RESEARCH ARTICLE

Cross-sectional and longitudinal assessments of function in prodromal-to-mild Alzheimer's disease: A comparison of the ADCS-ADL and A-IADL-Q scales

Edmond Teng[1] | Yihao Li[1] | Paul T. Manser[1] | Karen Pickthorn[1] | Brandon D. Butcher[1] | Mira Blendstrup[1] | Christopher Randolph[2] | Sietske A.M. Sikkes[3,4]

[1]Genentech, Inc., South San Francisco, California, USA

[2]WCG Clinical Endpoint Solutions, Princeton, New Jersey, USA

[3]Alzheimer Center Amsterdam, Amsterdam Neuroscience, Amsterdam University Medical Center, VU University Amsterdam, Amsterdam, the Netherlands

[4]Department of Clinical, Neuro- and Developmental Psychology, VU University, Amsterdam, the Netherlands

Correspondence
Edmond Teng, Genentech, Inc., 1 DNA Way, South San Francisco, CA 94080, USA.
E-mail: teng.edmond@gene.com

Abstract

Introduction: Prior observational work in a heterogeneous cohort of participants with mild cognitive impairment suggests the Amsterdam Instrumental Activities of Daily Living Questionnaire (A-IADL-Q) may have greater sensitivity for functional decline than the more established Alzheimer's Disease Cooperative Study–Activities of Daily Living (ADCS-ADL) scale. However, the relative utility of the A-IADL-Q versus the ADCS-ADL for clinical trials in early Alzheimer's disease (AD) remains uncertain.

Methods: We compared baseline and longitudinal performance of the A-IADL-Q and ADCS-ADL in participants with biomarker-confirmed prodromal (pAD; $n = 158$) or mild (mAD; $n = 283$) AD enrolled in the 18-month Tauriel study of semorinemab (NCT03289143).

Results: The A-IADL-Q exhibited numerically stronger discrimination between pAD and mAD participants at baseline per Cohen's d analyses and similar sensitivity to longitudinal decline across cohorts over 18 months relative to the ADCS-ADL.

Discussion: The comparable performance of the ADCS-ADL and A-IADL-Q supports the utility of the A-IADL-Q in early AD clinical trials.

KEYWORDS
activities of daily living, Alzheimer's disease, assessment, clinical trials, function

Highlights

- The Amsterdam Instrumental Activities of Daily Living Questionnaire (A-IADL-Q) may be more sensitive than the Alzheimer's Disease Cooperative Study–Activities of Daily Living Scale (ADCS-ADL) for distinguishing prodromal and mild Alzheimer's disease (AD).

- A-IADL-Q and ADCS-ADL are similarly sensitive to decline in early AD over 18 months.
- Comparable performance of these indices supports A-IADL-Q use in future AD trials.
- Additional AD clinical trial data could extend findings across more diverse cohorts.

1 | INTRODUCTION

Per regulatory guidance from both the US Food and Drug Administration[1] and the European Medicines Agency,[2] the assessment of function in clinical trials for Alzheimer's disease (AD) is essential for establishing the efficacy of novel AD therapeutics.[3] Currently, the most widely used instrument for measuring function in AD clinical trials is the Alzheimer's Disease Cooperative Study–Activities of Daily Living Scale (ADCS-ADL).[4] While the ADCS-ADL appears to be sufficiently sensitive to consistently detect treatment-related changes in the rate of functional decline in moderate AD dementia, poorer sensitivity is seen at earlier stages of disease, such as mild AD dementia.[5]

As interventions for AD, particularly those targeting amyloid beta (Aβ), increasingly focus on patient populations at earlier stages of disease,[6] the continued utility of the ADCS-ADL for clinical trials in prodromal (i.e., mild cognitive impairment [MCI]) to mild (i.e., mild AD dementia) AD remains uncertain. One approach used in recent trials in prodromal-to-mild AD is to implement modified versions of the ADCS-ADL that have either been optimized for MCI (ADCS-ADL-MCI[7]) or focus on instrumental activities of daily living (iADLs), which become impaired at earlier stages of disease progression (ADCS-iADL[8]). However, even these modified versions of the ADCS-ADL have only inconsistently demonstrated treatment benefits for functional outcomes with interventions that have more robustly shown efficacy on global and cognitive endpoints.[9,10]

An alternative approach has been to develop novel functional assessments that may be more sensitive to iADLs that might become impaired at earlier stages of AD progression. One such measure is the Amsterdam Instrumental Activities of Daily Living Questionnaire (A-IADL-Q),[11] which incorporates assessments of more complex iADLs, such as the use of modern everyday technologies that may have emerged after the ADCS-ADL scale was developed. Prior published work with the A-IADL-Q has primarily focused on cross-sectional[12,13] and longitudinal[14,15] analyses across a diverse and multicultural range of observational cohorts using an item-response theory (IRT)–derived scoring algorithm. In one such cohort, a head-to-head comparison to the ADCS-ADL suggested that the A-IADL-Q may have greater sensitivity for longitudinal decline in MCI and similar sensitivity for longitudinal decline in mild dementia.[15]

It remains unclear whether the sensitivity and overall performance of the A-IADL-Q seen in observational studies comprised of clinically diagnosed participants with heterogenous underlying etiologies translates to interventional clinical trial settings with biomarker-confirmed AD participants. We addressed this question by examining the cross-sectional and longitudinal performance of A-IADL-Q in an 18-month Phase 2 study in Aβ-positive prodromal-to-mild AD participants relative to concurrent assessments with the ADCS-ADL and 13-item version of the Alzheimer's Disease Assessment Scale-Cognitive Subscale (ADAS-Cog13).[16]

2 | METHODS

The Tauriel study (NCT03289143) was a Phase 2, multi-center, randomized, double-blind, placebo-controlled, parallel-group clinical trial that assessed the safety and efficacy of semorinemab in prodromal-to-mild AD. Overall results for the 73-week blinded portion of the study, which was conducted between October 18, 2017 and July 16, 2020 at 97 sites in North America, Europe, and Australia, have previously been reported, and no clinical efficacy relative to placebo was observed with any of three doses of semorinemab that were administered.[17] As such, pooled data from all treatment arms are included in the analyses presented here.

