DELIRIUM & HRQoL (from Bentvelzen et al. 2017 JAMDA)

		Delirium			HRQoL			
No	Rating Criteria	**CAM**	**DRS-R/-98**	**DI**	**QOL-AD**	**DEMQOL**	**QUALID**	**QUALIDEM**
1	Inter-rater reliability (/4)	3	4	4	2	0	2	2
2	Test-retest reliability (/4)	0	0	0	2	2	2	2
3	Internal consistency (/2)	0	1	1	1	1	1	1
4	Content validity (/2)	2	1	2	2	2	2	2
5	Concurrent validity (/4)	2	2	4	2	2	2	2
6	Discriminant validitiy (/4)	2	4	0	4	4	2	2
7	Sensitivity (/4)	4	4	0	0	0	0	0
8	Specificity (/4)	4	4	0	0	0	0	0
9	Responsiveness (/4)	0	2	4	2	4	2	2
10	Dementia types (/2)	0	0	0	4	2	0	2
11	Clinical settings (/2)	2	0.5	2	2	1	1	2
12	Education/literacy (/2)	2	2	2	1.5	1	1	0
13	Translations (/2)	2	2	1	2	2	2	2
14	International acceptance (/4)	2	2	0	0	0	0	0
15	Administration time (/4)	4	2	2	1	1	4	2
16	A: Ease of use (/4)	4	2	4	-	4	4	3
	B: Respondent burden (/4)	-	-	-	4	-	-	-
17	Qualifications required (/4)	2	2	2	3	2	2	2
18	Cost of tool/training (/4)	4	4	4	4	4	4	4
	Weighted score (/60)	**39**	**38.5**	**32**	**34**	**32**	**31**	**30**

1 Reliability 1: inter-rater
4 excellent (ICC/κ ≥ .90)
2 adequate (ICC/κ .70 to .89)
0 low (ICC/κ < .70) or no data

2 Reliability 2: test-retest
4 excellent (ICC/κ ≥ .90)
2 adequate (ICC/κ .70 to .89)
0 low (ICC/κ < .70) or no data

3 Reliability 3: internal consistency
2 excellent (Cronbach's α ≥ .90)
1 adequate to good (Cronbach's α from .70 to .89)
0 low (Cronbach's α < .70) or no data

4 Validity 1: Content validity–domain of interest is comprehensively sampled by the items
2 domain comprehensively sampled
1 domain reasonably well sampled
0 important aspects of domain are not sampled or irrelevant items included

5 Validity 2: Concurrent validity–expected correlations with similar validated measures
4 high (|r/κ| ≥ .70)
2 moderate (|r/κ| from .40 to .69)
0 low concurrent validity ((|r/κ| < .30), or no data

6 Validity 3: Discriminant validity - cross-sectional (eg, dementia vs depression; low vs high levels of severity/impairment; AD vs FTD etc.)
4 can distinguish between >2 clinically important categories of respondents
2 can distinguish between 2 categories of respondents
0 no evidence

7 Validity 4: Sensitivity to diagnosis/category
4 high (≥.85)
2 moderate (.70 to .84)
0 low (<.70)

8 Validity 5: Specificity to diagnosis/category
4 high (≥.85)
2 moderate (.70 to .84)
0 low (<.70)

9 Validity 6: Responsivenessdability to detect clinically important change over time (eg, because of course of the condition or in response to intervention)
4 availability of minimum clinically important difference (MCID) in appropriate metrics (eg, standardized response means) at the individual patient level on external clinical criteria
2 can detect statistically significant changes over time in hypothesized direction on external clinical criteria, but no metrics available to quantify MCID at the individual patient level
0 no evidence for responsiveness

10 Generalizability 1: validity in different dementia populations (eg, AD, FTD, PD etc.)
2 > 2 types of dementia
1 two different types of dementia
0 only 1 type of dementia

11 Generalizability 2: validity in different clinical settings (ie, nursing home, community, primary care, specialist)
2 > 2 types of setting
1 two different types of setting
0 only 1 type of setting

12 Generalizability 3: validity in patients with low education/literacy
2 scale shown to be resistant to low education/literacy, or effects of education/literacy shown but alternative cut-offs or corrections published
1 effect of low education/literacy on validity, but no alternative cut-offs or corrections available
0 not investigated

13 Generalizability 4: validity in multiple countries/languages
2 multiple countries or languages
1 different countries but only 1 language
0 1 country and language

14 Recommended in published international dementia guidelines
4 ≥ 2 countries
2 1 country
0 0 countries

15 Administration time (minutes)
4 ≤ 5
2 6–15
0 > 15

16A Ease of administration and scoring (for clinician-administered tools)
4 does not require algorithm to score or special equipment
2 requires an algorithm to compute score OR special equipment
0 requires an algorithm to compute score AND special equipment

16B Burden on respondent (for self-reported or proxy tools)
4 items are worded simply
2 minor challenges for respondent (eg, minority of items are worded in a complex manner)
0 reasonable degree of burden on respondent (majority of items worded in a complex manner)

17 Clinical qualifications required to administer tool
4 untrained rater (eg, general nursing staff, patient/informant)
2 paraprofessional/staff member (eg, clinical nurse; research assistant)
0 professional (eg, doctor, occupational therapist, or neuropsychologist)

18 Cost of the tool and training for clinicians
4 no charge for tool or for training
2 small 1-time costs to acquire tool or for training
0 costs charged each time tool is used