This item is the archived peer-reviewed author-version of:

Validity of the three apathy items of the Geriatric Depression Scale (GDS-3A) in measuring apathy in older persons

Reference:

Bertens Anne Suzanne, Moonen Justine E.F., de Waal Margot W.M., Foster-Dingley Jessica C., de Ruijter Wouter, Gussekloo Jacobijn, van der Mast Rosa, de Craen Anton J.M.- Validity of the three apathy items of the Geriatric Depression Scale (GDS-3A) in measuring apathy in older persons

International journal of geriatric psychiatry: a journal of the psychiatry of late life and allied sciences - ISSN 0885-6230 - (2016), p. 1-8

Full text (Publishers DOI): <http://dx.doi.org/doi:10.1002/GPS.4484>

To cite this reference: <http://hdl.handle.net/10067/1380720151162165141>

**VALIDITY OF THE THREE APATHY ITEMS OF THE GERIATRIC DEPRESSION
SCALE (GDS-3A) IN MEASURING APATHY IN OLDER PERSONS**

**Anne Suzanne Bertens, MD^{1,2*}, Justine E.F. Moonen, MD^{2*}, Margot W.M. de Waal,
PhD³, Jessica C. Foster-Dingley, MSc², Wouter de Ruijter, MD, PhD³, Jacobijn
Gussekloo MD, PhD³, Roos C. van der Mast, MD, PhD^{2,4}, Anton J.M. de Craen, PhD^{5†}**

*** Authors A.S. Bertens and J.E.F. Moonen contributed equally**

¹ Department of Radiology, Leiden University Medical Centre, Leiden, the Netherlands

² Department of Psychiatry, Leiden University Medical Centre, Leiden, the Netherlands

³ Department of Public Health and Primary Care, Leiden University Medical Centre, Leiden, the Netherlands

⁴ Department of Psychiatry, CAPRI-University of Antwerp, Antwerp, Belgium

⁵Department of Gerontology and Geriatrics, Leiden University Medical Centre, Leiden, the Netherlands

†Dr. Anton J.M. de Craen passed away in January 2016.

Running head: Validity of the GDS-3A in measuring apathy

Key words: apathy; Geriatric Depression Scale – 3A (GDS-3A); validity; Apathy Scale;
older persons

Key points:

- The Geriatric Depression Scale (GDS)-3A, a subset of the GDS-15, is increasingly used in research to measure apathy.

- The GDS-3A has a low sensitivity and high specificity as a measure of apathy in two populations of older persons.
- Using the GDS-3A in research might yield estimates biased towards the null in case of non-differential misclassification.

Funding: The DANTE Study Leiden was funded by a grant from ZonMw, the Netherlands Organization for Health Research and Development, Program Priority Medicines for the Elderly, grant number 40-41600-98-9014. The PROMODE Study was funded by a grant from ZonMw, the Netherlands Organisation for Health Research and Development, grant number 945-07-502; and a grant from Rivierduinen, Centre for mental health care.

Corresponding author: Anne Suzanne Bertens, Leiden University Medical Centre, Department of Radiology, C-02-S, P.O. Box 9600, 2300 RC Leiden, The Netherlands. Telephone: +31 71 5265948. Fax: +31 71 5248256. E-mail: a.s.bertens@lumc.nl

Word count body text: 3473

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

ABSTRACT

Objective The Geriatric Depression Scale(GDS)-3A, a three-item subset of the GDS-15, is increasingly used as a measure for apathy in research settings to assess factors associating with this neuropsychiatric syndrome. We aimed to assess how accurately the GDS-3A discriminates between presence and absence of apathy in two populations of community-dwelling older persons, using the Apathy Scale as reference standard.

Methods Baseline data were used from 427 participants of the Discontinuation of Antihypertensive Treatment in Elderly people (DANTE) Study Leiden and 1118 participants of the PROactive Management Of Depression in the Elderly (PROMODE) Study, all ≥ 75 years and with available GDS-3A and Apathy Scale measurements. A cut-off score of ≥ 14 was used for presence of apathy according to the Apathy Scale. Areas under the receiver operating characteristic curve (AUC) were calculated. Based on the likelihood ratios for GDS-3A scores, a cut-off of ≥ 2 was used for presence of apathy according to the GDS-3A to calculate test characteristics.

Results The AUC was 0.68 (95% confidence interval 0.62-0.73) in the DANTE Study and 0.72 (0.67-0.77) in the PROMODE Study. In the DANTE Study sensitivity was 29.3% (21.4-38.1) and specificity was 88.5% (84.4-91.8), whereas in the PROMODE Study sensitivity was 32.8% (24.5-41.1) and specificity 92.6% (90.9-94.2). Stratification on population characteristics did not yield more favourable test characteristics.

Conclusion The GDS-3A has low sensitivity and high specificity as a measure of apathy in two populations of older persons. Using the GDS-3A in research might yield estimates biased towards the null in case of non-differential misclassification.