2.1 | Consent statement

This study was approved by each center's institutional review board/ethics committee and conducted in accordance with the Declaration of Helsinki and the International Conference on Harmonization E6 Guidelines for Good Clinical Practice. All participants and/or their legally authorized representatives provided written informed consent.

2.2 | Participants

Participants eligible for the Tauriel study were between 50 to 80 years old (inclusive) at time of screening, met diagnostic criteria for MCI[18] or dementia[19] due to AD, and had: Mini-Mental State Examination (MMSE)[20] scores between 20 and 30 (inclusive), global scores on the Clinical Dementia Rating (CDR)[21] of 0.5 or 1, Repeatable

Battery for the Assessment of Neuropsychological Status (RBANS)[22] Delayed Memory Index scores of ≤85, and evidence of significant cerebral amyloid pathology confirmed by Aβ positron emission tomography scan ([18F]florbetaben, [18F]florbetapir, [18F]flutemetamol, or [18F]NAV4694 via visual read)[23–26] or cerebrospinal fluid (CSF) Aβ(1-42) levels [≤1000 pg/mL, Elecsys β-amyloid (1-42) CSF immunoassay; Roche Diagnostics]. Participants were stratified into prodromal (i.e., MCI) or mild AD cohorts per investigator interpretation and central review of the respective diagnostic criteria[18,19] prior to formal testing with ADCS-ADL or A-IADL-Q. Concurrent treatment with approved symptomatic AD medications (e.g., acetylcholinesterase inhibitors, memantine) was permitted if dosing had been stable for ≥2 months prior to the beginning of screening. Our analyses included Tauriel participants with baseline assessments on both the ADCS-ADL and A-IADL-Q.

2.3 | Outcome measures

Two informant-based functional measures, the ADCS-ADL[4] and A-IADL-Q,[11] were administered at baseline and weeks 25, 49, and 73. The ADCS-ADL was analyzed using the total score across all items. The A-IADL-Q was scored two ways. The first approach used a scaled average (A-IADL-Q SA) of scored responses, which only included iADLs that the participant had performed in their premorbid state and only accounted for deficits that the informant attributed to cognitive impairment. The second approach (A-IADL-Q IRT) used an IRT graded response model[27] that accounts for relative differences in iADL difficulty and has previously been shown to generate a unidimensional latent trait[11] that was normally distributed in a memory clinic population.[28] Using scores derived from the IRT approach, participants can be characterized as having "no problems" (≥60), "mild problems" (50–59), "moderate problems" (40–49), or "severe problems" (<40) with their functional abilities.[29] For both scoring systems, higher scores represent better functional performance. Cognitive assessments with the ADAS-Cog13[16] were concurrently administered at baseline and weeks 25, 49, and 73. Although prior work suggests relatively limited sensitivity of the ADAS-Cog13 in early AD,[30] more robust declines were seen with this measure in the Tauriel study,[17] likely due to the inclusion criterion requiring significant episodic memory impairment as measured by the RBANS.[31]

2.4 | Statistical analyses

Primary statistical analyses were performed with R (v.3.3.2).[32] Baseline comparisons between the prodromal AD (pAD) and mild AD (mAD) cohorts were conducted using t tests for continuous measures and chi-squared tests for categorical variables. Cross-sectional associations between baseline scores on the ADCS-ADL, the two A-IADL-Q scoring schemas (SA and IRT), and the ADAS-Cog13 were explored with Spearman correlations. Receiver operating characteristic (ROC)

RESEARCH IN CONTEXT

1. **Systematic Review**: The authors reviewed the literature using traditional (e.g., PubMed) sources and meeting abstracts/presentations. Comparisons of the Amsterdam Instrumental Activities of Daily Living Questionnaire (A-IADL-Q) and Alzheimer's Disease Cooperative Study–Activities of Daily Living Scale (ADCS-ADL) in more heterogenous observational cohorts of participants exhibiting cognitive impairment suggest that the A-IADL-Q may have greater sensitivity at earlier stages of cognitive decline, but the relative utility of these instruments in more homogenous clinical trial cohorts with biomarker-confirmed Alzheimer's disease (AD) has not yet been explored.

2. **Interpretation**: Our findings indicate that the A-IADL-Q and ADCS-ADL performed similarly in cross-sectional and longitudinal analyses of participants with early (prodromal-to-mild) AD enrolled in the Tauriel study of semorinemab and supports the use of the A-IADL-Q as an outcome measure in future AD clinical trials.

3. **Future Directions**: Additional comparisons of the A-IADL-Q and ADCS-ADL are needed to confirm whether the A-IADL-Q is more sensitive for detecting subtle functional decline at earlier stages of AD pathogenesis in clinical trial settings.

analyses were used to determine the utility of the ADCS-ADL and A-IADL-Q scoring schema for distinguishing between pAD and mAD participants. For the modified intent-to-treat (mITT) population (participants with ADCS-ADL/A-IADL-Q assessments at baseline and at least one post-baseline time point), longitudinal rates of change on the ADCS-ADL and A-IADL-Q were estimated using mixed models for repeated measures (MMRM) models, adjusting for baseline diagnosis (pAD vs. mAD), apolipoprotein E (*APOE*) genotype (ε4+ vs. ε4–) and baseline performance on each assessment. Longitudinal correlations between annualized linear rates of change for these functional measures and the ADAS-Cog13 were explored with Spearman correlations. Effect sizes for change from baseline at weeks 25, 48, and 73 were determined using Cohen's d statistic. Additional analyses explored the potential clinical significance of changes on the ADCS-ADL and A-IADL-Q IRT metrics. For the ADCS-ADL, prior published work based on expert opinion has suggested that a change of 2 points[33] could be considered clinically meaningful. For the A-IADL-Q IRT, separate analyses considered participants who experienced a 2.2-point drop (which has been identified as the minimal important change [MIC] for decline on this measure[34]) or who transitioned to a lower functional category[29] as exhibiting clinically meaningful decline.

