Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Normative scoring and discordance between performance-based and patient-reported cognition in young adults with congenital heart disease: a MINDS-ACHD cross-sectional analysis

Dong Wei
Children's Hospital of Nanjing Medical University
Yaqi Zhang
Children's Hospital of Nanjing Medical University
Zheng Wei
The Affiliated Hospital of Yanbian University, Korean Autonomous Prefecture
Zhaocong Yang
Children's Hospital of Nanjing Medical University
Xuming Mo
mohsuming15@njmu.edu.cn

Children's Hospital of Nanjing Medical University

Research Article

Keywords: adult congenital heart disease, cognition, patient-reported outcomes, NIH Toolbox, normative scores, discordance

Posted Date: November 25th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11419785/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

Background

Performance tests and patient reports assess different aspects of cognition in congenital heart disease (CHD). Their overlap within individuals, and its sensitivity to the normative reference used for performance scores, are less clear. We examined these questions in young adults with CHD.

Methods

This cross-sectional secondary analysis included 403 adults aged 18–30 years with moderate or severely complex CHD from MINDS-ACHD. NIH Toolbox fully corrected and Neuro-QoL Cognitive Function T scores below 40 defined four paired categories. Age-corrected and uncorrected standard scores below 85 were compared within the same participants. Site-stratified bootstrap resampling provided 95% confidence intervals (CIs). Sequential norm comparisons and Hospital Anxiety and Depression Scale profiles were examined descriptively.

Results

With fully corrected scoring, 136 participants had low performance scores and 88 had low patient-report scores; 32 had both. Performance-only classification exceeded patient-report-only classification by 11.91 percentage points (95% CI, 5.96–17.87). Moving to uncorrected scoring reclassified 106 participants (26.3%), reducing performance-low classifications from 136 to 30 and both-low classifications from 32 to 8. Only uncorrected scoring reversed the predominant discordant category. Sequential comparisons showed 37 changes from fully corrected to age-corrected scoring and 87 from age-corrected to uncorrected scoring, including nine participants who changed at both steps. Under the primary classification, patient-report-only participants had higher mean anxiety (11.93 versus 5.52) and depression (6.57 versus 2.77) scores than performance-only participants.

Conclusions

Performance-based and patient-reported low scores identified substantially different groups of young adults with CHD, and the normative reference altered their overlap. The higher mood-symptom burden in the patient-report-only group supports considering anxiety and depression when interpreting reported cognitive difficulties in adult CHD follow-up.

Background

Young adults with congenital heart disease (CHD) must manage a lifelong cardiac condition while pursuing education, employment, and increasing independence. Cognitive and psychological difficulties

can complicate these demands, even when cardiovascular status is stable. Neurodevelopmental vulnerabilities may persist from childhood, and adult congenital heart disease (ACHD) reviews and scientific statements emphasize assessment across the life course [1–4]. Recent ACHD consensus statements and clinical guidance place neurocognitive and psychological care within comprehensive follow-up [5,6]. Identifying who may need further evaluation therefore requires attention to how cognitive difficulties are measured.

Standardized performance tests and patient reports assess different aspects of cognition. A performance task samples ability under structured conditions; patient report describes perceived effectiveness in daily activities. A recent umbrella review found weak average correspondence between the two approaches across medical and neuropsychiatric populations [7]. Smaller ACHD studies have also examined performance testing alongside perceived cognitive difficulties and psychosocial outcomes [8,9]. Separate percentages of low scores cannot establish whether the same people are identified. A paired classification makes this distinction visible: low scores on both measures, on performance testing alone, on patient report alone, or on neither.

The Multi-Institutional Neurocognitive Discovery Study in Adult Congenital Heart Disease (MINDS-ACHD) collected paired measurements at 14 North American centers [10]. Its main report described the separate prevalence and predictors of performance-based and patient-reported dysfunction [11,12]. We extend that work by examining within-person overlap and reclassification. NIH Toolbox fully corrected, age-corrected, and uncorrected scores apply different normative references to the same responses [13,14]. We compared their effects on paired cognitive classification in a fixed public-use sample. This measurement-focused analysis informs comparisons across ACHD studies; mood and school or work participation provide secondary descriptive context.

Methods

Study design and data source

This secondary cross-sectional analysis used the MINDS-ACHD public-use dataset. The parent study enrolled adults aged 18 to 30 years with moderate or severely complex CHD at 14 North American centers. Enrollment ran from November 2020 through March 2023, and study procedures were completed by May 2023 [10,11]. Participants had to complete the assessments independently; parent-study exclusions included current hospitalization, recent procedures requiring sedation or anesthesia, and conditions preventing independent testing. Known genetic syndromes were not an automatic exclusion.

We used the deidentified public-use archive from the Pediatric Heart Network [15]. Demographic, cognitive-score, Hospital Anxiety and Depression Scale (HADS), and medical-history files were linked using blinded participant identifiers. The official codebooks and National Institutes of Health (NIH) Toolbox scoring guide defined the score fields and clinical variables. Reporting follows the Strengthening

the Reporting of Observational Studies in Epidemiology (STROBE) recommendations for cross-sectional studies [16]. Supplementary results are provided in Additional file 1.

Analytic cohort

The completed-study demographics file contained 467 unique participants. The primary paired cohort required a numeric NIH Toolbox Cognition Total Composite Fully Corrected T score and a Neuro-QoL Cognitive Function T score. We checked participant uniqueness, instrument and score type, score ranges, and membership in the completed-study cohort. The primary analysis used observed score pairs without imputation.

We compared the paired cohort with completed-study participants lacking a complete primary pair using signed standardized differences. For continuous variables, the paired-minus-nonpaired mean difference was divided by the square root of the average group variances; for binary indicators, the proportion difference was divided by the square root of the average binomial variances. Available denominators were retained for each characteristic. We also tabulated joint availability of all three performance scores and patient report, and checked raw instrument records for the site contributing no primary pairs.

Cognitive measures and normative scores

Performance-based cognition was measured with the NIH Toolbox Cognition Total Composite, labelled version 1.1 in the public-use release. The battery samples episodic memory, executive function, attention, working memory, language, and processing speed; adult validation and composite reliability have been reported [17–19]. The primary Fully Corrected T score has a normative mean of 50 and standard deviation (SD) of 10. Its normative correction accounts for age, education, sex, and race/ethnicity [13]. We selected it to describe performance relative to demographically comparable individuals and to align with the MINDS-ACHD main report. The Age-Corrected Standard Score adjusts for age alone. The Uncorrected Standard Score uses the general normative distribution without those demographic corrections. Both standard-score scales have a normative mean of 100 and SD of 15 and served as explanatory comparisons. Higher scores indicate better performance.