61 INTRODUCTION

62 Apathy is an important yet often overlooked neuropsychiatric behavioural syndrome that is
63 common in older persons (Lyketsos *et al.*, 2002; van Reekum *et al.*, 2005). Apathy is
64 characterized by diminished motivation and initiative, reduced goal-directed behaviour, loss
65 of interest and emotional indifference (Levy and Dubois, 2006; Marin, 1991). Its presence in
66 older persons is associated with worse cognitive functioning (Masters *et al.*, 2015; Richard *et*
67 *al.*, 2012), reduced therapeutic response (van Reekum *et al.*, 2005), lower quality of life
68 (Gerritsen *et al.*, 2005; Groeneweg-Koolhoven *et al.*, 2014), and high caregiver distress
69 (Thomas *et al.*, 2001). While symptoms of apathy can occur as part of depression, apathy is
70 increasingly being recognised as a syndrome in its own right, also in the absence of a
71 depressed mood (Levy *et al.*, 1998; Marin, 1991; Robert *et al.*, 2009). No valid, easily
72 applicable screening tool for apathy is available for use in general clinical practice (Clarke *et*
73 *al.*, 2011; Radakovic *et al.*, 2015). Because of its association with adverse health outcomes,
74 apathy is increasingly prioritised on research agendas, with subsequent expanding knowledge
75 on its specific prognostic (Ligthart *et al.*, 2012; Van der Mast *et al.*, 2008) and possibly
76 causal factors (Benoit and Robert, 2011; Grool *et al.*, 2014; Stella *et al.*, 2014).
77
78 Data from large observational studies could be particularly useful in identifying potentially
79 modifiable risk factors for apathy at old age. However, only few studies have used specific
80 instruments to prospectively collect data on symptoms of apathy, whereas symptoms of
81 depression are often measured (Harris *et al.*, 2007; Ligthart *et al.*, 2012). To screen for
82 symptoms of depression at old age, the Geriatric Depression Scale (GDS)-15 is frequently
83 used in clinical practice as well as in research, showing a good reliability and validity (D'Ath
84 *et al.*, 1994). In factor analyses, a subset of three GDS-15 items has repeatedly been
85 identified as a cluster of symptoms that assesses apathy (Adams, 2001; Adams *et al.*, 2004;

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

86 Kim *et al.*, 2013). This GDS-3-apathy subscale (GDS-3A) comprises the items: (1) Have you
87 dropped many of your activities and interest?; (2) Do you prefer to stay at home, rather than
88 going out and doing new things?; and (3) Do you feel full of energy? The GDS-3A subscale
89 is increasingly being used in research to identify participants with apathy in studying
90 associating factors (Grool *et al.*, 2014; Ligthart *et al.*, 2012; Van der Mast *et al.*, 2008).
91 However, only limited evidence exists for the discriminatory value of the GDS-3A for the
92 presence or absence of clinically relevant apathy. Van der Mast *et al.* compared the GDS-3A
93 with the Apathy Scale in a sample of community-dwelling 90-year-olds, rendering a
94 sensitivity of 68.6%, a specificity of 84.9%, a positive predictive value of 77.8%, and a
95 negative predictive value of 77.8% (Van der Mast *et al.*, 2008). To the best of our knowledge,
96 this is the only study providing epidemiological test characteristics for the GDS-3A and
97 therefore it is yet undetermined whether these discriminatory qualities of the GDS-3A also
98 hold for other populations of older persons.
99
100 In the Discontinuation of Antihypertensive Treatment in Elderly people (DANTE) Study
101 Leiden and the Proactive Management of Depression in the Elderly (PROMODE) Study both
102 the GDS-3A and the Apathy Scale were assessed as part of a neuropsychological evaluation,
103 providing the unique opportunity to compare these questionnaires in two large cohorts of
104 older persons. Therefore, the aim of conducting the current study was to assess how
105 accurately the GDS-3A discriminates between presence and absence of apathy in two
106 populations of community-dwelling older persons, compared to the Apathy Scale.

METHODS

Study populations

We used baseline data of two Dutch randomised controlled trials, the DANTE Study and the PROMODE Study.

Discontinuation of Antihypertensive Treatment in Elderly people (DANTE) Study Leiden

The primary aim of the DANTE trial was to assess whether discontinuation of antihypertensive treatment in older persons with mild cognitive deficits improves cognitive, psychological and general daily functioning (Moonen *et al.*, 2015). Participants aged ≥ 75 years, using antihypertensive medication, with a current systolic blood pressure ≤ 160 mmHg, without serious cardiovascular disease, and without a diagnosis of dementia were recruited from primary care practices between May 2011 and July 2013. Of the 5537 selected older persons, 2002 consented to participate in a Mini Mental State Examination screening. A total of 1301 persons did not meet the MMSE selection criterion (MMSE score of 21-27), 67 did not meet other selection criteria, 204 declined to participate, and 3 had missing data on the Apathy Scale and/or the GDS-15, leaving 427 participants for further analyses.

PROactive Management Of Depression in the Elderly (PROMODE) Study

The primary aim of the PROMODE trial was to investigate the (cost-) effectiveness of a stepped-care intervention programme among older persons with depressive symptoms (van der Weele *et al.*, 2012). A total of 2759 participants aged 75 years and above were recruited from primary care practices between April 2007 and July 2008. A total of 366 persons were excluded because of a MMSE score of less than 19 points, a limited life expectancy, recent loss of partner, a diagnosis of dementia or a current treatment for depression. Of the remaining 2393 persons invited to participate, 1054 were non-responders, 101 did not meet

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

132 selection criteria and 120 persons had inadequate assessment or missing data of the Apathy
133 Scale, leaving 1118 participants for analyses (Groeneweg-Koolhoven *et al.*, 2014).
134
135 The medical ethics committee of the Leiden University Medical Center approved both the
136 DANTE and PROMODE Study, and informed consent was obtained from all participants. In
137 the DANTE Study all participants had mild cognitive deficits and therefore gave informed
138 consent after written and verbal description of the study was given in the presence of a close
139 relative serving as a proxy decision maker (Van Rookhuijzen *et al.*, 2014).