TABLE 1 Demographic and clinical characteristics of the prodromal AD and mild AD participants included in cross-sectional and longitudinal analyses.

	Cross-sectional Cohort			Longitudinal cohort		
	pAD (n = 158)	mAD (n = 283)	P	pAD (n = 144)	mAD (n = 267)	P
Age, mean (SD)	69.9 (7.0)	69.6 (6.8)	0.70	70.0 (7.1)	69.5 (6.9)	0.48
Female, n (%)	84 (53.2%)	161 (56.9%)	0.45	76 (52.8%)	154 (57.7%)	0.34
Race, n (%)						
White	147 (93.0%)	261 (92.2%)	0.76	133 (92.4%)	246 (92.1%)	0.93
Education, n (%)						
≥ HS graduate	130 (82.3%)	225 (79.5%)	0.48	119 (82.6%)	210 (78.7%)	0.34
APOE ε4+, n (%)	114 (72.2%)	214 (75.6%)	0.42	104 (72.2%)	198 (74.2%)	0.67
Symp AD Tx, n (%)	81 (51.3%)	222 (78.5%)	<0.01	75 (52.1%)	211 (79.0%)	<0.01
MMSE, mean (SD)	24.8 (2.4)	22.3 (2.4)	<0.01	25.0 (2.3)	22.4 (2.4)	<0.01
CDR-SB, mean (SD)	2.7 (1.0)	4.6 (1.5)	<0.01	2.7 (1.0)	4.5 (1.5)	<0.01
ADAS-Cog13, mean (SD)	24.9 (7.0)	31.0 (7.2)	<0.01	24.7 (7.0)	30.8 (7.2)	<0.01

Abbreviations: AD, Alzheimer's disease; ADAS-Cog13, 13-item version of Alzheimer's Disease Assessment Scale–Cognitive subscale; APOE, apolipoprotein E; CDR-SB, Clinical Dementia Rating–Sum of Boxes; HS, high school; mAD, mild Alzheimer's disease; MMSE, Mini-Mental State Examination; pAD, prodromal Alzheimer's disease; RBANS-DMI, SD, standard deviation; Symp AD Tx, symptomatic treatment with an acetylcholinesterase inhibitor and/or memantine.

3 | RESULTS

3.1 | Baseline demographic and clinical characteristics

The baseline demographic and clinical characteristics of the pAD and mAD cohorts that contributed data to the cross-sectional and longitudinal analyses are described in Table 1. Cross-sectional analyses included 441 participants (158 pAD, 283 mAD), and longitudinal analyses included 411 participants (144 pAD, 267 mAD). For both the cross-sectional and longitudinal analysis cohorts, the pAD and mAD groups were similar in age and sex distribution, and had similar proportions of participants who were White, had at least a high school education, and were APOE ε4 carriers. In both cohorts, the pAD group had significantly higher MMSE scores, significantly lower CDR-SB and ADAS-Cog13 scores, and a lower proportion of participants on symptomatic AD drugs relative to the mAD group.

3.2 | Baseline ADCS-ADL and A-IADL-Q scores

Baseline ADCS-ADL and A-IADL-Q scores for the pAD and mAD groups are shown in Figure 1 (boxplot) and the Figure S1 in supporting information (density plot). Significantly higher scores were seen for the pAD group relative to the mAD group using the ADCS-ADL (t[439] = 8.19, P < 0.001), A-IADL-Q SA (t[439] = 10.77, P < 0.001), and A-IADL-Q IRT (t[439] = 11.19, P < 0.001) indices. Likewise, categorical analyses of A-IADL-Q IRT scores indicated distinct patterns of impairment between groups (χ^2[441] = 93.56, P < 0.001); greater proportions of pAD participants were classified as having "no problems" (pAD: 40.5%; mAD: 7.8%) or "mild problems" (pAD: 41.1%; mAD
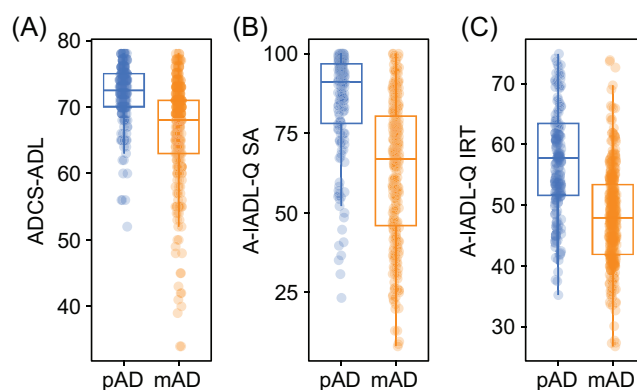

FIGURE 1 Boxplots of baseline scores for the pAD and mAD cohorts on the (A) ADCS-ADL and the A-IADL-Q scored via (B) SA (A-IADL-Q SA) and (C) IRT (A-IADL-Q IRT) approaches. ADCS-ADL, Alzheimer's Disease Cooperative Study-Activities of Daily Living Scale; A-IADL-Q, Amsterdam Instrumental Activities of Daily Living Questionnaire; IRT, item response theory; mAD, mild Alzheimer's disease; pAD, prodromal Alzheimer's disease; SA, scaled average

35.0%), while greater proportions of mAD participants were classified as having "moderate problems" (pAD: 15.8%; mAD: 41.3%) or "severe problems" (pAD: 2.5%; mAD: 15.9%). Across all participants, moderate correlations were seen between ADCS-ADL scores and both A-IADL-Q scoring schemas (SA: r_s = 0.63; IRT: r_s = 0.62), with stronger correlations seen in the mAD cohort (SA: r_s = 0.57; IRT: r_s = 0.56) relative to the pAD cohort (SA: r_s = 0.38; IRT: r_s = 0.36). A-IADL-Q scores calculated via the SA and IRT approaches were highly correlated (all participants: r_s = 0.99; pAD: r_s = 0.90; mAD: r_s = 0.90).