Patient-reported cognitive function was measured with the Neuro-QoL Bank version 2.0 Cognitive Function T score. The item bank assesses perceived difficulties in memory, attention, decision-making, and everyday cognitive tasks. Scores have a normative mean of 50 and SD of 10; higher scores indicate better perceived function [10,20]. The primary thresholds were Fully Corrected T score below 40 and Neuro-QoL T score below 40, more than 1 SD below each normative mean. The equivalent threshold on either NIH Toolbox standard-score scale was below 85. These research categories describe low scores, not a clinical diagnosis of cognitive impairment. Scores exactly at the threshold were classified as not low.

Results

Analytic cohort

Among 467 participants who completed study procedures, 403 (86.3%) had both the Fully Corrected NIH Toolbox Cognition Total Composite T score and Neuro-QoL Cognitive Function T score, forming the primary paired cohort across 13 blinded sites (Fig. 1). The paired cohort had a mean age of 24.55 years (SD 3.40), 55.2% were female, and the four CHD groups were represented in proportions similar to the completed cohort (Table 1). Sixty-four participants lacked a complete primary pair; score-availability patterns and site distributions are detailed in Supplementary Table S1.

Table 1
Characteristics of completed, paired, and nonpaired participants

Characteristic	Completed cohort	Primary paired	No complete pair	Standardized difference
Age, mean (SD), y	24.54 (3.41)	24.55 (3.40)	24.46 (3.46)	0.03
Female, n/N (%)	262/466 (56.2)	222/402 (55.2)	40/64 (62.5)	-0.15
White, n (%)	376 (80.5)	327 (81.1)	49 (76.6)	0.11
Black or African American, n (%)	29 (6.2)	24 (6.0)	5 (7.8)	-0.07
Asian, n (%)	25 (5.4)	20 (5.0)	5 (7.8)	-0.12
Other/multiracial, n (%)	36 (7.7)	32 (7.9)	4 (6.3)	0.07
Hispanic or Latino ethnicity, n (%)	49 (10.5)	41 (10.2)	8 (12.5)	-0.07
High school or less, n (%)	144 (30.8)	124 (30.8)	20 (31.3)	-0.01
Some college/associate degree, n (%)	139 (29.8)	120 (29.8)	19 (29.7)	0.00
College/postgraduate degree, n (%)	184 (39.4)	159 (39.5)	25 (39.1)	0.01
d-Transposition of the great arteries, n (%)	82 (17.6)	70 (17.4)	12 (18.8)	-0.04
Tetralogy of Fallot, n (%)	123 (26.3)	106 (26.3)	17 (26.6)	-0.01
Single ventricle, n (%)	132 (28.3)	112 (27.8)	20 (31.3)	-0.08
Other moderate/severely complex CHD, n (%)	130 (27.8)	115 (28.5)	15 (23.4)	0.12
Open-heart procedures, median (IQR)	2 (1–3)	2 (1–3)	2 (1–3)	-0.13
Cerebrovascular history, n (%)	24 (5.1)	22 (5.5)	2 (3.1)	0.12

Completed cohort, n = 467; primary paired cohort, n = 403; no complete pair, n = 64. Percentages use available denominators; not-provided race and ethnicity are retained in their denominators. The signed standardized difference is primary paired minus no complete pair, using means for continuous variables (including open-heart procedure count) and proportions for category indicators, as defined in Methods. Race and ethnicity are separate released variables. Full race, site, and score-availability distributions are in Supplementary Table S1. CHD, congenital heart disease; IQR, interquartile range; SD, standard deviation.

Paired overlap under fully corrected scoring

Fully corrected scoring classified 211 participants (52.36%) as low on neither measure, 104 (25.81%) as performance-only low, 56 (13.90%) as patient-report-only low, and 32 (7.94%) as low on both (Table 2). Of the 136 participants with low performance scores, 32 (23.5%) also had low patient-report scores; among the 88 with low patient-report scores, 32 (36.4%) also had low performance scores. Total discordance was 39.70% (95% CI, 34.99–44.42). Performance-only classification exceeded patient-report-only classification by 11.91 percentage points (95% CI, 5.96–17.87; exact McNemar P = 0.00018).

Table 2
Paired cognitive classifications under three normative references in the same 403 participants

Category or metric	**Fully corrected T < 40**	**Age-corrected SS < 85**	**Uncorrected SS < 85**
Neither low	211 (52.36) [47.64, 57.07]	231 (57.32) [52.61, 62.03]	293 (72.70) [68.24, 76.92]
Performance-only low	104 (25.81) [21.84, 30.02]	84 (20.84) [17.12, 24.81]	22 (5.46) [3.23, 7.69]
Patient-report-only low	56 (13.90) [10.42, 17.37]	55 (13.65) [10.42, 17.12]	80 (19.85) [16.13, 23.82]
Both low	32 (7.94) [5.46, 10.67]	33 (8.19) [5.71, 10.92]	8 (1.99) [0.74, 3.47]
Performance low, n (%)	136 (33.75) [29.28, 38.46]	117 (29.03) [24.81, 33.50]	30 (7.44) [4.96, 9.93]
Total discordance, %	39.70 [34.99, 44.42]	34.49 [30.02, 38.96]	25.31 [21.09, 29.53]
Performance-only minus patient-report-only, pp	11.91 [5.96, 17.87]	7.20 [1.49, 12.90]	-14.39 [-19.11, -9.68]

Category cells show n (%) with 95% CIs for percentages in square brackets. The fixed denominator is 403; patient-report low is Neuro-QoL T < 40 under every norm. Fully corrected scoring is primary; age-corrected and uncorrected scoring are explanatory comparisons. Performance-low counts include performance-only and both-low participants. CIs use 5,000 site-stratified participant bootstrap replicates. The exact two-sided McNemar test under fully corrected scoring gave P = 0.00018. Agreement coefficients and independent-marginal benchmarks are in Supplementary Table S9. pp, percentage points; CI, confidence interval.

Raw agreement was 60.30%, with Cohen's kappa of 0.028 (95% CI, -0.065 to 0.121). Observed discordance was close to the independent-marginal benchmark of 40.85%; the difference was − 1.14 percentage points (95% CI, -5.00 to 2.62). Continuous scores were weakly associated: Pearson r = 0.042

(95% CI, -0.058 to 0.142) and Spearman ρ = 0.070 (95% CI, -0.026 to 0.170) (Fig. 2). Additional agreement estimates, including Gwet's AC1, are reported in Supplementary Table S9.

Norm-dependent classification

Changing the performance-score norm changed both the frequency and direction of discordance in the same 403 participants (Table 2; Fig. 3). With age-corrected scoring, 34.49% (95% CI, 30.02–38.96) were discordant and performance-only classification remained more common by 7.20 percentage points (95% CI, 1.49–12.90). With uncorrected scoring, discordance was 25.31% (95% CI, 21.09–29.53), and patient-report-only classification became more common; the performance-only minus patient-report-only difference was − 14.39 percentage points (95% CI, -19.11 to -9.68).

Relative to fully corrected scoring, the change in total discordance was − 5.21 percentage points (95% CI, -7.94 to -2.48) under age-corrected scoring and − 14.39 percentage points (95% CI, -19.11 to -9.68) under uncorrected scoring. Direct paired contrasts and the definitions of these changes are provided in Supplementary Table S7.

