

➤ The Montgomery and Asberg Depression Rating Scale (MADRS)

Montgomery, S.A. & Asberg, M., (1979). A new depression scale designed to be sensitive to change. *The British Journal of Psychiatry*, 134, 382-389.

Instrument de mesure	Montgomery and Asberg Depression Rating Scale
Abréviation	MADRS
Auteur	Montgomery, S.A. & Asberg, M. (1979)
Thème	Dépression
Objectif	Evaluer la sévérité de la dépression et détecter les changements dus au(x) traitement(s)
Population	Patients chez lesquels un diagnostic de dépression a été posé
Utilisateurs	Cliniciens : psychiatres, médecins généralistes, psychologues ou infirmier(e)s
Nombre d'items	10
Participation du patient	Oui
Localisation de l'instrument de mesure	<p>Echelle en anglais : McDowell, I., (2006). A Guide to Rating Scales and Questionnaires. <i>Measuring Health, Oxford University Press, third edition</i>, 383-385</p> <p>Echelle en français : Bouvard, M. & Cottraux, J., (2005). Protocoles et échelles d'évaluation en psychiatrie et en psychologie. <i>Collection Pratiques en psychothérapie, Edition Masson, 4^e édition</i>, 195-197.</p>

OBJECTIF

L'objectif de la Montgomery and Asberg Depression Rating Scale est d'évaluer la sévérité de la dépression et de détecter les changements dus au(x) traitement(s).

PUBLIC CIBLE

Le public ciblé par la MADRS est celui des patients diagnostiqués comme souffrant de dépression.

DESCRIPTION

La Montgomery and Asberg Depression Rating Scale est une échelle de 10 items qui ont été choisis parmi les 65 éléments de la Comprehensive Psychopathological Rating Scale (CPRS) développée par Asberg *et al.* en 1978. Les éléments sélectionnés sont ceux qui étaient les plus fréquemment vérifiés et les plus sensibles au changement. L'échelle est remplie lors d'une interview clinique, le questionnaire part de questions larges sur les symptômes pour aboutir ensuite à une vue plus détaillée permettant l'évaluation précise de la gravité des symptômes. Les 10 éléments de l'échelle sont la tristesse apparente, la tristesse exprimée, l'incapacité à ressentir, les difficultés de concentration, la tension intérieure, les pensées pessimistes, les pensées suicidaires, la lassitude, la diminution de sommeil et la réduction de l'appétit. Ces 10 items se scorent de 0 à 6 (0 = pas de présence du symptôme, 6 = forte présence du symptôme). Seuls les scores pairs ont une valeur définitive, les scores impairs représentent des états intermédiaires.

Le score total est obtenu par addition des scores de chaque item, il varie entre 0 et 60. Plus le score est élevé, plus le trouble est présent. Snaith *et al.* (1986) proposent différents scores-seuils : des scores de 0 à 6 indiquent l'absence de symptômes, des scores de 7 à 19 indiquent une dépression légère, des scores de 20 à 34 indiquent une dépression modérée, des scores de 35 à 60 indiquent une dépression sévère. D'autres scores-seuils sont trouvés par Wilkin *et al.* (1992) : 12 pour la dépression légère, 24 pour une dépression modérée et 35 pour une dépression sévère.

FIABILITÉ

La consistance interne (*internal consistency*) de la MADRS atteint de bonnes valeurs. Elle varie entre 0.61 et 0.89. Maier and Philipp (1985) parlent d'un alpha de Cronbach de 0.86, Hammond *et al.* (1998) de 0.61, Bernstein *et al.* (2010) d'un alpha de Cronbach de 0.87, Doraiswamy *et al.* (2010) de 0.89 pour tout l'échantillon (N = 229), de 0.73 pour l'échantillon sans épisode de dépression majeure (N = 105), et de 0.61 pour l'échantillon avec épisode de dépression majeure (N = 124). Mundt *et al.* (2006) trouvent, eux, un alpha de Cronbach de 0.82. Hammond (1998), quant à lui, annonce un coefficient initial de 0.61 ; mais après suppression d'items, la consistance interne (estimée sur 5 items) atteint une valeur de 0.77. Toutes ces valeurs sont très élevées et indiquent une très bonne consistance interne.

L'équivalence (*equivalence*) est évaluée via la fidélité inter-juges. Montgomery and Asberg (1979) obtiennent des valeurs entre 0.89 et 0.97. Davidson *et al.* (1986) trouvent une valeur de 0.76 pour le score total, tandis que chaque item varie individuellement entre 0.57 et 0.76. Bramley *et al.* (1988) ont

une fidélité inter-juges de 0.98. Iannunzio *et al.* (2006) obtient une valeur de fidélité de 0.91. Les valeurs sont élevées et reflètent une bonne équivalence.