Cohen's d scores indicated that each of these approaches clearly discriminated between the pAD and mAD cohorts, though numerically larger Cohen's d scores were seen with the A-IADL-Q indices (SA: 1.06,
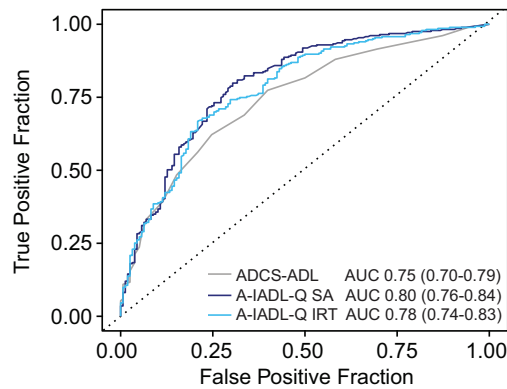
FIGURE 2 ROC curve for distinguishing participants diagnosed with prodromal AD versus mild AD at baseline using scores on the ADCS-ADL and the A-IADL-Q (scored via SA [A-IADL-Q SA] and IRT [A-IADL-Q SA IRT] approaches). AD, Alzheimer's disease; ADCS-ADL, Alzheimer's Disease Cooperative Study-Activities of Daily Living Scale; A-IADL-Q, Amsterdam Instrumental Activities of Daily Living Questionnaire; IRT, item response theory; ROC, receiver operating characteristic; SA, scaled average

95% confidence interval [CI] 0.83–1.29; IRT 1.11, 95% CI 0.87–1.34) relative to the ADCS-ADL (0.82, 95% CI 0.60–1.04). Similarly, ROC analyses (Figure 2) suggest moderately strong discrimination between pAD and mAD using each approach, with numerically larger area under the curve (AUC) values with the A-IADL-Q indices (SA: 0.80, 95% CI 0.75–0.84; IRT 0.78, 95% CI 0.74–0.83) relative to the ADCS-ADL (0.75, 95% CI 0.70–0.80).

Modest cross-sectional correlations between ADAS-Cog and ADCS-ADL scores have previously been reported in AD cohorts, with stronger correlations seen at more advanced stages of disease.[35] Similarly, analyses of baseline data from the Tauriel study yielded modest cross-sectional correlations between ADAS-Cog13 and ADCS-ADL scores across all participants (Figure 3A; $r_s = -0.47$, $P < 0.001$), with lower Spearman correlation coefficients seen in pAD ($r_s = -0.33$, $P < 0.001$) and mAD ($r_s = -0.36$, $P < 0.001$) subgroups. Likewise, modest overall correlations were seen between baseline ADAS-Cog13 and A-IADL-Q scores (SA [Figure 3B]: $r_s = -0.43$, $P < 0.001$; IRT [Figure 3C]: $r_s = -0.42$, $P < 0.001$), with lower Spearman correlation coefficients again seen in pAD (SA: $r_s = -0.32$, $P < 0.001$; IRT: $r_s = -0.29$, $P < 0.001$) and mAD (SA: $r_s = -0.28$, $P < 0.001$; IRT: $r_s = -0.28$, $P < 0.001$) subgroups.

3.3 | Longitudinal ADCS-ADL and A-IADL-Q scores

We subsequently assessed longitudinal change on the ADCS-ADL and A-IADL-Q metrics in the pAD and mAD cohorts over 18 months (Figures 4A–C). Declines in ADL performance were seen over this interval in both cohorts on all measures, more markedly in the mAD cohort (detectable by week 25) than in the pAD cohort (detectable by week 48). When raw scores are considered, the A-IADL-Q SA metric demonstrated both a broader dynamic range and greater variance

than both the ADCS-ADL and A-IADL-Q IRT metrics. When normalized to variance using Cohen's d (Figure 4D–F), similar effect sizes for change from baseline over time were observed with each metric in both cohorts, indicating that they demonstrate similar sensitivity for functional decline for patient population over 18 months.

Longitudinal correlations between changes on the ADAS-Cog13 and the ADCS-ADL or A-IADL-Q indices were performed using annualized linear rates of change on each measure (Figures 3D-3F). In the overall study cohort, modest correlations between cognitive and functional decline were seen with each measure (ADCS-ADL: $r_s = -0.38$, $P < 0.001$; A-IADL-Q SA: $r_s = -0.39$, $P < 0.001$; A-IADL-Q IRT: $r_s = -0.39$, $P < 0.001$). On the ADCS-ADL (Figure 3D), stronger longitudinal correlations were seen in mAD ($r_s = -0.40$, $P < 0.001$) relative to pAD ($r_s = -0.24$, $P = 0.003$). However, for the A-IADL-Q indices (Figures 3E and 3F), similar longitudinal correlations were seen in mAD (SA: $r_s = -0.35$, $P < 0.001$; IRT: $r_s = -0.38$, $P < 0.001$) and pAD (SA: $r_s = -0.38$, $P < 0.001$; IRT: $r_s = -0.36$, $P < 0.001$).

We also sought to determine what proportion of participants experienced potentially clinically meaningful declines per changes on the ADCS-ADL or A-IADL-Q IRT metrics. Using a threshold of a 2-point decrease on the ADCS-ADL[33] resulted in the identification of the vast majority of participants as experiencing clinically meaningful decline at each time point: 83.1% (pAD: 79.7%; mAD 85.0%) at week 25, 87.6% (pAD: 89.1%; mAD: 86.7%) at week 49, and 93.0% (pAD: 93.9%; mAD: 92.5%) at week 73. Using the previously defined MIC for decline on the A-IADL-Q IRT (−2.2 points relative to baseline),[34] clinically meaningful declines were observed in substantial proportions of participants at each time point: 46.2% (pAD: 38.6%; mAD 50.2%) at week 25, 63.4% (pAD: 52.5%; mAD: 69.3%) at week 49, and 72.5% (pAD: 65.6%; mAD: 76.2%) at week 73. A categorical approach to this question is to examine the proportion of participants at each time point that declined by ≥1 functional level based on A-IADL-Q IRT classifications:[29] 27.2% (pAD: 23.1%; mAD: 28.9%) at week 25, 44.9% (pAD: 32.8%; mAD: 51.2%) at week 49, and 52.5% (pAD: 47.2%; mAD: 55.7%) at week 73. When participants with a baseline classification of "severe problems" (i.e., lowest possible functional level) were excluded from this analysis, slightly higher proportions of participants fulfilled these criteria: 33.6% (pAD: 24.1%; mAD: 40.5%) at week 25, 50.1% (pAD: 33.8%; mAD: 60.2%) at week 49, and 58.4% (pAD: 48.3%; mAD: 65.2%) at week 73.