140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156

Measurements

Geriatric Depression Scale (GDS)

In both studies, the GDS-15 (Yesavage *et al.*, 1982) was administered at baseline by trained research personnel to assess presence of depressive symptoms within the last few weeks. The GDS-15 is a short version of the GDS-30 (Yesavage *et al.*, 1982), and shows a good reliability and validity (D'Ath *et al.*, 1994). The scale consists of 15 items that can be scored as present or absent (range 0-15 point, with higher scores indicating more symptoms of depression). A score of five or higher is indicative of clinically relevant depressive symptoms (D'Ath *et al.*, 1994). The apathy subscale (GDS-3A) of the GDS-15 consists of the following 3 items (score range 0-3 points): (1)Have you dropped many of your activities and interest?; (2)Do you prefer to stay at home, rather than going out and doing new things?; and (3)Do you feel full of energy?

157 *Apathy Scale*

158 In both studies trained research personnel assessed the Apathy Scale to record presence of
159 symptoms of apathy. The Apathy Scale, a semi-structured interview combining input from
160 participants, proxy and clinical impression, is an abbreviated version of the Apathy
161 Evaluation Scale (Marin, 1991), and has a good one-week test-retest reliability, inter-rater
162 reliability and internal validity (Starkstein *et al.*, 1992). The Apathy Scale consists of 14
163 items that are scored on a four-point Likert scale (range 0-42 points, with higher scores
164 indicating more symptoms of apathy). A score of at least 14 points is considered to be
165 indicative for the presence of clinically relevant apathy (Starkstein *et al.*, 1992). The research
166 personnel was not blinded for the GDS-15 scores.

168 *Additional measurements*

169 In both studies socio-demographic characteristics were assessed at baseline using
170 standardized interviews. Level of education was dichotomized at primary education (six years
171 of schooling) and use of alcohol was dichotomized at 14 consumptions per week. Global
172 cognitive functioning was assessed with the MMSE (score range 0-30 points, with higher
173 scores indicating better cognitive function) (Folstein *et al.*, 1975). In the DANTE Study,
174 presence of cardiovascular disease was obtained from the general practitioners using
175 structured questionnaires, and defined as myocardial infarction or coronary reperfusion
176 procedure longer than three years ago, and/or peripheral arterial disease, as persons with
177 serious or recent cardiovascular disease (such as a history of stroke, transient ischemic attack
178 or heart failure and/or a myocardial infarction/coronary reperfusion procedure within the last
179 three years) were excluded from participation. No information regarding presence of
180 cardiovascular disease was available in the PROMODE Study.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

182 **Statistical analyses**

183 Characteristics of the study populations are presented as mean (standard deviation (SD)),
184 number (%) or median (interquartile range (IQR)) when appropriate. In a receiver operating
185 characteristic (ROC) curve the true positive rate (sensitivity) was plotted against the false
186 positive rate (100-specificity) for different cut-off points of the GDS-3A, and the area under
187 the curve was calculated. For each score on the GDS-3A, likelihood ratios were calculated:
188 the proportion of participants with a specific test score in the presence of apathy was divided
189 by the proportion of participants without apathy with that same test score. A two by two
190 contingency table was created to determine the discriminatory accuracy of the GDS-3A using
191 a cut-off of two or more points in assessing presence of apathy (according to ≥ 14 on the
192 Apathy Scale). Sensitivity and specificity were calculated, as well as likelihood ratios for a
193 positive (LR+) and negative (LR-) test using the following formulas: $LR+ = \text{sensitivity}/(1 -$
194 $\text{specificity})$ and $LR- = (1 - \text{sensitivity})/\text{specificity}$.
195
196 Spearman's correlation coefficients were computed for the GDS-3A and the Apathy Scale.
197 To assess the influence of each item on the correlation of the GDS-3A with the Apathy Scale,
198 these analyses were repeated after omitting either the first, second or third item of the GDS-
199 3A. Furthermore, to assess whether discriminatory accuracy of the GDS-3A depended on
200 population characteristics, we performed analyses in strata of age (of 5 years from 75 years
201 onwards), gender, cognitive function (dichotomised at the median MMSE score), level of
202 education (dichotomised at 6 years), and presence of cardiovascular disease (yes or no; only
203 in the DANTE Study). Data were analysed using SPSS, version 22.0 and, Stata, version 12.0.

RESULTS

In Table 1 the characteristics are shown of the 427 participants of the DANTE Study (mean age 81.3 (SD 4.6)) and the 1118 participants of the PROMODE Study (mean age 81.8 (SD 4.9)), with mean Apathy Scale scores of 11.3 (SD 4.7) and 7.5 (SD 4.6), respectively.

Presence of apathy according to a score of ≥ 14 points on the Apathy Scale was 28.8% in the DANTE Study and 10.9% in the PROMODE Study. INSERT TABLE 1

Figure 1 presents the ROC curves, with areas under the ROC curves of 0.68(95% confidence interval 0.62–0.73) in the DANTE Study and 0.72(0.67–0.77) in the PROMODE Study.