La corrélation intra-classes (équivalence – *equivalence*) est estimée par des valeurs de 0.66 à 0.82 (Maier *et al.*, 1988), de 0.86 (Kørner *et al.*, 1990) et de 0.98 (Iannunzio *et al.*, 2006). Ce sont de bonnes valeurs exprimant une bonne fiabilité.

La corrélation test-retest (stabilité – *stability*) obtient une valeur de 0.90. (Mundt *et al.*, 2006), ce qui représente une excellente stabilité.

VALIDITE

La validité concurrente (*criterion-related validity*) entre la MADRS et des évaluations cliniques de la sévérité de la dépression obtient des bonnes valeurs : 0.71 (Maier *et al.*, 1988) ; 0.75 (Maier, 1990) ; 0.83 (Snaith *et al.*, 1986), 0.86 (Agrell and Dehlin, 1989).

La corrélation (*criterion-related validity*) entre la MADRS et la Hamilton Rating Scale for Depression obtient des valeurs de 0.46 (Davidson *et al.*, 1986), de 0.82 (Kørner *et al.*, 1990), de 0.85 (Maier *et al.*, 1988), et de 0.70 (Montgomery and Asberg, 1979). Iannunzio *et al.* (2006) établit une corrélation entre la MADRS et diverses versions de la Hamilton Rating Scale for Depression : 0.90 avec la version à 17 items (HAM-D₁₇), 0.88 avec la version à 31 items (HAM-D₃₁), 0.89 avec la sous-échelle à 6 items (HAM-D₆), et 0.96 avec le HMI (HAM-D/MADRS Interview). A l'exception de l'étude de Davidson *et al.* (1986), toutes les valeurs sont très bonnes et démontrent une bonne validité.

La validité concurrente (*criterion-related validity*) de la MADRS est aussi estimée en regard d'autres échelles. Kørner *et al.* (1990) trouvent une corrélation de 0.92 et de 0.89 avec la Bech-Rafaelsen Melancholia Scale (BRMS). Maier *et al.* (1988) évaluent la corrélation entre la MADRS et la Raskin Depression Scale à 0.71. Snaith et Taylor (1985) rapportent une corrélation de 0.81 entre la MADRS et l'échelle de dépression de la Hospital Anxiety and Depression (HAD) scale. Avec la QIDSC₁₆ (Quick Inventory of Depressive Symptomatology-Clinician-rated), la corrélation est de 0.86. Avec la QIDS-SR₁₆ (self-reported Inventory of Depressive Symptomatology), elle est de 0.84 (Doraiswamy *et al.*, 2010). Toutes ces valeurs sont excellentes et dénotent une bonne validité.

La validité concurrente (*criterion-related validity*) est de 0.88 avec la CGI-S (Clinical Global Impression scale for severity) et de 0.61 avec la PGI-S (Patient version of the Clinical Global Impression scale for severity) (Mundt *et al.*, 2006).

La validité du construit (*construct validity*) est évaluée à 0.81 avec l'échelle de dépression de la Hospital Anxiety and Depression Scale (HADS) mais seulement 0.37 avec l'échelle d'anxiété (Snaith *et al.*, 1985). Cela prouve une bonne validité. Maier and Philipp (1985) trouvent une valeur mitigée de 0.60, ainsi que Maier *et al.*, (1988) et Maier (1990) qui obtiennent des valeurs de 0.61 et 0.63.

L'aire sous la courbe ROC (*area under the curve*) obtient une valeur de 0.69 (Doraiswamy *et al.*, 2010). Cette valeur est mitigée.

Agrell and Dehlin (1989) trouvent une sensibilité (*sensitivity*) très élevée (88%), tandis que les valeurs variant de 0.75 à 0.99 selon les scores-seuils (Doraiswamy *et al.*, 2010). Ces valeurs indiquent une très bonne validité.

La spécificité (*specificity*) est excellente car elle est représentée par des valeurs de 0.74 à 0.92 lorsque les scores-seuils varient de 8 à 14 (Doraiswamy *et al.*, 2010). Elle est très bonne aussi chez Agrell et Dehlin (1989) qui trouvent une spécificité de 70%.

La valeur predictive positive (*positive predictive value*) varie entre 0.83 et 0.93 selon le score-seuil (Doraiswamy *et al.*, 2010). Ces valeurs sont excellentes et expriment une bonne validité.