4 | DISCUSSION

Our cross-sectional and longitudinal analyses of ADCS-ADL and A-IADL-Q data from a clinical trial cohort of early AD participants indicate that both instruments discriminate between pAD and mAD at baseline and detect longitudinal functional decline across 18 months. The A-IADL-Q results, which were robust across two different scoring algorithms, reinforce the utility of this instrument for assessing function in clinical trials of novel therapeutics targeting this segment of the AD patient population, and raise the possibility that it may be more sensitive than the ADCS-ADL for the detection of functional impair-
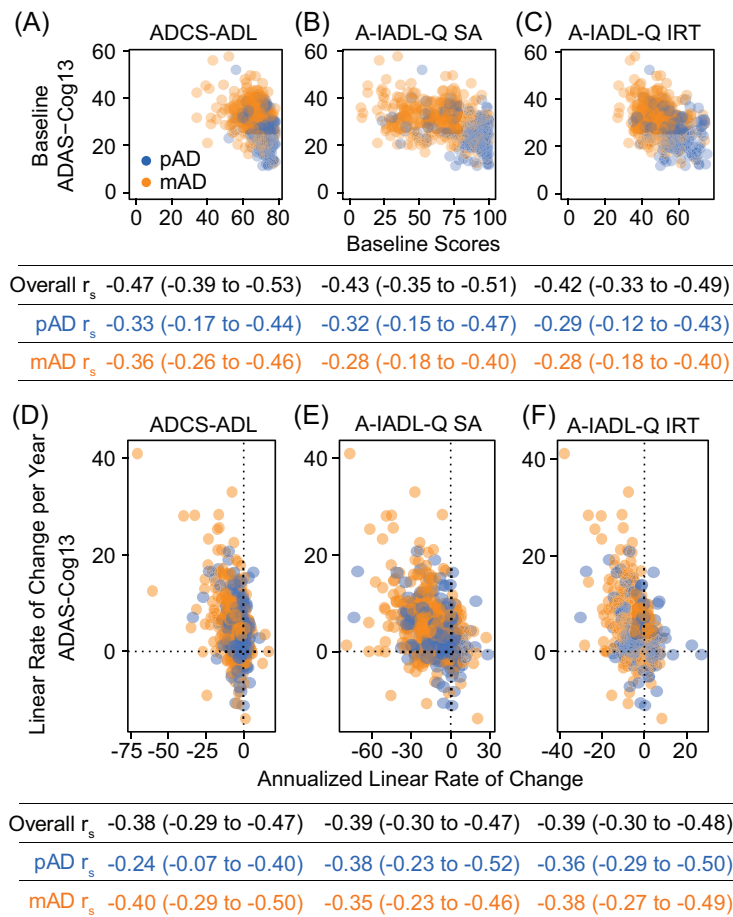
FIGURE 3 Scatterplots of baseline (A–C) and longitudinal (D–F) correlations between scores on the ADAS-Cog13 and the ADCS-ADL (A and D) and the A-IADL-Q SA (B and E) and A-IADL-Q IRT (C and F) approaches in the overall Tauriel participant population and in the pAD and mAD cohorts. Parentheses in table indicate 95% confidence intervals for Spearman correlation coefficients. ADAS-Cog13, Alzheimer's Disease Assessment Scale–Cognitive subscale 13-item version; ADCS-ADL, Alzheimer's Disease Cooperative Study-Activities of Daily Living Scale; A-IADL-Q IRT, Amsterdam Instrumental Activities of Daily Living Questionnaire scored via item response theory; A-IADL-Q SA, Amsterdam Instrumental Activities of Daily Living Questionnaire scored via scaled average; mAD, mild Alzheimer's disease; pAD, prodromal Alzheimer's disease.

	ADCS-ADL	A-IADL-Q SA	A-IADL-Q IRT
Overall r_s	-0.47 (-0.39 to -0.53)	-0.43 (-0.35 to -0.51)	-0.42 (-0.33 to -0.49)
pAD r_s	-0.33 (-0.17 to -0.44)	-0.32 (-0.15 to -0.47)	-0.29 (-0.12 to -0.43)
mAD r_s	-0.36 (-0.26 to -0.46)	-0.28 (-0.18 to -0.40)	-0.28 (-0.18 to -0.40)

	ADCS-ADL	A-IADL-Q SA	A-IADL-Q IRT
Overall r_s	-0.38 (-0.29 to -0.47)	-0.39 (-0.30 to -0.47)	-0.39 (-0.30 to -0.48)
pAD r_s	-0.24 (-0.07 to -0.40)	-0.38 (-0.23 to -0.52)	-0.36 (-0.29 to -0.50)
mAD r_s	-0.40 (-0.29 to -0.50)	-0.35 (-0.23 to -0.46)	-0.38 (-0.27 to -0.49)

ment and comparable in sensitivity for measuring additional functional decline with disease progression.

Cross-sectionally, scores on the ADCS-ADL and the A-IADL-Q were moderately well correlated (using either scoring approach for the A-IADL-Q), suggesting that across the entire study population, they measure similar underlying functional constructs. However, relatively weaker correlations were seen between the two instruments in the pAD cohort relative to the mAD cohort. One possible explanation for the relative divergence seen in the pAD cohort is that the A-IADL-Q may identify deficits arising at earlier disease stages, when the ADCS-ADL scale may be less sensitive.[36]

Using the ADCS-ADL and both scoring approaches for the A-IADL-Q, more profound functional deficits were seen at baseline in the mAD cohort relative to the pAD cohort. These results are consistent with prior cross-sectional analyses of the observational Catch-Cog cohort, which demonstrated that scores on both the A-IADL-Q and ADCS-ADL differed between participants with MCI and AD dementia.[36] In both the current and Catch-Cog analyses, numerically larger effect sizes are seen with the A-IADL-Q relative to the ADCS-ADL, though larger Cohen's d values are seen with the ADCS-ADL in our dataset (0.82) than in the Catch-Cog dataset (0.46). The relatively better discrimination between pAD and mAD participants with the A-IADL-Q may arise from the greater number and broader range of activities surveyed in this instrument relative to the ADCS-

ADL. This may in turn be reflected by the broader distribution of scores seen with the two scoring approaches for the A-IADL-Q versus the ADCS-ADL as ceiling effects may hamper the latter assessment in early AD.[36] While similar Cohen's d values were seen with the SA (1.06) and IRT (1.11) scoring of the A-IADL-Q, the latter approach may further mitigate ceiling effects, particularly among pAD participants (Figure S1).