INSERT FIGURE 1

In Table 2 the likelihood ratios for the individual GDS-3A scores are shown. In the DANTE Study, the likelihood ratio was 1.96(1.18–3.25) for a score of two and 5.36(2.08–13.8) for a score of three on the GDS-3A, while this was 4.32(3.02–6.18) and 5.44(1.56–19.0) respectively in the PROMODE Study. INSERT TABLE 2.

Since the percentage of participants scoring three points on the GDS-3A was low in both studies (n=19 for the DANTE Study and n=10 for the PROMODE Study), a cut-off score of ≥ 2 was used for the calculation of sensitivity and specificity (Table 3). The GDS-3A had a sensitivity of 29.3% (21.4–38.1) and a specificity of 88.5% (84.4–91.8) in the DANTE Study, and 32.8% (24.5–41.1) and 92.6% (90.9–94.2), respectively, in the PROMODE Study.

Correlation coefficients between the GDS-3A and the Apathy Scale were 0.42 in the DANTE Study and 0.39 in the PROMODE Study (both p-values<0.001). Coefficients remained largely similar after omitting either the first, second or third item of the GDS-3A. Stratified analyses according to age, gender, cognitive function, presence of cardiovascular disease, and level of education rendered largely similar test characteristics in both studies (data not shown). INSERT TABLE 3.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

229 **DISCUSSION**

230 In our study investigating the applicability of the GDS-3A in research settings, the GDS-3A
231 only moderately discriminated between presence and absence of clinically relevant apathy in
232 two populations of older persons, when using the Apathy Scale as reference standard. Using a
233 cut-off of ≥ 2 for presence of apathy according to the GDS-3A, sensitivity compared to the
234 Apathy Scale was 29.3% in the DANTE Study and 32.8% in the PROMODE study, whereas
235 in both studies specificity was high (88.5% and 92.6%, respectively).

236

237 To the best of our knowledge, our study is the second to report epidemiological test
238 characteristics for the GDS-3A, and the first to do so in two large cohorts of older persons
239 with a wider age range. The likelihood ratios for the GDS-3A scores increased with
240 increasing test scores in both studies, but poorly discriminated for scores of one or two (1.53
241 and 1.96 in the DANTE Study, respectively, and 1.56 and 4.32 in the PROMODE Study).
242 Although a score of three had a moderately high likelihood ratio in both studies (5.36 in the
243 DANTE Study and 5.44 in the PROMODE Study), the number of participants in this
244 category was very low in both studies, limiting power for analyses. Therefore we decided to
245 use a cut-off of ≥ 2 for calculations of test characteristics in these populations, with modest
246 corresponding likelihood ratios because of a low sensitivity and high specificity. In both the
247 DANTE Study and the PROMODE Study the GDS-3A scores moderately but significantly
248 correlated with Apathy Scale scores and the correlation did not depend on the performance of
249 a single item.

250

251 A possible explanation for the moderate performance of the GDS-3A is that, although
252 interviewer-administered, it records self-reported symptoms. The Apathy Scale however, is a
253 semi-structured interview allowing the interviewer to incorporate his or her own clinical

judgement and information obtained from proxies. Since disease awareness might be low in apathy (van Reekum *et al.*, 2005), this difference might explain the low sensitivity of the GDS-3A. Although lack of self-awareness might at least partly explain the low sensitivity of the GDS-3A, our findings in the DANTE Study and the PROMODE Study contrasts with findings of Van der Mast *et al.* Using data from the Leiden 85-plus Study, they reported a sensitivity of 68.6% using the same cut-off values for both the GDS-3A and the Apathy Scale (Van der Mast *et al.*, 2008). The difference in test characteristics might be explained by a higher prevalence of apathy according to the Apathy Scale in the Leiden 85-plus study, which was present in 51 out of 117 participants (43.6%) and might be due to the much higher age in the Leiden 85-plus Study. It is increasingly recognized that also sensitivity and specificity may vary across patient populations (Leeflang *et al.*, 2013; Moons and Harrell, 2003). For example with a lower disease prevalence, there may be more patients with less severe symptoms, and sensitivity can be lower (Leeflang *et al.*, 2009).

In contrast to the population-based Leiden 85-plus Study, participants from both our studies were selected for participation in clinical trials, were on average younger, had better cognitive function, higher level of education and better cardiovascular health (the latter information only available in the DANTE Leiden Study). However, stratified analyses showed that aforementioned characteristics did not explain the differences in findings between our current study and the Leiden 85-plus Study.