La valeur prédictive négative (*negative predictive value*) varie de 0.74 à 0.98 selon le score-seuil (Doraiswamy *et al.*, 2010). Quel que soit le score-seuil, la valeur prédictive négative est élevée et représente une bonne validité.

CONVIVIALITE

L'interview avec la MADRS dure de 20 à 60 minutes.

VARIANTE

Une variante de la MADRS est la MADRS-S, c'est une version de la MADRS à 9 items. La MADRS-S évalue l'humeur du patient, les sentiments de malaise, le sommeil, l'appétit, la faculté de concentration, les initiatives, l'engagement émotionnel, le pessimisme et la joie de vivre. Chaque élément est évalué de 0 à 3, avec des intervalles intermédiaires (0.5, 1.5, 2.5). Le score total est calculé par la somme des résultats de chaque élément. Comme pour la MADRS, plus le score est élevé, plus le trouble est présent.

La MADRS-S a une excellente fiabilité car elle obtient un alpha de Cronbach de 0.84 (consistance interne-*internal consistency*) et une corrélation intra-classes de 0.78.

Concernant sa fiabilité, la MADRS-S est corrélée (validité concurrente - *criterion-related validity*) à la MADRS avec 0.54. Cette valeur est mitigée, ce qui signifie que la MADRS-S n'est pas redondante avec la MADRS mais plutôt qu'elle est complémentaire. Avec un score-seuil de 5, la MADRS-S obtient une sensibilité (*sensitivity*) à 81,8% ; une spécificité (*specificity*) à 75,4% ; une valeur prédictive positive (*positive predictive value*) à 77,1% et une valeur prédictive négative (*negative predictive value*) à 80,3%. Ces valeurs sont élevées et démontrent une bonne fiabilité.

Une autre variante de la MADRS originelle est la IVR MADRS, c'est-à-dire la MADRS administrée de façon électronique, via le téléphone en utilisant une technologie interactive de réponse vocale (Interactive Voice Response (IVR) technology). La consistance interne (*internal consistency*) est évaluée à 0.79. Elle est légèrement inférieure à celle de la MADRS administrée par un clinicien (0.82). La stabilité (*stability*) obtient une valeur test-retest de 0.85. La corrélation (validité concurrente – *criterion-related validity*) entre la MADRS et la IVR MADRS est excellente grâce à une valeur de 0.82. La corrélation entre la CGI-S (Clinical Global Impression scale for severity) et la IVR MADRS est de 0.75 ; et celle entre la PGI-S (Patient version of the Clinical Global Impression scale for severity) et la IVR MADRS de 0.78. Ces valeurs indiquent une bonne corrélation.

REMARQUES

L'évaluation via la Montgomery and Asberg Depression Rating Scale ne permet pas de détecter les symptômes atypiques ou accessoires de la dépression comme la Hamilton Depression Rating Scale par exemple.

Un guide d'interview a été créé pour accroître la fiabilité et la validité de la MADRS. Ce guide est le HMI (Ham-D/MADRS Interview). Il est possible de le trouver dans l'article : Iannuzzo, R.W., Jaeger, J., Goldberg, J.F., Kafantaris, V., Sublette, M.E., (2006). Development and reliability of the HAM-D/MADRS Interview: An integrated depression symptom rating scale. *Psychiatry Research*, 145, 29-36.