Moving from fully corrected to uncorrected scoring reduced performance-low classifications from 136 (33.75%) to 30 (7.44%). All 106 reclassified participants (26.3%) moved from low to not low: 82 from performance-only to neither-low and 24 from both-low to patient-report-only. The both-low group consequently decreased from 32 to 8 (Fig. 3; Supplementary Table S2). The sequential comparisons differed. Fully corrected to age-corrected scoring changed 37 classifications: 28 from low to not low and nine in the reverse direction. Of those 28, 22 (78.6%) had a college or postgraduate degree. Among fully corrected low scorers, low-to-not-low transition rates were 22/36 (61.1%) in college/postgraduate graduates, 4/40 (10.0%) in those with some college/associate education, and 2/60 (3.3%) in those with high school or less. Age-corrected to uncorrected scoring moved 87 participants from low to not low, including 14 (16.1%) college/postgraduate graduates. Nine participants changed at both steps and returned to their original classification (Supplementary Table S10).

Clinical profiles of discordant categories

Patient-report-only participants had higher HADS scores than performance-only participants (Table 3; Fig. 4). Mean anxiety was 5.52 (SD 3.57) in performance-only participants and 11.93 (SD 4.37) in patient-report-only participants; corresponding depression scores were 2.77 (SD 2.89) and 6.57 (SD 3.81). The performance-only minus patient-report-only differences were − 6.41 points (95% CI, -7.73 to -5.13) for anxiety and − 3.80 points (95% CI, -4.94 to -2.69) for depression. Anxiety scores ≥ 11 occurred in 34/56 (60.7%) patient-report-only participants and 12/104 (11.5%) performance-only participants. In the patient-report-only group, the fully corrected performance T score averaged 52.45 (SD 7.60); 49/56 participants (87.5%) scored at least 45, placing most at least five points above the low-score boundary. Both-low participants also had higher mean HADS scores than neither-low participants.

Table 3
Cognitive scores, mood symptoms, and participation by primary paired category

Characteristic	Neither low (n = 211)	Performance-only (n = 104)	Patient-report-only (n = 56)	Both low (n = 32)
Performance T score, mean (SD)	51.51 (7.72)	31.57 (6.16)	52.45 (7.60)	29.50 (6.55)
Performance T score, median (IQR)	51.00 (45.00–57.00)	33.50 (27.75–37.00)	50.50 (46.75–58.25)	31.00 (24.00–34.25)
Neuro-QoL T score, mean (SD)	47.13 (4.69)	47.08 (5.76)	35.45 (3.51)	37.63 (1.91)
Neuro-QoL T score, median (IQR)	46.10 (43.50–49.90)	45.90 (42.70–49.90)	36.30 (33.25–38.23)	38.20 (36.60–39.50)
HADS anxiety, mean (SD)	6.84 (4.21)	5.52 (3.57)	11.93 (4.37)	11.38 (4.49)
HADS anxiety, median (IQR)	7.00 (3.00–10.00)	5.00 (3.00–7.25)	12.00 (9.00–15.00)	12.00 (8.75–15.00)
HADS anxiety ≥ 8, n (%)	92 (43.6)	26 (25.0)	47 (83.9)	25 (78.1)
HADS anxiety ≥ 11, n (%)	39 (18.5)	12 (11.5)	34 (60.7)	19 (59.4)
HADS depression, mean (SD)	2.40 (2.42)	2.77 (2.89)	6.57 (3.81)	5.84 (3.42)
HADS depression, median (IQR)	2.00 (1.00–3.00)	2.00 (1.00–4.00)	6.00 (4.00–9.00)	5.00 (3.75–7.25)
HADS depression ≥ 8, n (%)	9 (4.3)	7 (6.7)	18 (32.1)	8 (25.0)
HADS depression ≥ 11, n (%)	2 (0.9)	2 (1.9)	10 (17.9)	3 (9.4)
Current participation, n/N (%)	193/211 (91.5)	89/103 (86.4)	50/56 (89.3)	23/32 (71.9)

Characteristic	Neither low (n = 211)	Performance-only (n = 104)	Patient-report-only (n = 56)	Both low (n = 32)
Categories use fully corrected performance T < 40 and Neuro-QoL T < 40. HADS, Hospital Anxiety and Depression Scale; IQR, interquartile range; SD, standard deviation. Scores ≥ 8 and ≥ 11 are screening thresholds; the ≥ 11 group is a subset of the ≥ 8 group. Participation was missing for one performance-only participant; all listed cognitive and HADS scores were complete. Performance-only minus patient-report-only differences were − 6.41 (95% CI, -7.73 to -5.13) for mean anxiety, -3.80 (95% CI, -4.94 to -2.69) for mean depression, and − 2.88 percentage points (95% CI, -13.01 to 7.84) for participation. Median differences were − 7 (95% CI, -9 to -5) for anxiety and − 4 (95% CI, -6 to -3) for depression. Demographic and clinical profiles are in Supplementary Table S10.				
Counts were verified in the MINDS-ACHD public-use files. Primary pairing required the NIH Toolbox Cognition Total Composite Fully Corrected T score and Neuro-QoL Cognitive Function T score. The paired cohort included 403 participants (86.3% of the completed cohort) from 13 sites. Of 64 participants without a complete primary pair, 50 had only the patient-report score and 14 had neither primary score.				
Points retain exact observed coordinates. Point area is proportional to the number of participants sharing a coordinate; 369 distinct coordinates represent 403 participants. Dashed lines mark T scores of 40; low scores are strictly below 40. Color, shape, and quadrant labels identify the categories. Percentages use the denominator of 403. Pearson r = 0.042 (95% CI, -0.058 to 0.142).				
Supplementary Material				
Normative scoring and discordance between performance-based and patient-reported cognition in young adults with congenital heart disease: a MINDS-ACHD cross-sectional analysis				
Supplementary Table S1 Cognitive score availability and cohort characteristics				
Panel A. Blinded site distribution				

Current school or paid-work participation was reported by 89 of 103 participants (86.4%) in the performance-only group and 50 of 56 (89.3%) in the patient-report-only group. The prevalence difference was − 2.88 percentage points (95% CI, -13.01 to 7.84). The both-low group had the lowest observed participation proportion (23/32, 71.9%). CHD group and cerebrovascular history distributions are shown in Supplementary Table S10.

In models adjusted for age, gender, and CHD group, the performance-only minus patient-report-only mean difference was − 6.53 HADS anxiety points (95% CI, -7.86 to -5.24) and − 3.86 depression points (95% CI, -5.02 to -2.74). Estimates were similar after additional adjustment for center and education. All models used the same 160 participants. Full intervals and model stability summaries are in Supplementary Table S8.