Besides other, currently unmeasured, patient characteristics, differences in diagnostic accuracy might also be artificially caused by our use of the Apathy Scale, an imperfect reference standard (Biesheuvel *et al.*, 2007; Radakovic *et al.*, 2015). The Apathy Scale was developed in patients with Parkinson's disease (Starkstein *et al.*, 1992) and showed good interrater reliability ($r=0.81$), test-retest reliability ($r=0.90$), and good internal consistency (Cronbach's $\alpha=0.76$). Test characteristics for a cut-off of ≥ 14 compared to a clinical

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

279 diagnosis of apathy were calculated in 12 patients (sensitivity 66%, specificity 100%). The
280 only other study on psychometric properties of the Apathy Scale reported fair internal
281 consistency (Cronbach's $\alpha=0.69$) (Pedersen *et al.*, 2012). Although data on the performance
282 of the Apathy Scale in other populations is scarce, the questionnaire is derived from the well-
283 validated Apathy Evaluation Scale (AES) (Marin *et al.*, 1991; Radakovic *et al.*, 2015). The
284 Apathy Scale is shorter and has the possibility to combine information from the patient,
285 informant and clinician, making it a more favourable instrument. Even so, these imperfect
286 characteristics of our reference standard may underlie the moderate performance of the GDS-
287 3A in our study. However, it must be emphasized that the same reference standard was used
288 in the Leiden 85-plus Study, making results comparable for the purpose of this study.
289 Moreover, if reference standard misclassification would explain differences in sensitivity and
290 specificity, a pattern of higher sensitivity and lower specificity with increasing disease
291 prevalence would be observed (Biesheuvel *et al.*, 2007). As this was not consistently found
292 for the DANTE Study, PROMODE Study and Leiden 85-plus Study, it is less likely that
293 reference standard misclassification explains the variation in sensitivity and specificity across
294 the different studies.
295
296 Strengths of this study include the use of two relatively large sized, well-defined samples of
297 older persons and very few missing data on the GDS-3A and Apathy Scale, implying a low
298 risk of validation bias. Furthermore, the research protocols for the DANTE Study and the
299 PROMODE Study were designed similarly with regard to administering the questionnaires,
300 which contributes to comparability of results. However, there are several limitations of this
301 study that need to be taken into account when interpreting the results. First, it is important to
302 state that the current study was not designed to assess added diagnostic value, but is a test
303 accuracy study aiming to investigate the applicability of the GDS-3A in research settings.

Second, only 19 (4.4%, DANTE Study) and 10 (0.89%, PROMODE Study) of the participants had GDS-3A scores of three. Although numbers were too low to calculate valid diagnostic test characteristics for this cut-off in our study populations, increasing the cut-off for apathy might yield higher specificity but even lower sensitivity. This may be favoured in certain research settings, depending on the aim of the study. Third, the interviewers administering the Apathy Scale were not formally blinded for the GDS-3A scores, which could have led to information bias. Nonetheless, since the GDS-3A questions were incorporated in the GDS-15 questionnaire and our current aim was not the primary aim of the studies, we do not expect different results with a blinded study design. Furthermore, apathy can occur as a symptom of depression and as a syndrome in its own right (Levy *et al.*, 1998). As no formal diagnosis of depression was available in either study, we were not able to differentiate between apathy and depression. However, the GDS-3A items were identified as measuring apathy by several studies on construct validity (Adams, 2001; Adams *et al.*, 2004; Kim *et al.*, 2013) as well as an expert panel installed by Van der Mast *et al.* (Van der Mast *et al.*, 2008). We therefore deem it justified using these items as a measure for apathy. Last, both the DANTE and PROMODE participants were selected for clinical trials, limiting generalizability of the study results.

The main premise of this study was to assess the applicability of the GDS-3A in research settings. Given the moderate likelihood ratios in both the DANTE and PROMODE Study, our results suggest that the GDS-3A is not a useful tool in clinical practice to screen for presence of apathy. However, the GDS-3A may still be a useful scale in research, depending on the aim of the study. Because of its low sensitivity regardless of the cut-off used, the GDS-3A will not be suitable to determine the prevalence of apathy in specific populations. In studies focussing on potential risk and prognostic factors however, it can still be a useful

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

instrument as long as it can be assumed that the misclassification is non-differential. If so, effect sizes will be biased towards the null, both when the GDS-3A is used as a measure of outcome and determinant (Copeland *et al.*, 1977). Furthermore, the higher the prevalence of apathy in a population, the smaller the bias will be (Copeland *et al.*, 1977), making the GDS-3A better suited for research in older populations.

In conclusion, our results from two large study cohorts show that the GDS-3A is only very moderately accurate in discriminating between presence and absence of clinically relevant apathy in older persons. Although we think it is therefore a less favourable screening tool for clinical practice, the GDS-3A can still be useful in research if no other measurement of apathy is available. Because of the non-differential misclassification, and thus dilution of effect, the GDS-3A is preferably used in large studies. Especially for studies among older people, with long follow-up, the GDS-3A might be attractive to study risk factors and prognostic factors for apathy, since it is unlikely that these earlier studies used a validated questionnaire. However, since it becomes increasingly clear that apathy is an important and useful endpoint in studies among older people, future studies should use a more valid instrument. In all instances, caution has to be taken in interpreting negative findings as evidence for absence of an association.

ACKNOWLEDGEMENTS

The authors thank all the DANTE Study Leiden and PROMODE Study co-investigators.

CONFLICT OF INTEREST DISCLOSURES

We report no conflict of interest for all authors.

AUTHOR'S CONTRIBUTIONS

Bertens and Moonen: both authors contributed substantially to data acquisition of the DANTE Study, analysis and interpretation of the data, drafting of the manuscript and to revising the manuscript critically for important intellectual content.

De Waal: substantially contributed to the study concept and design of the PROMODE Study, analysis and interpretation of the data, and to revising the manuscript critically for important intellectual content.