RÉFÉRENCES

- Bernstein, I.H., Rush, A.J., Stegman, D., Macleod, L., Witte, B., & Trivedi, M.H., (2010). A Comparison of the QIDS-C16, QIDS-SR16, and the MADRS in an Adult Outpatient Clinical Sample. *CNS Spectr.*, 15(7), 458-468 (abstract).
- Bouvard, M. & Cottraux, J., (2005). Protocoles et échelles d'évaluation en psychiatrie et en psychologie. *Collection Pratiques en psychothérapie, Edition Masson, 4^e édition*, 174-175, 195-197.
- Bowling, A., (1997). A review of quality of life measurement scales. *Measuring Health, Open University Press, second edition*, 73-75.
- Doraiswamy, P.M., Bernstein, I.H., Rush, A.J., Kyutoku, Y., Carmody, T.J., Macleod, L., Venkatraman, S., Burks, M., Stegman, D., Witte, B., & Trivedi, M.H., (2010). Diagnostic utility of the Quick Inventory of Depressive Symptomatology (QIDS-C₁₆ and QIDS-SR₁₆) in the elderly. *Acta Psychiatr Scand*, 122, 226-234.
- Fantino, B. & Moore, N., (2009). The self-reported Montgomery-Asberg depression rating scale is a useful evaluative tool in major depressive disorder. *BMC Psychiatry*, 9, 26-31.
- Hammond, M.F., (1998). Rating depression severity in the elderly physically ill patient: reliability and factor structure of the Hamilton and the Montgomery-Asberg depression rating scales. *International Journal of Geriatric Psychiatry*, 13, 257-261.
- Iannuzzo, R.W., Jaeger, J., Goldberg, J.F., Kafantaris, V., & Sublette, M.E., (2006). Development and reliability of the HAM-D/MADRS Interview: An integrated depression symptom rating scale. *Psychiatry Research*, 145, 21-37.
- Katona, C.L.E., (1994). Depression in Old Age. *Wiley, Chichester*, p. 13.
- Montgomery, S.A. & Asberg, M., (1979). A new depression scale designed to be sensitive to change. *The British Journal of Psychiatry*, 134, 382-389.
- McDowell, I., (2006). A Guide to Rating Scales and Questionnaires. *Measuring Health, Oxford University Press, third edition*, 382-387.
- Mundt, J.C., Katzelnick, D.J., Kennedy, S.H., Eisfeld, B.S., Bouffard, B.B., & Greist, J.H., (2006). Validation of an IVRS version of the MADRS. *Journal of Psychiatric Research*, 40, 243-246.

LOCALISATION DE L'INSTRUMENT DE MESURE

- Echelle en anglais : McDowell, I., (2006). A Guide to Rating Scales and Questionnaires. *Measuring Health, Oxford University Press, third edition*, 383-385.
- Echelle en français : Bouvard, M. & Cottraux, J., (2005). Protocoles et échelles d'évaluation en psychiatrie et en psychologie. *Collection Pratiques en psychothérapie, Edition Masson, 4^e édition*, 195-197.

Montgomery and Asberg Depression Rating Scale

Montgomery, S.A. & Asberg, M. (1979)

Author (year)	Setting	Sample (n)	Design	Reliability	Validity
1. McDowell, I., (2006)			Book – description of the scale and of the studies over the scale	IC E	CrV CsV
2. Bernstein, I.H., Rush, A.J., Stegman, D., Macleod, L., Witte, B., Trivedi, M.H., (2010)		N = 278 consecutive outpatients		IC	
3. Bowling, A., (1997)			Description of the scale		CtV CrV
4. Doraiswamy, P.M., Bernstein, I.H., Rush, A.J., Kyutoku, Y., Carmody, T.J., Macleod, L., Venkatraman, S., Burks, M., Stegman, D., Witte, B., Trivedi, M.H., (2010)	Community and clinic patients	N = 229 community and clinic elderly subjects : - 105 not in a major depressive episode (MDE) - 124 in a major depressive episode (MDE)	Prospective study	IC	CrV ROC Sen Sp PPV NPV

5. Hammond, M.F., (1998)	Geriatric wards	N = 100 acute medical admissions aged 65 and over	Prospective study	IC	
6. Katona, C.L.E., (1994)		Geriatric stroke patients			CrV Sen Sp
7. Mundt, J.C., Katzelnick, D.J., Kennedy, S.H., Eisfeld, B.S., Bouffard, B.B., Greist, J.H., (2006)	General population	N = 66 subjects aged 22-64 years recruited through newspaper advertisements	Validation study	IC S	CrV
8. Iannuzzo, R.W., Jaeger, J., Goldberg, J.F., Kafantaris, V., Sublette, M.E., (2006)	Inpatient hospital	N = 50 inpatients subjects hospitalized for an acute manic or depressive episode	Prospective study	E	CrV

Betrouwbaarheid/ fiabilité: Stability (S), Internal Consistency (IC), Equivalence (E)

Validiteit/ validité: Face Validity (FV), Content Validity (CtV), Criterion Validity (CrV), Construct Validity (CsV)

Sensitivity (Sen), Specificity (Sp), Positive Predictive Value (PPV), Negative Predictive Value (NPV), Receiver Operating Curve (ROC), Likelihood Ratio (LR), Odds Ratio (OR), Area Under the Curve (AUC)