Regulatory guidance has emphasized the importance of assessing both cognitive and functional outcomes in AD trials to capture the full clinical impact of this disease.[1,2] Both ADCS-ADL and A-IADL-Q scores demonstrate similar (though only moderate) cross-sectional and longitudinal correlations with cognition as measured by the ADAS-Cog13, which are consistent with previously reported correlations between the ADCS-ADL and the ADAS-Cog14 in both mild and moderate AD cohorts.[35] These convergent results further support the partial independence of cognitive versus functional outcomes in AD.

While our cross-sectional analyses corroborate prior work suggesting that A-IADL-Q could be more sensitive than the ADCS-ADL for identifying functional impairment in prodromal AD,[34] our longitudinal analyses found that the sensitivity for detecting change in function over 18 months in early AD appeared to be similar across both instruments, including when the pAD and mAD cohorts were considered separately. Our findings contrast with longitudinal data from the
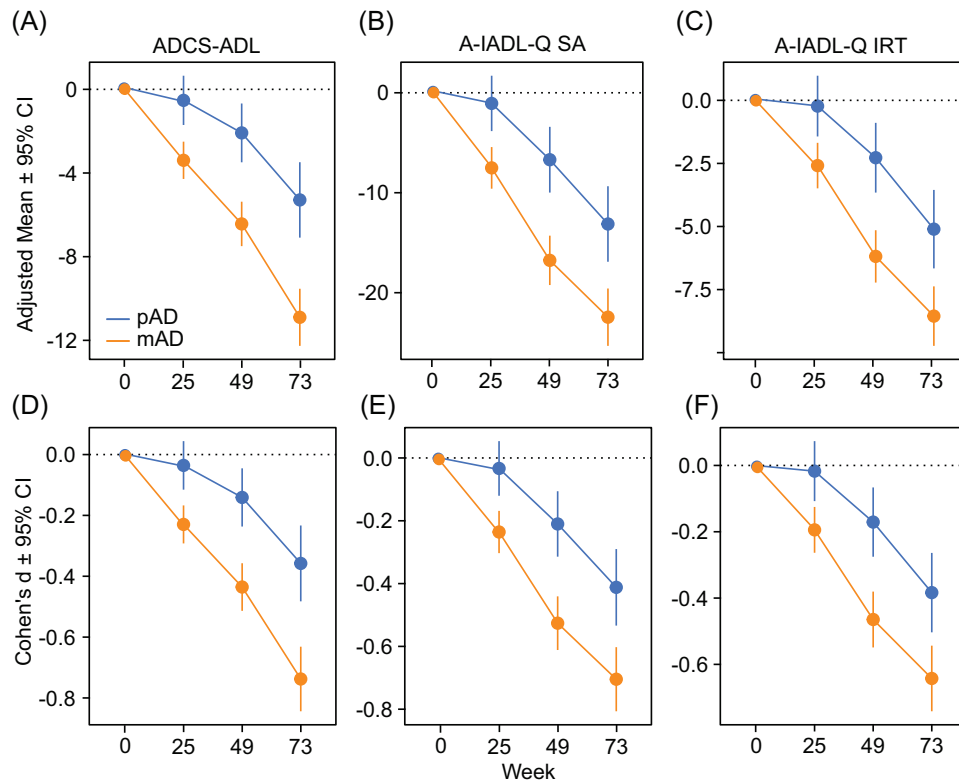
FIGURE 4 Longitudinal change on the ADCS-ADL and the A-IADL-Q as measured by adjusted mean change from baseline using an MMRM model (A–C) and effect size of change from baseline using Cohen's d (D–F). ADCS-ADL, Alzheimer's Disease Cooperative Study-Activities of Daily Living Scale; A-IADL-Q, Amsterdam Instrumental Activities of Daily Living Questionnaire; A-IADL-Q SA, Amsterdam Instrumental Activities of Daily Living Questionnaire scored via scaled average; A-IADL-Q IRT, Amsterdam Instrumental Activities of Daily Living Questionnaire scored via item response theory; A-IADL-Q SA, Amsterdam Instrumental Activities of Daily Living Questionnaire scored via scaled average; CI, confidence interval; MMRM, mixed models for repeated measures

Catch-Cog cohort, which suggested that relative to the ADCS-ADL, the A-IADL-Q was more sensitive when measuring functional decline in MCI but less sensitive when measuring functional decline in mild AD.[15] The differences between these results may be attributable to their study populations, as the inclusion criteria for the Tauriel study resulted in pAD and mAD cohorts that had greater cognitive and functional impairment at baseline relative to the analogous cohorts in the Catch-Cog study. As such, the more advanced disease in the Tauriel pAD cohort may have limited our ability to detect differences in sensitivity for longitudinal change between the ADCS-ADL and A-IADL-Q scales. Additional differences between the studies include the requirement for positive Aβ biomarkers in the Tauriel study[17] and the shorter duration of follow-up (12 months) and smaller size of the MCI (n = 75) and AD (n = 72) cohorts in the Catch-Cog study.[15]

Closer examination of the trajectories of change on the ADCS-ADL and A-IADL-Q indicate that significant longitudinal functional declines in the mAD cohort were evident by the week 24 time point and generally linear over the entire 18 months of follow-up. However, for the pAD participants, only small non-significant declines were seen on both scales by week 24, though more significant and linear declines were seen at subsequent assessments. Taken together, these findings suggest that more consistent and predictable longitudinal declines in pAD

on both the ADCS-ADL and A-IADL-Q may only become apparent over longer intervals.