Foster-Dingley: substantially contributed to data acquisition of the DANTE Study, analysis and interpretation of the data, and to revising the manuscript critically for important intellectual content.

De Ruijter: substantially contributed to the study concept and design of the DANTE Study, analysis and interpretation of the data, and to revising the manuscript critically for important intellectual content.

Gussekloo, Van der Mast: substantially contributed to the study concept and design of the DANTE Study and the PROMODE Study, analysis and interpretation of the data, and to revising the manuscript critically for important intellectual content.

De Craen: substantially contributed to the study concept and design of the DANTE Study, analysis and interpretation of the data, and to revising the manuscript critically for important intellectual content. Dr. Van der Mast takes responsibility for the integrity of this work as a whole, from inception to published article.

All authors contributed to and approved of the final manuscript.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

379 **SPONSOR’S ROLE**

380 The funders had no role the study’s design, in the collection, analysis and interpretation of the
381 data, in the writing of the manuscript, nor in the decision to submit the manuscript for
382 publication.

383

384

385

For Peer Review

REFERENCES

- Adams KB. 2001. Depressive symptoms, depletion, or developmental change? Withdrawal, apathy, and lack of vigor in the Geriatric Depression Scale. *Gerontologist* **41**: 768-777.
- Adams KB, Matto HC and Sanders S. 2004. Confirmatory factor analysis of the geriatric depression scale. *Gerontologist* **44**: 818-826.
- Benoit M and Robert PH. 2011. Imaging correlates of apathy and depression in Parkinson's disease. *J Neurol Sci* **310**: 58-60. DOI 10.1016/j.jns.2011.07.006
- Biesheuvel C, Irwig L and Bossuyt P. 2007. Observed differences in diagnostic test accuracy between patient subgroups: is it real or due to reference standard misclassification? *Clin Chem* **53**: 1725-1729. DOI 10.1373/clinchem.2007.087403
- Clarke DE, Ko JY, Kuhl EA, van Reekum R, Salvador R and Marin RS. 2011. Are the available apathy measures reliable and valid? A review of the psychometric evidence. *J Psychosom Res* **70**: 73-97. DOI 10.1016/j.jpsychores.2010.01.012
- Copeland KT, Checkoway H, McMichael AJ and Holbrook RH. 1977. Bias due to misclassification in the estimation of relative risk. *Am J Epidemiol* **105**: 488-495.
- D'Ath P, Katona P, Mullan E, Evans S and Katona C. 1994. Screening, detection and management of depression in elderly primary care attenders. I: The acceptability and performance of the 15 item Geriatric Depression Scale (GDS15) and the development of short versions. *Fam Pract* **11**: 260-266.
- Folstein MF, Folstein SE and McHugh PR. 1975. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res* **12**: 189-198.
- Gerritsen DL, Jongenelis K, Steverink N, Ooms ME and Ribbe MW. 2005. Down and drowsy? Do apathetic nursing home residents experience low quality of life? *Aging Ment Health* **9**: 135-141. DOI 10.1080/13607860412331336797

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

410 Groeneweg-Koolhoven I, de Waal MW, van der Weele GM, Gussekloo J and van der Mast
411 RC. 2014. Quality of life in community-dwelling older persons with apathy. *Am J Geriatr*
412 *Psychiatry* **22**: 186-194. DOI 10.1016/j.jagp.2012.10.024
413 Grool AM, Geerlings MI, Sigurdsson S, Eiriksdottir G, Jonsson PV, Garcia ME,
414 Siggeirsdottir K, Harris TB, Sigmundsson T, Gudnason V and Launer LJ. 2014. Structural
415 MRI correlates of apathy symptoms in older persons without dementia: AGES-Reykjavik
416 Study. *Neurology* **82**: 1628-1635. DOI 10.1212/WNL.0000000000000378
417 Harris TB, Launer LJ, Eiriksdottir G, Kjartansson O, Jonsson PV, Sigurdsson G,
418 Thorgeirsson G, Aspelund T, Garcia ME, Cotch MF, Hoffman HJ and Gudnason V. 2007.
419 Age, Gene/Environment Susceptibility-Reykjavik Study: multidisciplinary applied
420 phenomics. *Am J Epidemiol* **165**: 1076-1087. DOI 10.1093/aje/kwk115
421 Kim G, DeCoster J, Huang CH and Bryant AN. 2013. A meta-analysis of the factor structure
422 of the Geriatric Depression Scale (GDS): the effects of language. *Int Psychogeriatr* **25**: 71-
423 81. DOI 10.1017/s1041610212001421
424 Leeflang MM, Bossuyt PM and Irwig L. 2009. Diagnostic test accuracy may vary with
425 prevalence: implications for evidence-based diagnosis. *J Clin Epidemiol* **62**: 5-12. DOI
426 10.1016/j.jclinepi.2008.04.007
427 Leeflang MM, Rutjes AW, Reitsma JB, Hooft L and Bossuyt PM. 2013. Variation of a test's
428 sensitivity and specificity with disease prevalence. *Cmaj* **185**: E537-544. DOI
429 10.1503/cmaj.121286
430 Levy ML, Cummings JL, Fairbanks LA, Masterman D, Miller BL, Craig AH, Paulsen JS and
431 Litvan I. 1998. Apathy is not depression. *J Neuropsychiatry Clin Neurosci* **10**: 314-319.
432 Levy R and Dubois B. 2006. Apathy and the functional anatomy of the prefrontal cortex-
433 basal ganglia circuits. *Cereb Cortex* **16**: 916-928. DOI 10.1093/cercor/bhj043