Sensitivity analyses

The primary discordance estimate ranged from 38.24% to 40.16% after excluding participants with recorded cerebrovascular history, entries in the released genetic-condition fields, or both (Supplementary Table S4). Leave-one-site-out estimates ranged from 37.57% to 41.21%, and the direction difference

remained positive after every omission (9.39–14.45 percentage points; Supplementary Figure S1). Available-case analyses of 406 participants closely matched the fixed-cohort estimates: discordance was 34.73% for age-corrected scores and 25.37% for uncorrected scores (Supplementary Table S5). More stringent thresholds produced sparse cells (Supplementary Table S3).

The direction reversal persisted when scores equal to the 1-SD boundary were included as low (Supplementary Table S6).

Discussion

Performance-based and patient-reported low scores showed limited overlap in this young-adult CHD cohort. Under fully corrected scoring, fewer than one quarter of participants with low performance scores also had low patient-report scores. Changing to uncorrected scoring reclassified 26.3% of participants and reduced the both-low group from 32 to 8. Uncorrected, but not age-corrected, scoring shifted the predominant discordant category toward patient-report-only. In the primary fully corrected classification, the patient-report-only group also reported higher anxiety and depression scores than the performance-only group.

The within-person design holds participants and patient-report scores constant across norm comparisons. The three norms compare performance with different reference groups. Fully corrected scores account for age and selected demographic characteristics, age-corrected scores condition on age alone, and uncorrected scores use a general normative comparison [13]. Parallel 1-SD cutoffs therefore describe different relative standings. Age-corrected scoring is particularly relevant to comparisons with same-age peers in this young-adult cohort. Nitsch and colleagues demonstrated changes in NIH Toolbox interpretation after demographic correction in acquired brain injury [14]. Our analysis shows how this choice also changes overlap with patient reports in CHD. Subsequent NIH Toolbox Version 3 norming and executive-task scoring studies further illustrate why the test and scoring version should be identified [24,25].

Education is particularly relevant in CHD because attainment may reflect learning difficulties, school absence, and access to support across development [1,3]. Conditioning on attainment describes performance relative to people with similar education; it can also remove variation relevant to the overall developmental burden. The fully corrected-to-age-corrected comparison changes education, sex, and race/ethnicity correction together. Among fully corrected low scorers, 61.1% of college/postgraduate graduates became not low with age-corrected scoring, compared with 10.0% of those with some college/associate education and 3.3% of those with high school or less. By contrast, most of the 87 low-to-not-low changes between age-corrected and uncorrected scoring occurred outside the college/postgraduate group. The endpoint comparison of 106 changes combines distinct education-related transition patterns across the normative references.

The weak continuous-score correlations are consistent with the broader performance/report literature and smaller ACHD studies [7–9]. Observed discordance (39.70%) was close to the independent-marginal

benchmark (40.85%), so its absolute percentage alone adds little evidence about correspondence. Under uncorrected scoring, raw agreement and AC1 increased as the neither-low category grew and low-score margins changed. AC1 also uses a different chance-agreement term from kappa [21]. The paired categories and individual reclassification paths are therefore more informative than ranking norms by a single agreement coefficient (Supplementary Table S9).

Uncorrected scoring moved 82 participants from performance-only to neither-low and 24 from both-low to patient-report-only. These transitions reduced the number of discordant participants by 58 while reducing the both-low group from 32 to 8. The direction reversal persisted with boundary-inclusive scoring; sparse cells limited assessment at 1.5- and 2-SD cutoffs.

Patient-report-only participants had higher anxiety and depression scores in both unadjusted and adjusted comparisons; 60.7% reached the HADS anxiety threshold of 11, compared with 11.5% of performance-only participants. Their performance scores were generally well above the low-score boundary. This pattern is consistent with findings in MINDS-ACHD, stroke, and broader ACHD research [11,26–28], and supports attending to mood symptoms when interpreting reported cognitive difficulties. Neuro-QoL and HADS are both self-reports, so negative affect and shared reporting processes could contribute to the association. Cross-sectional data cannot distinguish mood-related cognitive perceptions from cognitive difficulties accompanying mood symptoms. Interviews, informant reports, or functional tasks could clarify these profiles.

School or paid-work participation was 86.4% and 89.3% in the two discordant groups, with an imprecise between-group difference. The both-low group had the lowest observed participation (71.9%). This binary measure captures participation but not work hours, accommodations, or difficulties sustaining a role. Functional assessments could determine whether the cognitive categories differ in these aspects of everyday life.

In ACHD follow-up, a performance test and a patient report should not be treated as interchangeable ways to identify low-scoring patients. ACHD-specific health-status research likewise distinguishes patient report from objective clinical status [29]. Our findings distinguish tested performance, perceived cognitive difficulty, and accompanying mood symptoms within the assessment priorities of ACHD guidance [5,6,30]. For comparisons across studies, the normative reference, score version, cutoff, and paired denominator determine how the low-score categories can be interpreted.

Limitations

The cohort was restricted to 18- to 30-year-olds with moderate or severely complex CHD who could complete testing independently. Sixty-four completed participants lacked fully corrected scores, including 13 at one site with no released performance records, creating potential selection bias (Supplementary Table S1). Composite scores do not separate the contributions of age, education, sex, and race/ethnicity to reclassification. The study compared classifications rather than validating them

against an independent clinical or functional criterion; longitudinal data would be needed to assess their prognostic value.

Conclusions

Performance-based and patient-reported low scores identified substantially different groups of young adults with moderate or severely complex CHD. Moving from fully corrected to uncorrected scoring reclassified 106 of 403 participants (26.3%) and reduced the both-low group from 32 to 8. Sequential comparisons revealed different education-related transition patterns, which the endpoint comparison alone obscured. The higher anxiety and depression scores in the patient-report-only group support considering mood symptoms alongside reported cognitive difficulties in ACHD follow-up.

Abbreviations

ACHD
adult congenital heart disease
CHD
congenital heart disease
CI
confidence interval
HADS
Hospital Anxiety and Depression Scale
IQR
interquartile range
MINDS-ACHD
Multi-Institutional Neurocognitive Discovery Study in Adult Congenital Heart Disease
NIH
National Institutes of Health
Neuro-QoL
Quality of Life in Neurological Disorders
SD
standard deviation
STROBE
Strengthening the Reporting of Observational Studies in Epidemiology.

Declarations

Ethics approval and consent to participate

This secondary analysis used publicly available, deidentified MINDS-ACHD records. No additional institutional review board approval or informed consent was required. Ethics approval and participant

consent for the original study were obtained by the MINDS-ACHD investigators [10,11].

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Funding

This work was supported by the National Natural Science Foundation of China (82270310) and the Jiangsu Provincial Social Development Program (BE2023662). The funders had no role in study design, data analysis, manuscript preparation, or the decision to publish.

