Table S1: The results of two one-sided tests (TOST) null-hypothesis and equivalence tests of different short GDS versions against the established 15-item GDS according to Sheikh. The most accurate short versions 4-item D’Ath, 7-item Broekman and 10-item Van Marwijk were also compared against each other.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | GDS Mean ± SD | | | Null-hypothesis test | | | | Equivalence test | | |
|  | MDD  N=17 | Non-depressed N=43 | Degrees of  freedom | | t-value | p-value | t-value | | p-value | Effect size d |
| Reference 15-item GDS Sheikh | 9.18 ± 3.45 | 3.81 ± 3.03 | reference | | | | | | | | |
| 12-item Sutcliffe | 7.82 ± 3.26 | 2.84 ± 2.65 | 27.35 | | 1.004 | 0.324 | -0.687 | | 0.249 | -0.196822 |
| 10-item D’Ath | 6.00 ± 2.31 | 2.26 ± 2.02 | 22.01 | | 1.898 | 0.0709 | 0.302 | | 0.617 | 0.08652147 |
| 10-item Van Marwijk | 6.82 ± 2.30 | 2.95 ±2.10 | 22.48 | | 1.226 | 0.233 | -0.379 | | 0.354 | -0.1085816 |
| 8-item Allgaier | 5.29 ± 2.44 | 1.44 ± 1.87 | 22.01 | | 2.657 | **0.0144\*** | 1.061 | | 0.850 | 0.3039711 |
| 7-item Broekman | 4.76 ± 1.89 | 1.26 ± 1.63 | 20.32 | | 3.005 | **0.00691\*** | 1.446 | | 0.918 | 0.4142717 |
| 5-item Hoyl | 3.24 ± 1.25 | 1.40 ± 1.05 | 17.57 | | 3.423 | **0.00312\*** | 1.936 | | 0.965 | 0.5546542 |
| 5-item Cheng | 3.41 ± 1.58 | 0.98 ± 1.34 | 18.62 | | 3.571 | **0.00209\*** | 2.054 | | 0.973 | 0.5884606 |
| 4-item D’Ath | 2.47 ± 1.01 | 0.65 ± 0.95 | 17.35 | | 4.141 | **0.000658\*** | 2.660 | | 0.992 | 0.7620765 |
| 4-item Van Marwijk | 2.82 ± 0.89 | 1.40 ± 0.85 | 16.97 | | 3.555 | **0.00244\*** | 2.086 | | 0.974 | 0.5976284 |
| 1-item D’Ath | 0.89 ± 0.33 | 0.28 ± 0.45 | 16.20 | | 5.035 | **0.000117\*** | 3.592 | | 0.999 | 1.02909 |
| 1-item Van Marwijk | 0.65 ± 0.49 | 0.07 ± 0.26 | 16.13 | | 5.247 | **0.000078\*** | 3.805 | | 0.999 | 1.090113 |
| Reference GDS 7-item Broekman | | | | | | | | | | |
| 10-item Van Marwijk | 6.82 ± 2.30 | 2.95 ±2.10 | 35.59 | | -2.562 | **0.0148\*** | -0.765 | | 0.775 | -0.2191686 |
| 4-item D’Ath | 2.82 ± 0.89 | 0.28 ± 0.45 | 20.17 | | 1.689 | 0.107 | 0.134 | | 0.553 | 0.03839032 |
| Reference GDS 4-item D’Ath | | | | | | | | | | |
| 10-item Van Marwijk | 6.82 ± 2.30 | 2.95 ±2.10 | 18.79 | | 3.958 | **0.000859\*** | 2.437 | | 0.988 | 0.6981881 |

LEGEND: SD Standard Deviation. MDD Major depression disorder group (n=17) was compared against the non-depressed group (n=43) on GDS scores using a standardized effect size d=0.5 with lower bound dL=-0.49 and upper bound dU=0.49 in the TOSTER package in the open-source programming language R with formula for effect size d=equivalence test t\*sqrt[(1/17+1/43)]. 1

\*Significant p-value < alpha 0.05. The two one-sided tests (null-hypothesis and equivalence test) were considered separately. For a definitive conclusion both tests should show significant p-value.

REFERENCE: 1Lakens, Scheel and Isager. 2018 Equivalence Testing for Psychological Research: A Tutorial. Advances in Methods and Practices in Psychological Science, 1(2), 259-269. <https://doi.org/10.1177/2515245918770963>