- 434 Ligthart SA, Richard E, Fransen NL, Eurelings LS, Beem L, Eikelenboom P, van Gool WA
435 and Moll van Charante EP. 2012. Association of vascular factors with apathy in community-
436 dwelling elderly individuals. *Arch Gen Psychiatry* **69**: 636-642. DOI
437 10.1001/archgenpsychiatry.2011.1858
- 438 Lyketsos CG, Lopez O, Jones B, Fitzpatrick AL, Breitner J and DeKosky S. 2002. Prevalence
439 of neuropsychiatric symptoms in dementia and mild cognitive impairment: results from the
440 cardiovascular health study. *JAMA* **288**: 1475-1483.
- 441 Marin RS. 1991. Apathy: a neuropsychiatric syndrome. *J Neuropsychiatry Clin Neurosci* **3**:
442 243-254.
- 443 Marin RS, Biedrzycki RC and Firinciogullari S. 1991. Reliability and validity of the Apathy
444 Evaluation Scale. *Psychiatry Res* **38**: 143-162.
- 445 Masters MC, Morris JC and Roe CM. 2015. "Noncognitive" symptoms of early Alzheimer
446 disease: A longitudinal analysis. *Neurology* **84**: 617-622. DOI
447 10.1212/wnl.0000000000001238
- 448 Moonen JEF, Foster-Dingley JC, De Ruijter W, Van der Grond J, Bertens AS, Van Buchem
449 MA, Gussekloo J, Middelkoop HA, Wermer MJH, Westendorp RGJ, De Craen AJM and Van
450 der Mast RC. 2015. Effects of the Discontinuation of Antihypertensive Treatment in Elderly
451 People Study on cognitive functioning. The DANTE Study Leiden. A Randomized clinical
452 trial. *JAMA Internal Medicine* **175**: 1622-1630. DOI 10.1001/jamainternmed.2015.4103
- 453 Moons KGM and Harrell FE. 2003. Sensitivity and Specificity should be De-emphasized in
454 Diagnostic Accuracy Studies. *Academic Radiology* **10**: 670-672. DOI 10.1016/s1076-
455 6332(03)80087-9
- 456 Pedersen KF, Alves G, Larsen JP, Tysnes OB, Moller SG and Bronnick K. 2012.
457 Psychometric properties of the Starkstein Apathy Scale in patients with early untreated

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

458 Parkinson disease. *Am J Geriatr Psychiatry* **20**: 142-148. DOI
459 10.1097/JGP.0b013e31823038f2

460 Radakovic R, Harley C, Abrahams S and Starr JM. 2015. A systematic review of the validity
461 and reliability of apathy scales in neurodegenerative conditions. *Int Psychogeriatr* **27**: 903-
462 923. DOI 10.1017/S1041610214002221

463 Richard E, Schmand B, Eikelenboom P, Yang SC, Ligthart SA, Moll van Charante EP and
464 van Gool WA. 2012. Symptoms of apathy are associated with progression from mild
465 cognitive impairment to Alzheimer's disease in non-depressed subjects. *Dement Geriatr Cogn*
466 *Disord* **33**: 204-209. DOI 10.1159/000338239

467 Robert P, Onyike CU, Leentjens AF, Dujardin K, Aalten P, Starkstein S, Verhey FR,
468 Yessavage J, Clement JP, Drapier D, Bayle F, Benoit M, Boyer P, Lorca PM, Thibaut F,
469 Gauthier S, Grossberg G, Vellas B and Byrne J. 2009. Proposed diagnostic criteria for apathy
470 in Alzheimer's disease and other neuropsychiatric disorders. *Eur Psychiatry* **24**: 98-104. DOI
471 10.1016/j.eurpsy.2008.09.001

472 Starkstein SE, Mayberg HS, Preziosi TJ, Andrezejewski P, Leiguarda R and Robinson RG.
473 1992. Reliability, validity, and clinical correlates of apathy in Parkinson's disease. *J*
474 *Neuropsychiatry Clin Neurosci* **4**: 134-139.

475 Stella F, Radanovic M, Aprahamian I, Canineu PR, de Andrade LP and Forlenza OV. 2014.
476 Neurobiological correlates of apathy in Alzheimer's disease and mild cognitive impairment: a
477 critical review. *J Alzheimers Dis* **39**: 633-648. DOI 10.3233/JAD-131385

478 Thomas P, Clement JP, Hazif-Thomas C and Leger JM. 2001. Family, Alzheimer's disease
479 and negative symptoms. *Int J Geriatr Psychiatry* **16**: 192-202.