Author Contribution

DW conceived the study, extracted and curated the data, performed the statistical analyses, drafted the manuscript, and prepared the figures and tables. YZ contributed to study conception and statistical analyses, prepared the figures and tables, and revised the initial draft. ZW contributed to study conception and statistical analyses, verified the data, and revised the figures and tables. ZY extracted the data, verified the results, and revised the manuscript. XM supervised the study, oversaw the methodology, coordinated the project, and reviewed and revised the manuscript.

Acknowledgements

The NIH/NHLBI Pediatric Heart Network MINDS-ACHD dataset was used in preparation of this work. Data were downloaded from www.pediatricheartnetwork.orgon09/27/2026 [15].

Data Availability

This secondary analysis used publicly available, deidentified MINDS-ACHD records. No additional institutional review board approval or informed consent was required. Ethics approval and participant consent for the original study were obtained by the MINDS-ACHD investigators.

References

1. 1. Sood E, Newburger JW, Anixt JS, Cassidy AR, Jackson JL, Jonas RA, et al. Neurodevelopmental Outcomes for Individuals With Congenital Heart Disease: Updates in Neuroprotection, Risk-Stratification, Evaluation, and Management: A Scientific Statement From the American Heart Association. Circulation. 2024;149(13):e997-e1022. doi:10.1161/cir.0000000000001211.
2. 2. Wolfe K, Peyvandi S. Neurodevelopmental outcomes in congenital heart disease: modifiable and nonmodifiable substrates. Curr Opin Cardiol. 2025;40(4):259–264. doi:10.1097/hco.0000000000001218.
3. 3. Kovacs AH, Bellinger DC. Neurocognitive and psychosocial outcomes in adult congenital heart disease: a lifespan approach. Heart. 2021;107(2):159–167. doi:10.1136/heartjnl-2016-310862.
4. 4. Keir M, Ebert P, Kovacs AH, Smith JMC, Kwan E, Field TS, et al. Neurocognition in Adult Congenital Heart Disease: How to Monitor and Prevent Progressive Decline. Can J Cardiol. 2019;35(12):1675–1685. doi:10.1016/j.cjca.2019.06.020.
5. 5. Jackson JL, Wolfe KR, Gaynor JW, Suarez L, Umfleet L, Puder R, et al. Prioritizing Neurocognitive and Psychosocial Outcomes in Adults With Congenital Heart Disease: A Consensus Statement From the Adults With Congenital Heart Disease/Neuro 2024 Conference. J Am Heart Assoc. 2026;15(2):e042672. doi:10.1161/jaha.125.042672.
6. 6. Gurvitz M, Krieger EV, Fuller S, Davis LL, Kittleson MM, Aboulhosn JA, et al. 2025 ACC/AHA/HRS/ISACHD/SCAI Guideline for the Management of Adults With Congenital Heart Disease: A Report of the American College of Cardiology/American Heart Association Joint Committee on Clinical Practice Guidelines. J Am Coll Cardiol. 2026;87(7):822–976. doi:10.1016/j.jacc.2025.09.006.
7. 7. Van Patten R, Mulhauser K, Austin TA, Bellone JA, Cotton E, Chan L, et al. The association between subjective and objective cognitive functioning from a transdiagnostic perspective: An umbrella review and meta-analysis. Clin Psychol Rev. 2025;121:102648. doi:10.1016/j.cpr.2025.102648.
8. 8. Verrall CE, Tran DL, Kasparian NA, Williams T, Oxenham V, Ayer J, et al. Cognitive Functioning and Psychosocial Outcomes in Adults with Complex Congenital Heart Disease: A Cross-sectional Pilot Study. Pediatr Cardiol. 2024;45(3):529–543. doi:10.1007/s00246-023-03376-7.
9. 9. Brunmeier A, Reis MP, Earing MG, Umfleet L, Ginde S, Bartz PJ, et al. Identifying self-reported neurocognitive deficits in the adult with congenital heart disease using a simple screening tool. Congenit Heart Dis. 2018;13(5):728–733. doi:10.1111/chd.12646.
10. 10. Cohen S, Gurvitz M, Burns KM, Wheaton O, Panigrahy A, Umfleet L, et al. Design of a multi-institutional neurocognitive discovery study in adult congenital heart disease (MINDS-ACHD). Am Heart J. 2023;262:131–139. doi:10.1016/j.ahj.2023.04.002.
11. 11. Cohen S, Gurvitz M, Burns KM, Wheaton O, Umfleet LG, Brown N, et al. Prevalence and Predictors of Neurocognitive Dysfunction in Adults With Congenital Heart Disease: Results From MINDS-ACHD. J Am Coll Cardiol. 2025;86(9):640–655. doi:10.1016/j.jacc.2025.06.051.

12. 12. Peyvandi S, Williams RG. Neurocognitive Outcomes in Adults With Congenital Heart Disease: MIND the Gap. J Am Coll Cardiol. 2025;86(9):656–658. doi:10.1016/j.jacc.2025.07.005.
13. 13. Casaletto KB, Umlauf A, Beaumont J, Gershon R, Slotkin J, Akshoomoff N, et al. Demographically Corrected Normative Standards for the English Version of the NIH Toolbox Cognition Battery. J Int Neuropsychol Soc. 2015;21(5):378–391. doi:10.1017/s1355617715000351.
14. 14. Nitsch KP, Casaletto KB, Carlozzi NE, Tulsky DS, Heinemann AW, Heaton RK. Uncorrected versus demographically-corrected scores on the NIH Toolbox Cognition Battery in persons with traumatic brain injury and stroke. Rehabil Psychol. 2017;62(4):485–495. doi:10.1037/rep0000122.
15. 15. Pediatric Heart Network. MINDS-ACHD public-use dataset and documentation. https://www.pediatricheartnetwork.org/studies/minds-achd/. Accessed 27 Sep 2026.
16. 16. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. PLoS Med. 2007;4(10):e296. doi:10.1371/journal.pmed.0040296.
17. 17. Weintraub S, Dikmen SS, Heaton RK, Tulsky DS, Zelazo PD, Slotkin J, et al. The cognition battery of the NIH toolbox for assessment of neurological and behavioral function: validation in an adult sample. J Int Neuropsychol Soc. 2014;20(6):567–578. doi:10.1017/s1355617714000320.
18. 18. Heaton RK, Akshoomoff N, Tulsky D, Mungas D, Weintraub S, Dikmen S, et al. Reliability and validity of composite scores from the NIH Toolbox Cognition Battery in adults. J Int Neuropsychol Soc. 2014;20(6):588–598. doi:10.1017/s1355617714000241.
19. 19. Karr JE, Ingram EO, Pinheiro CN, Ali S, Iverson GL. Test-Retest Reliability and Reliable Change on the NIH Toolbox Cognition Battery. Arch Clin Neuropsychol. 2024;39(6):702–713. doi:10.1093/arclin/acae011.
20. 20. Cella D, Lai JS, Nowinski CJ, Victorson D, Peterman A, Miller D, et al. Neuro-QOL: brief measures of health-related quality of life for clinical research in neurology. Neurology. 2012;78(23):1860–1867. doi:10.1212/WNL.0b013e318258f744.
21. 21. Gwet KL. Computing inter-rater reliability and its variance in the presence of high agreement. Br J Math Stat Psychol. 2008;61(Pt 1):29–48. doi:10.1348/000711006X126600.
22. 22. Zigmond AS, Snaith RP. The hospital anxiety and depression scale. Acta Psychiatr Scand. 1983;67(6):361–370. doi:10.1111/j.1600-0447.1983.tb09716.x.
23. 23. Snaith RP. The Hospital Anxiety And Depression Scale. Health Qual Life Outcomes. 2003;1:29. doi:10.1186/1477-7525-1-29.
24. 24. Ho EH, Han YC, LaForte EM, Kaat AJ, Dworak EM, Ece B, et al. Norming of the NIH Toolbox Cognition Battery Version 3. Assessment. Published online 15 November 2025. doi:10.1177/10731911251386277.
25. 25. Shono Y, Ece B, Ho EH, Kaat AJ, LaForte EM, Ayturk E, et al. A comparison of scoring algorithms for the NIH Toolbox executive function tasks in a U.S. norming sample. Psychol Assess. 2024;36(12):760–771. doi:10.1037/pas0001350.