When assessments of clinically meaningful functional decline on the A-IADL-Q IRT were analyzed using continuous[34] or categorical[29] approaches, the majority of participants met these thresholds by week 73 using either method. As might be expected, higher rates of clinically meaningful decline on the A-IADL-Q IRT were seen with longer follow-up, in mAD relative to pAD, and with continuous relative to categorical analyses. Overall, rates of clinically meaningful decline on the ADCS-ADL were far higher than on the A-IADL-Q. However, direct comparisons between these analyses may not be appropriate, given that the published thresholds on the ADCS-ADL are based on expert opinion[32,37] while the A-IADL-Q IRT thresholds were derived through qualitative caregiver-driven approaches.[29,34]

A number of additional factors may affect the interpretation of our results. A number of study visits and assessments occurred after the March 11, 2020 World Health Organization declaration of the COVID-19 pandemic,[17] including ~60% of the week 73 ADCS-ADL and A-IADL-Q assessments. While regional restrictions imposed at the beginning of the pandemic may have disrupted iADL performance, longitudinal trajectories on these indices appear to be essentially linear between weeks 25 and 73 (Figure 4), suggesting only minimally effects of such restrictions on our findings. Pandemic-associated restrictions

also caused ~15% of week 73 assessments to be collected telephonically as opposed to in clinic.[17] Although there are no published studies comparing telephonic versus in-person administration of the ADCS-ADL or the A-IADL-Q, the instructions of the ADCS-ADL scale explicitly indicate that it can be assessed in both settings, and the informant-based multiple-choice nature of both assessments would appear to be robust in either setting. The relatively limited ethnic and racial diversity of participants in the Tauriel study may restrict the generalizability of our analyses of the ADCS-ADL and A-IADL-Q, as previous reports have suggested minor cross-region differences on these indices.[13,38]

Overall, our results indicate that the ADCS-ADL and A-IADL-Q performed similarly for measuring rates of functional decline in early AD participants enrolled in the Tauriel study and support the inclusion of the A-IADL-Q as a key outcome measure in future trials of AD therapeutics. The A-IADL-Q may be particularly useful for longitudinal assessments of iADLs in populations at earlier stages of AD progression than those included in the current study and/or with study designs focused on time-to-event analyses for meaningful clinical decline, though additional comparisons to the ADCS-ADL will be needed to determine the relative utility of these measures in those settings.

ACKNOWLEDGMENTS

CONFLICT OF INTEREST STATEMENT

Dr. Teng, Dr. Li, Dr. Manser, Dr. Pickthorn, Dr. Butcher, and Ms. Blendstrup are employees of Genentech, Inc. and shareholders in F. Hoffmann La Roche, Ltd. Dr. Randolph has nothing to disclose. Dr. Sikkes receives funding from Health~Holland, Topsector Life Sciences & Health (LSHM19051, LSHM20084), license fees for the use of the Amsterdam IADL Questionnaire from Genentech, Inc., Medavante, Janssen, Axon Neuroscience, Toyama, Roche, Vivoryon, Alzheon, VtV Therapeutics, Green Valley, and consultancy fees from Biogen, Boehringer, Toyama, and Lundbeck, all through her institution (Amsterdam UMC). Author disclosures are available in the supporting information.

DATA AVAILABILITY STATEMENT

Qualified researchers may request access to individual patient-level data through the clinical study data request platform (https://vivli.org/). Further details on Roche's criteria for eligible studies are available here: https://vivli.org/members/ourmembers/. For further details on Roche's Global Policy on the Sharing of Clinical Information and how to request access to related clinical study documents, see https://www.roche.com/innovation/process/clinical-trials/data-sharing/.

ORCID

Edmond Teng https://orcid.org/0000-0002-1375-857X

REFERENCES

1. US Food and Drug Administration. Alzheimer's disease: developing drugs for treatment guidance for industry. 2018.
2. Committee for Medicinal Products for Human Use. *Guideline on the Clinical Investigation of Medicines for the Treatment of Alzheimer's Disease.* CHMP; 2018.
3. Snyder PJ, Kahle-Wrobleski K, Brannan S, et al. Assessing cognition and function in Alzheimer's disease clinical trials: do we have the right tools? *Alzheimers Dement.* 2014;10:853-860.
4. Galasko D, Bennett D, Sano M, et al. An inventory to assess activities of daily living for clinical trials in Alzheimer's disease. The Alzheimer's Disease Cooperative Study. *Alzheimer Dis Assoc Disord.* 1997;11(2):S33-S39.
5. Siemers E, Holdridge KC, Sundell KL, Liu-Seifert H. Function and clinical meaningfulness of treatments for mild Alzheimer's disease. *Alzheimers Dement: Diag Assess Dis Monit.* 2016;2:105-112.
6. Sperling RA, Jack CR, Jr, Aisen PS. Testing the right target and right drug at the right stage. *Sci Transl Med.* 2011;3:111cm33.
7. Pedrosa H, De Sa A, Guerreiro M, et al. Functional evaluation distinguishes MCI patients from healthy elderly people–the ADCS/MCI/ADL scale. *J Nutr Health Aging.* 2010;14:703-709.
8. Galasko D, Kershaw PR, Schneider L, Zhu Y, Tariot PN. Galantamine maintains ability to perform activities of daily living in patients with Alzheimer's disease. *J Am Geriatr Soc.* 2004;52: 1070-1076.
9. Mintun MA, Lo AC, Duggan Evans C, et al. Donanemab in early Alzheimer's disease. *N Engl J Med.* 2021;384:1691-1704.
10. Budd Haeberlein S, Aisen PS, Barkhof F, et al. Two randomized Phase 3 studies of aducanumab in early Alzheimer's disease. *J Prev Alzheimers Dis.* 2022;9:197-210.
11. Sikkes SA, de Lange-de Klerk ES, Pijnenburg YA, et al. A new informant-based questionnaire for instrumental activities of daily living in dementia. *Alzheimers Dement.* 2012;8:536-543.
12. Facal D, Carabias MAR, Pereiro AX, et al. Assessing everyday activities across the dementia spectrum with the Amsterdam IADL Questionnaire. *Curr Alzheimer Res.* 2018;15:1261-1266.
13. Dubbelman MA, Verrijp M, Facal D, et al. The influence of diversity on the measurement of functional impairment: an international validation of the Amsterdam IADL Questionnaire in eight countries. *Alzheimers Dement: Diag Assess Dis Monit.* 2020;12:e12021.
14. Koster N, Knol DL, Uitdehaag BM, Scheltens P, Sikkes SA. The sensitivity to change over time of the Amsterdam IADL Questionnaire. *Alzheimers Dement.* 2015;11:1231-1240.
15. Jutten RJ, Harrison JE, Brunner AJ, et al. The cognitive-functional composite is sensitive to clinical progression in early dementia: longitudinal findings from the Catch-Cog study cohort. *Alzheimers Dement: Trans Res Clin Interv.* 2020;6:e12020.
16. Mohs RC, Knopman D, Petersen RC, et al. Development of cognitive instruments for use in clinical trials of antidementia drugs: additions to the Alzheimer's Disease Assessment Scale that broaden its scope. The Alzheimer's Disease Cooperative Study. *Alzheimer Dis Assoc Disord.* 1997;11(2):S13-S21.
17. Teng E, Manser PT, Pickthorn K, et al. Safety and efficacy of semorinemab in individuals with prodromal to mild Alzheimer disease: a randomized clinical trial. *JAMA Neurol.* 2022;79:758-767.
18. Albert MS, DeKosky ST, Dickson D, et al. The diagnosis of mild cognitive impairment due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement.* 2011;7:270-279.
19. McKhann GM, Knopman DS, Chertkow H, et al. The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement.* 2011;7:263-269.

20. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res.* 1975;12:189-198.

21. Hughes CP, Berg L, Danziger WL, Coben LA, Martin RL. A new clinical scale for the staging of dementia. *Br J Psychiatry.* 1982;140:566-572.

22. Randolph C, Tierney MC, Mohr E, Chase TN. The Repeatable Battery for the Assessment of Neuropsychological Status (RBANS): preliminary clinical validity. *J Clin Exp Neuropsychol.* 1998;20:310-319.

23. Piramal Imaging. Neuraceq (florbetaben F 18 injection) [package insert]. Updated March 2014. Accessed December 28, 2020. https://www.accessdata.fda.gov/drugsatfda_docs/label/2014/204677s000lbl.pdf

24. Lilly Eli, and Company. Amyvid (florbetapir F 18 injection) [package insert]. Updated April 2012. Accessed December 29, 2020. https://www.accessdata.fda.gov/drugsatfda_docs/label/2012/202008s000lbl.pdf

25. GE Healthcare. Vizamyl (flutemetamol F 18 injection) [package insert]. Updated April 2016. Accessed December28, 2020. https://www.accessdata.fda.gov/drugsatfda_docs/label/2016/203137s005lbl.pdf

26. Ng KP, Pascoal TA, Mathotaarachchi S, et al. Monoamine oxidase B inhibitor, selegiline, reduces (18)F-THK5351 uptake in the human brain. *Alzheimers Res Ther.* 2017;9:25.

27. Samejima F. Estimation of latent ability using a response pattern of graded scores. *Psychometrika.* 1969;34:1-97.

28. Sikkes SA, Knol DL, Pijnenburg YA, de Lange-de Klerk ES, Uitdehaag BM, Scheltens P. Validation of the Amsterdam IADL Questionnaire, a new tool to measure instrumental activities of daily living in dementia. *Neuroepidemiology.* 2013;41:35-41.

29. Dubbelman MA, Terwee CB, Verrijp M, Visser LNC, Scheltens P, Sikkes SAM. Giving meaning to the scores of the Amsterdam instrumental activities of daily living questionnaire: a qualitative study. *Health Qual Life Outcomes.* 2022;20:47.

30. Podhorna J, Krahnke T, Shear M, Harrison JE. Alzheimer's Disease Assessment Scale-Cognitive subscale variants in mild cognitive impairment and mild Alzheimer's disease: change over time and the effect of enrichment strategies. *Alzheimers Res Ther.* 2016;8.

31. Teng E, Manser PT, Shah M, et al. The use of episodic memory tests for screening in clinical trials for early Alzheimer's disease: a comparison of the Free and Cued Selective Reminding Test (FCSRT) and the Repeatable Battery for the Assessment of Neuropsychological Status (RBANS). *J Prev Alzheimers Dis.* 2023;10:41-49.

32. R Core Team. R: a language and environment for statistical computing. *R Foundation for Statistical Computing.* R Core Team. 2016.

33. Dysken MW, Sano M, Asthana S, et al. Effect of vitamin E and memantine on functional decline in Alzheimer disease: the TEAM-AD VA cooperative randomized trial. *JAMA.* 2014;311:33-44.

34. Dubbelman MA, Verrijp M, Terwee CB, et al. Determining the minimal important change of everyday functioning in dementia: pursuing clinical meaningfulness. *Neurology.* 2022;99:e954.

35. Liu-Seifert H, Siemers E, Selzler K, et al. Correlation between cognition and function across the spectrum of Alzheimer's disease. *J Prev Alzheimers Dis.* 2016;3:138-144.

36. Jutten RJ, Harrison JE, Lee Meeuw Kjoe PR, et al. Assessing cognition and daily function in early dementia using the cognitive-functional composite: findings from the Catch-Cog study cohort. *Alzheimers Res Ther.* 2019;11:45.

37. Callahan CM, Boustani MA, Schmid AA, et al. Targeting functional decline in Alzheimer disease: a randomized trial. *Ann Intern Med.* 2017;166:164-171.

38. Henley DB, Dowsett SA, Chen YF, et al. Alzheimer's disease progression by geographical region in a clinical trial setting. *Alzheimers Res Ther.* 2015;7:43.

SUPPORTING INFORMATION

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