480 Van der Mast RC, Vinkers DJ, Stek ML, Bek MC, Westendorp RG, Gussekloo J and De
481 Craen AJ. 2008. Vascular disease and apathy in old age. The Leiden 85-Plus Study. *Int J*
482 *Geriatr Psychiatry* **23**: 266-271. DOI 10.1002/gps.1872

- 483 van der Weele GM, de Waal MW, van den Hout WB, de Craen AJ, Spinhoven P, Stijnen T,
484 Assendelft WJ, van der Mast RC and Gussekloo J. 2012. Effects of a stepped-care
485 intervention programme among older subjects who screened positive for depressive
486 symptoms in general practice: the PROMODE randomised controlled trial. *Age Ageing* **41**:
487 482-488. DOI 10.1093/ageing/afs027
- 488 van Reekum R, Stuss DT and Ostrander L. 2005. Apathy: why care? *J Neuropsychiatry Clin*
489 *Neurosci* **17**: 7-19. DOI 10.1176/appi.neuropsych.17.1.7
- 490 Van Rookhuijzen AE, Touwen DP, De Ruijter W, Engberts DP and Van der Mast RC. 2014.
491 Deliberating clinical research with cognitively impaired older people and their relatives: an
492 ethical add-on study to the protocol "Effects of Temporary Discontinuation of
493 Antihypertensive Treatment in the Elderly (DANTE) with Cognitive Impairment". *Am J*
494 *Geriatr Psychiatry* **22**: 1233-1240. DOI 10.1016/j.jagp.2013.04.005
- 495 Yesavage JA, Brink TL, Rose TL, Lum O, Huang V, Adey M and Leirer VO. 1982.
496 Development and validation of a geriatric depression screening scale: a preliminary report. *J*
497 *Psychiatr Res* **17**: 37-49.

Table 1. Characteristics of the DANTE Study and the PROMODE Study

	DANTE	PROMODE
Number of participants	427	1118
Age, years (mean, SD)	81.3 (4.6)	81.8 (4.9)
Female gender (n, %)	257 (60.2%)	684 (61.1)
Presence of CVD (n, %) ^a	48 (11.2%)	- ^b
MMSE score (median, IQR)	26 (25 - 27)	28 (27 - 29)
Lower level of education (n, %) ^c	142 (33.3%)	333 (29.8%)
Apathy Scale (mean, SD)	11.3 (4.7)	7.5 (4.6)
Apathy according to Apathy Scale (n, %) ^d	123 (28.8%)	122 (10.9%)
Apathy Scale score in those with apathy (mean, SD)	17.2 (3.3)	16.6 (2.8)
Scores on the GDS-3A (n, %)		
0	225 (52.8%)	718 (64.2%)
1	131 (30.7%)	286 (25.6%)
2	52 (12.2%)	104 (9.3%)
3	19 (4.4%)	10 (0.89%)
Depressive symptoms present (n, %) ^e	45 (10.5%)	64 (5.7%)

SD, standard deviation; CVD, cardiovascular disease; MMSE, Mini Mental State Examination; GDS, Geriatric Depression Scale

a: Cardiovascular diseases comprise myocardial infarction or coronary intervention > 3 years ago, or presence of peripheral artery disease

b: No data on cardiovascular morbidity available in the PROMODE Study

c: Level of education is dichotomized at 6 years

d: Apathy according to the Apathy Scale: score of ≥ 14

e: Depressive symptoms present according to the GDS-15: score of ≥ 5

Table 2. Performance of the GDS-3A in the DANTE Study and the PROMODE Study

DANTE Study (n=427)				
GDS-3A score	N	Apathy present ^a	Apathy absent	Likelihood ratio for GDS-3A score
3	19	13	6	5.36 (2.08 - 13.8)
2	52	23	29	1.96 (1.18 - 3.25)
1	131	50	81	1.53 (1.15 - 2.03)
0	225	37	188	0.49 (0.37 - 0.65)
Total	427	123	304	

PROMODE Study (n=1118)				
Cut-off point	N	Apathy present ^a	Apathy absent	Likelihood ratio for GDS-3A score
3	10	4	6	5.44 (1.56 - 19.0)
2	104	36	68	4.32 (3.02 - 6.18)
1	286	46	240	1.56 (1.21 - 2.02)
0	718	36	682	0.43 (0.33 - 0.57)
Total	1118	122	996	

Data are presented as numbers, or as likelihood ratios with 95% confidence intervals.

GDS, Geriatric Depression Scale

a: Apathy present: Apathy Scale score ≥ 14

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Table 3. Presence of apathy according to the GDS-3A and the Apathy Scale in the DANTE Study and the PROMODE Study

DANTE		Apathy Scale ≥ 14		
		Positive	Negative	
GDS-3A ≥ 2	Positive	36	35	71
	Negative	87	269	356
		123	304	427
Sensitivity		29.3% (21.4 - 38.1)		
Specificity		88.5% (84.4 - 91.8)		
LR+		2.54 (1.68 - 3.85)		
LR-		0.80 (0.71 - 0.90)		
PROMODE		Apathy Scale ≥ 14		
		Positive	Negative	
GDS-3A ≥ 2	Positive	40	74	114
	Negative	82	922	1004
		122	996	1118
Sensitivity		32.8% (24.5 - 41.1)		
Specificity		92.6% (90.9 - 94.2)		
LR+		4.41 (3.15 - 6.17)		
LR-		0.73 (0.64 - 0.82)		

Data are presented as numbers, percentages with 95% confidence intervals, or likelihood ratios with 95% confidence intervals.

GDS, Geriatric Depression Scale; LR, likelihood ratio

504
505

Figure 1. Receiver Operating Characteristic (ROC) curve for GDS-3A compared to the Apathy Scale in the DANTE Study and PROMODE Study

332x507mm (200 x 200 DPI)