26. 26. Kusec A, Demeyere N. Relationship of subjective and objective cognition with post-stroke mood differs between early and long-term stroke. Clin Neuropsychol. 2025;39(6):1651–1672. doi:10.1080/13854046.2024.2417865.
27. 27. Kovacs AH, Luyckx K, Thomet C, Budts W, Enomoto J, Sluman MA, et al. Anxiety and Depression in Adults With Congenital Heart Disease. J Am Coll Cardiol. 2024;83(3):430–441. doi:10.1016/j.jacc.2023.10.043.
28. 28. Moons P, Van Bulck L, Daelman B, Luyckx K. Mental health in adult congenital heart disease. Int J Cardiol Congenit Heart Dis. 2023;12:100455. doi:10.1016/j.ijcchd.2023.100455.
29. 29. Ko JM, Kutty S, Van Bulck L, Jackson JL, Caruana M, Jameson SM, et al. Association Between a Novel Adult Congenital Heart Disease-Specific Patient-Reported Health Status Metric and Objective Clinical Status. Circ Heart Fail. 2026;19(1):e012860. doi:10.1161/circheartfailure.125.012860.
30. 30. Ilardi D, Zaidi AN, Calderon J, Choi JJ, Kasparian NA, Pike NA, et al. Neurocognitive and Psychological Interventions for Adults With Congenital Heart Disease: A Consensus Statement From the Adult Congenital Heart Disease/Neuro 2024 Conference. J Am Heart Assoc. 2026;15(2):e042921. doi:10.1161/jaha.125.042921.

Figures

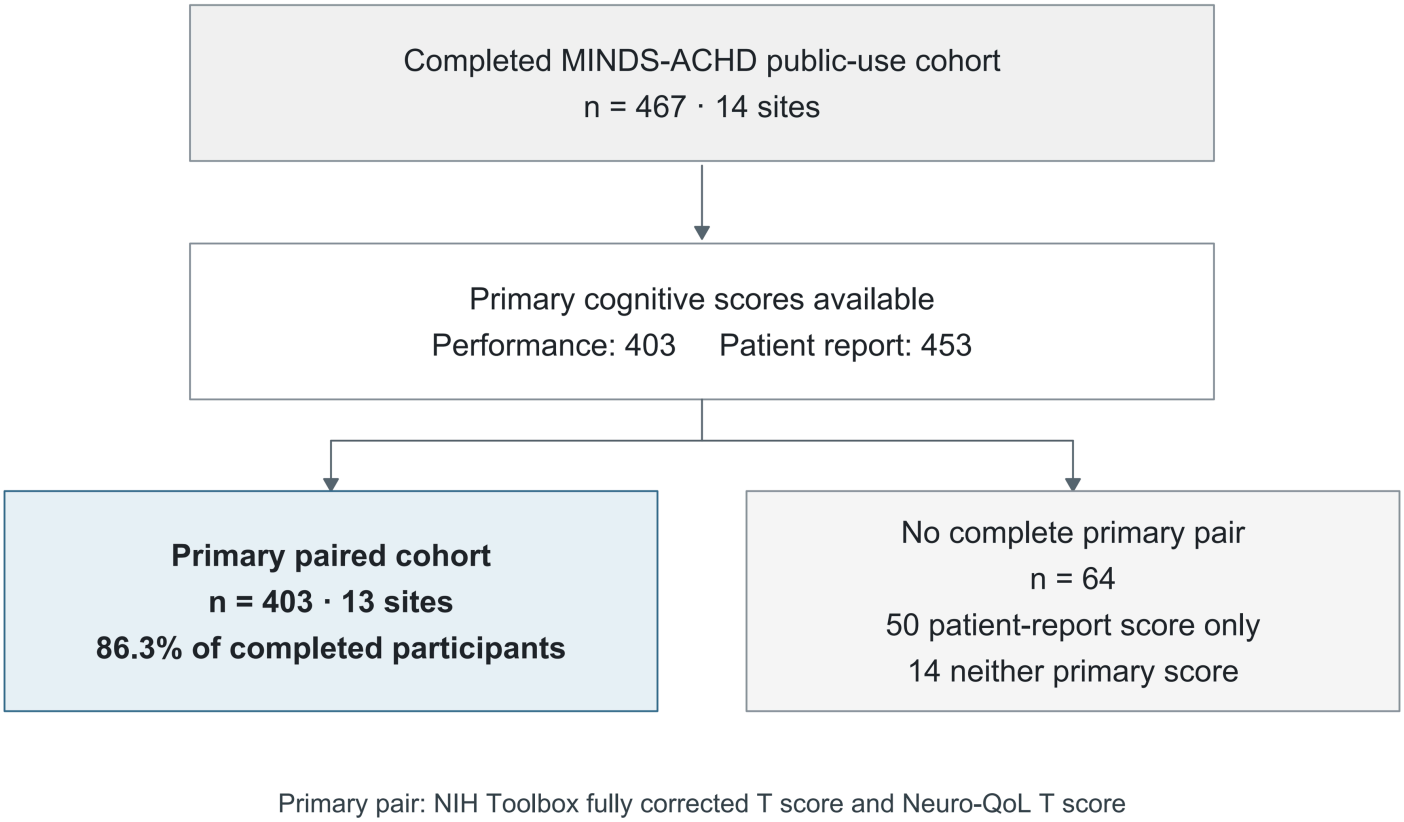

Figure 1

Selection of the paired analytic cohort.

Counts were verified in the MINDS-ACHD public-use files. Primary pairing required the NIH Toolbox Cognition Total Composite Fully Corrected T score and Neuro-QoL Cognitive Function T score. The paired cohort included 403 participants (86.3% of the completed cohort) from 13 sites. Of 64 participants without a complete primary pair, 50 had only the patient-report score and 14 had neither primary score.

