrome and others with intellectual disabilities: a diagnostic procedure for detecting dementia in people with Down's syndrome. *Z Gerontol Geriatr*. 2020;53(6):546-551. doi:10.1007/S00391-019-01591-7/METRICS
27. R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing; 2020 Accessed: May 10, 2021. [Online]. Available: https://www.r-project.org/
28. Sawilowsky SS. New effect size rules of thumb. *J Modern Appl Stat Meth*. 2009;8:597-599. 10.56801/10.56801/V8.I.452
29. Ho DE, Imai K, King G, Stuart EA. MatchIt: nonparametric preprocessing for parametric causal inference. *J Stat Softw*. 2011;42(8):1-28. 10.18637/JSS.V042.I08
30. Unal I. Defining an optimal cut-point value in ROC analysis: an alternative approach. *Comput Math Methods Med*. 2017;2017:3762651. doi:10.1155/2017/3762651
31. Kievit RA, Frankenhuis WE, Waldorp LJ, Borsboom D. Simpson's paradox in psychological science: a practical guide. *Front Psychol*. 2013;4:513. doi:10.3389/FPSYG.2013.00513
32. Baksh RA, Pape SE, Chan LF, Aslam AA, Gulliford MC, Strydom A. Multiple morbidity across the lifespan in people with Down syndrome or intellectual disabilities: a population-based cohort study using electronic health records. *Lancet Public Health*. 2023;8(6):e453-e462. doi:10.1016/S2468-2667(23)00057-9
33. Alcolea D, Pegueroles J, Muñoz L, et al. Agreement of amyloid PET and CSF biomarkers for Alzheimer's disease on Lumipulse. *Ann Clin Transl Neurol*. 2019;6(9):1815. doi:10.1002/ACN3.50873
34. Harrison P, Oakland T. *Adaptive Behavior Assessment System (ABAS-II), Western Ps*. PEARSON; 2003.

35. Handen BL, Lott IT, Christian BT, et al. The Alzheimer's Biomarker Consortium-Down syndrome: rationale and methodology. *Alzheimers Dement*. 2020;12(1):e12065. doi:10.1002/DAD2.12065⟨?PMU?⟩

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Soriano LH, Sandkühler K, Videla L, et al. Discrepancies in assessing intellectual disability levels in adults with Down syndrome: Implications for dementia diagnosis. *Alzheimer's Dement*. 2025;21:e70307. https://doi.org/10.1002/alz.70307