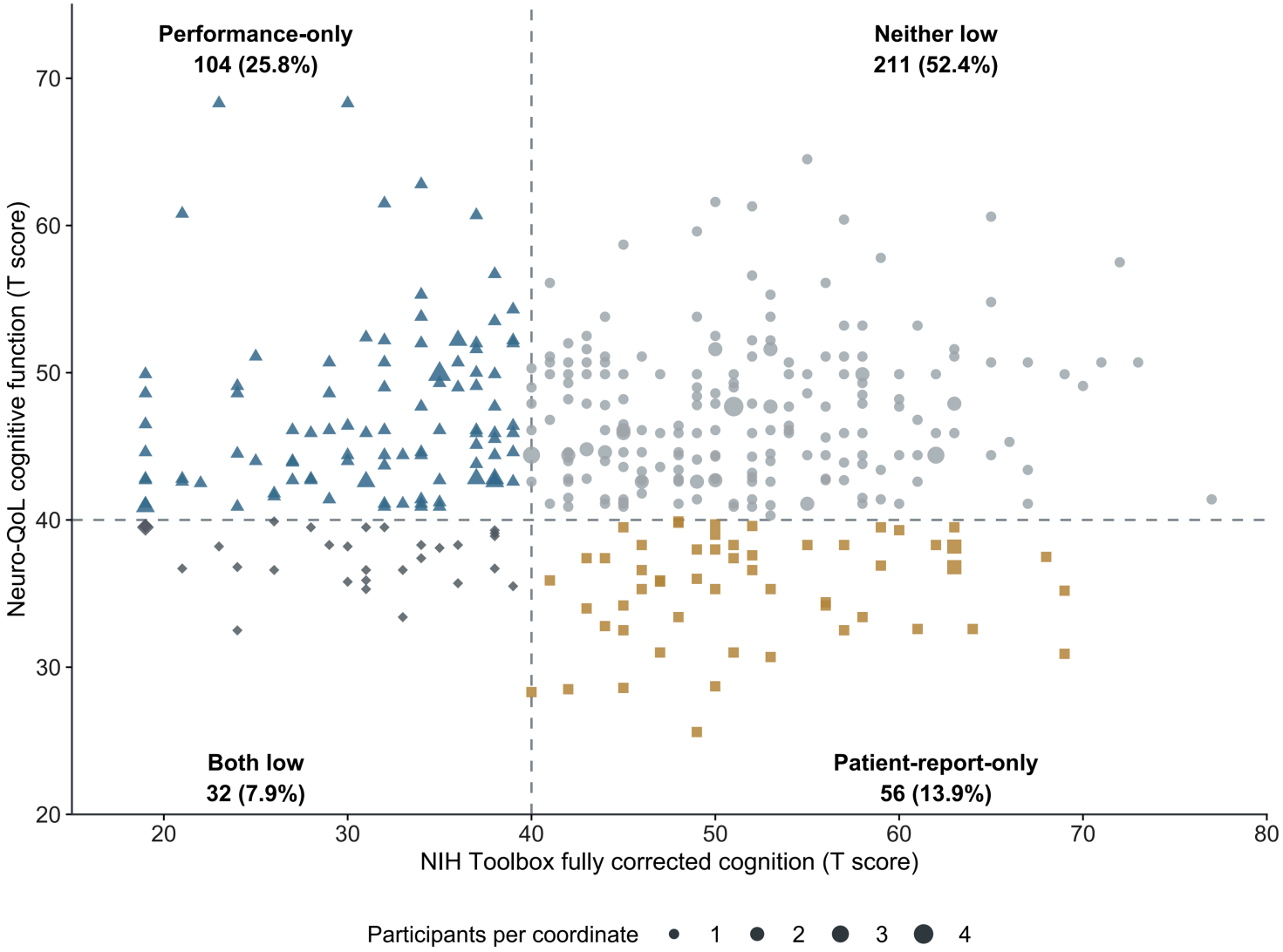

Figure 2

Overlap between performance-based and patient-reported cognitive scores.

Points retain exact observed coordinates. Point area is proportional to the number of participants sharing a coordinate; 369 distinct coordinates represent 403 participants. Dashed lines mark T scores of 40; low scores are strictly below 40. Color, shape, and quadrant labels identify the categories. Percentages use the denominator of 403. Pearson r=0.042 (95% CI, -0.058 to 0.142).

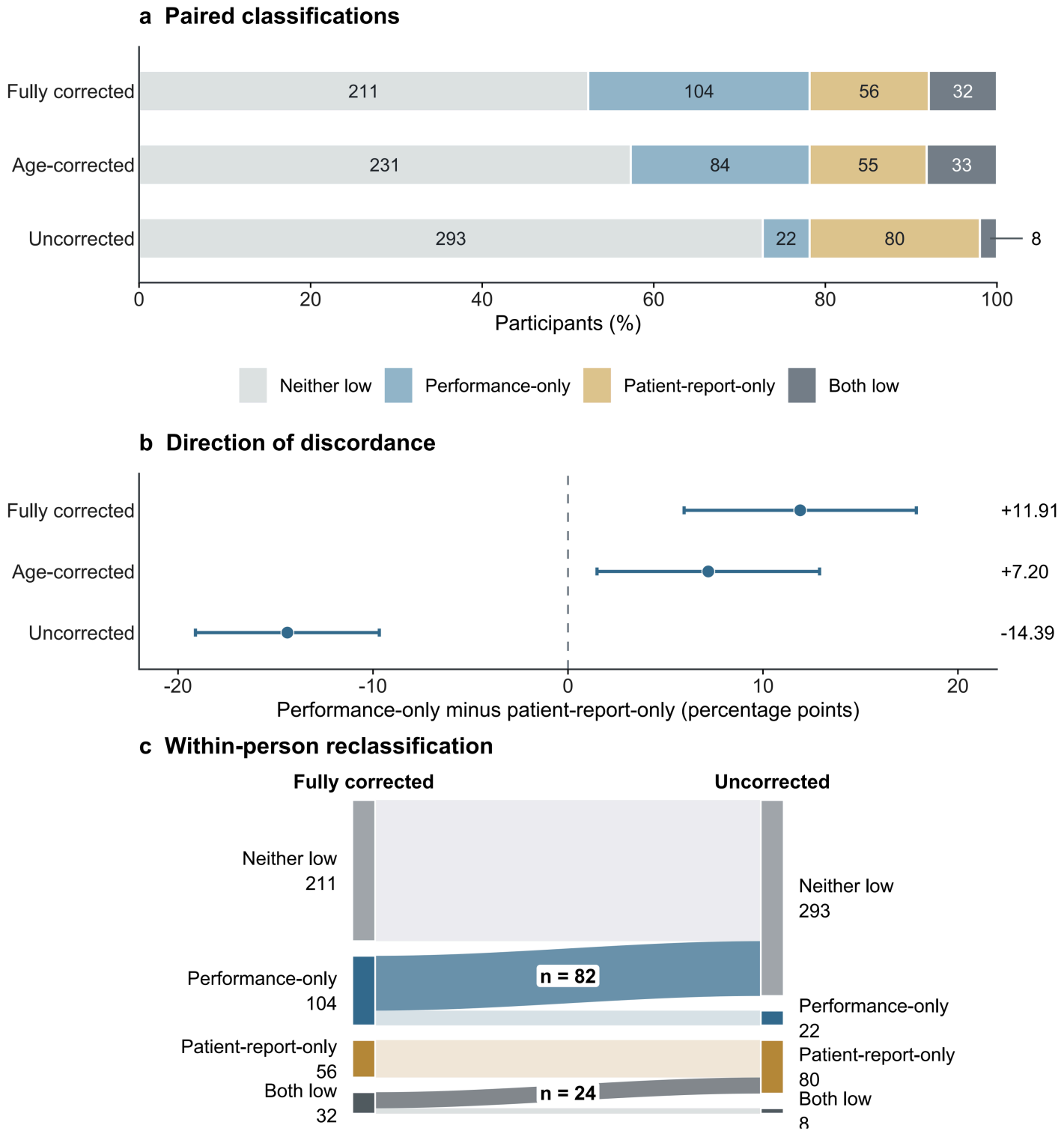

Figure 3

Cognitive classifications and reclassification across normative scores.

(a) Four-category distributions among the same 403 participants. Fully corrected T scores below 40, age-corrected standard scores below 85, and uncorrected standard scores below 85 define performance-low status. The Neuro-QoL threshold remains below 40. Segment labels give participant

counts. (b) Performance-only minus patient-report-only prevalence, in percentage points. Horizontal bars show 95% percentile CIs from 5,000 site-stratified participant bootstrap replicates. Zero denotes equal prevalence of the discordant categories. Direct paired contrasts are in Supplementary Table S7. (c) Reclassification from fully corrected to uncorrected scoring at the same assessment. Flow widths represent counts, and all flows total 403. Darker paths identify 82 participants moving from performance-only to neither-low and 24 moving from both-low to patient-report-only. Node labels give counts.

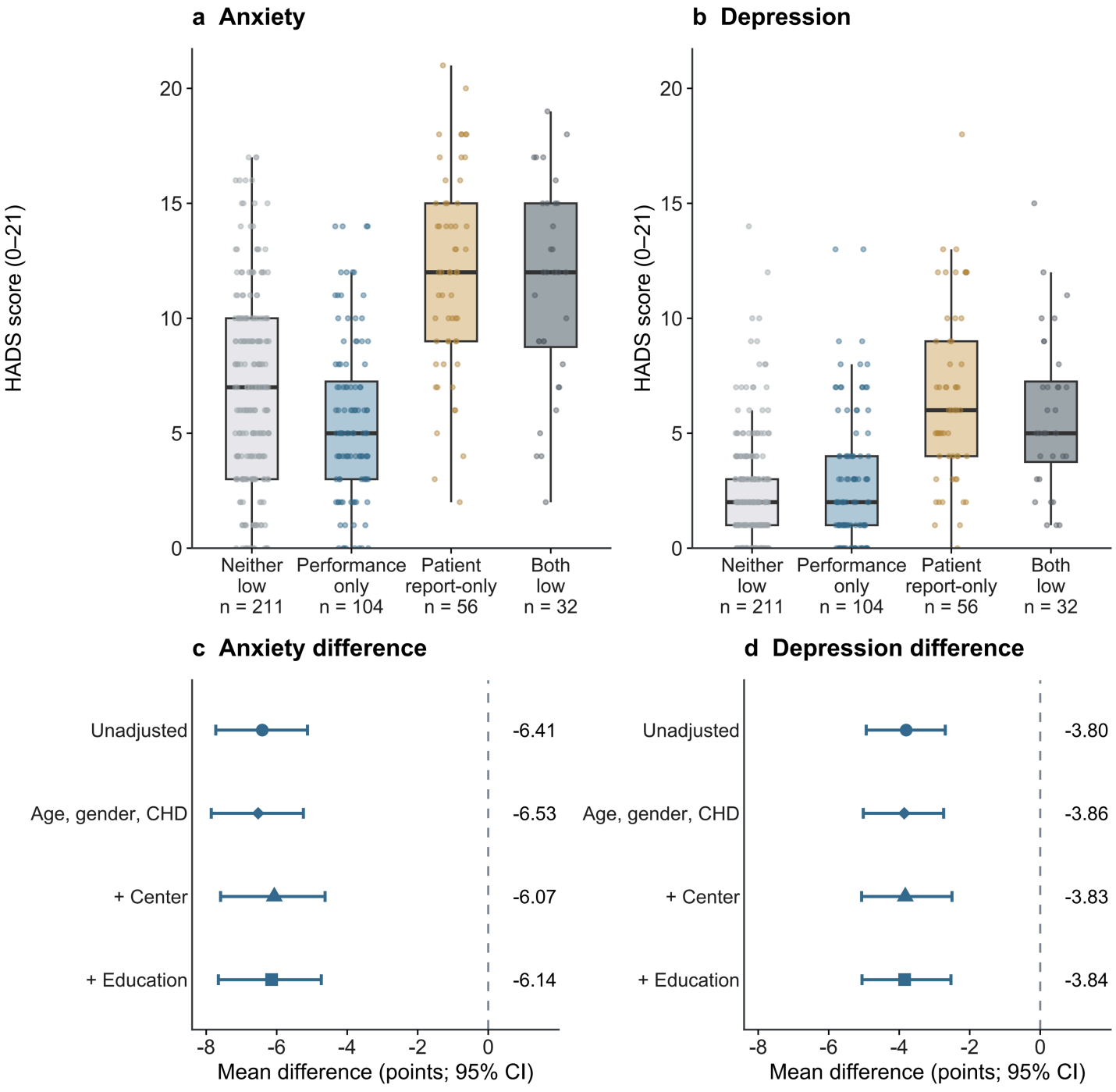

Figure 4

HADS symptom distributions and discordant-group comparisons.

(a) Anxiety and (b) depression totals range from 0 to 21, with higher scores indicating greater reported symptom burden. Boxes show medians and interquartile ranges; whiskers extend to observations within 1.5 interquartile ranges. Points show every observation, offset horizontally only. All four groups had complete HADS scores. (c) Anxiety and (d) depression mean differences compare 104 performance-only with 56 patient-report-only participants. Symbols identify the unadjusted estimate and models adjusting for age, gender, and CHD group, additionally for center, and additionally for education. Horizontal bars show 95% percentile CIs from 5,000 full-cohort bootstrap replicates stratified by center. Models were refitted among discordant participants in each replicate. Negative values indicate lower scores in the performance-only group. Adjusted comparisons are post hoc exploratory analyses; model details are in Supplementary Table S8.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- STROBEChecklist.docx
- SupplementaryMaterial.docx