Results reliability	Results validity	Commentary
<p>1. IC</p> <p>An alpha of 0.86 was found in a study of 151 depressed patients (Maier and Philipp, 1985), but a much lower alpha, 0.61, has also been reported (Hammond <i>et al.</i>, 1998).</p> <p>E interrater reliability</p> <p>Montgomery and Asberg (1979) reported inter-rater reliability ranged from 0.89 to 0.97 for various combinations of raters in small samples of 12 to 30 patients.</p> <p>Davidson <i>et al.</i> (1986) reported Spearman inter-rater correlations of 0.76 for the total score, whereas correlations for individual items ranged from 0.57 to 0.76.</p> <p>A Pearson correlation of 0.98 between raters has been reported (Bramley <i>et al.</i>, 1988)</p> <p>E</p> <p>Intraclass coefficients for the MADRS fell between 0.66 to 0.82 (Maier <i>et al.</i>, 1988). An equivalent figure of 0.86 was reported by Kørner <i>et al.</i> (1990).</p>	<p>CrV</p> <p>Maier <i>et al.</i> (1988) compared the MADRS with clinical assessments of severity of depression; the correlation was 0.71.</p> <p>Equivalent figures from a second study were 0.75 for the MADRS (Maier, 1990).</p> <p>In Snaith's study, MADRS scores correlated $\rho = 0.83$ with a clinician's global severity rating (Snaith <i>et al.</i>, 1986).</p> <p>Davidson <i>et al.</i> (1986) reported Spearman correlations averaging only 0.46 between HRSD ratings and the MADRS.</p> <p>Other coefficients are higher, including Spearman correlations with the HRSD of 0.82 and with the Bech-Rafaelsen Melancholia Scale (BRMS) of 0.92 (Kørner <i>et al.</i>, 1990), and Pearson correlations of 0.85 with the HRSD, 0.89 with the BMRS, and 0.71 with the Raskin Depression Scale (Maier <i>et al.</i>, 1988).</p> <p>CsV</p> <p>Snaith and Taylor reported convergent and discriminant validity : MADRS scores correlated 0.81 with the depression score of the Hospital Anxiety and Depression Scale (HADS) and 0.37 with its anxiety score (Snaith <i>et al.</i>, 1985).</p> <p>Maier and Philipp (1985) found that MADRS had higher mean discriminatory power (coefficient = 0.60) than the Hamilton (coefficient = 0.39). By contrast, in</p>	

	two other studies the MADRS had slightly lower correlations with global assessments of change (0.61 and 0.63) (Maier <i>et al.</i> , 1988 ; Maier, 1990).	
2. IC a = 0.87 for MADRS 10		
3. E interrater reliability Comparisons between two English raters, two Swedish raters and one English and one Swedish rater, rating 11 to 30 patients, produced correlations between 0.89 and 0.97 (Montgomery <i>et al.</i> , 1979)	CtV 54 English and 52 Swedish patients completed a 65-item psychopathology scale. Analysis identified the 17 most commonly occurring symptoms in depressive illness. Subsequent analyses, using 64 patients on different types of antidepressive drugs, were then used to create a 10-item depression scale. CrV The scale was tested against the Hamilton Depression Scale with a reported correlation of 0.70. (Montgomery and Asberg, 1979). Snaith and Taylor (1985) reported a high correlation between the scale and the depression scale of their self-rated Hospital Anxiety and Depression (HAD) scale (0.81).	
4. IC Cronbach's α = 0.89 (All, N = 229), 0.73 (no MDE, N = 105), 0.61 (yes MDE, N = 124)	CrV Scale scores were highly correlated: r = 0.86 for the MADRS and QIDSC ₁₆ (Quick Inventory of Depressive Symptomatology-Clinician-rated); 0.84 for the MADRS and QIDS-SR ₁₆ (self-reported Inventory of Depressive Symptomatology)	

	<p>ROC</p> <p>The area under the ROC curve was 0.69 for the MADRS</p> <p>Sen</p> <p>Cutoff 8 : 0.99 ; cutoff 10 : 0.89 ; cutoff 12 : 0.83 ; cutoff 14 : 0.75</p> <p>Sp</p> <p>Cutoff 8 : 0.74 ; cutoff 10 : 0.82 ; cutoff 12 : 0.87 ; cutoff 14 : 0.92</p> <p>PPV</p> <p>Cutoff 8 : 0.83 ; cutoff 10 : 0.87 ; cutoff 12 : 0.89 ; cutoff 14 : 0.93</p> <p>NPV</p> <p>Cutoff 8 : 0.98 ; cutoff 10 : 0.85 ; cutoff 12 : 0.79 ; cutoff 14 : 0.74</p>	
<p>5. IC</p> <p>The initial coefficient was 0.61. Successive deletion of items to maximise α resulted in a five-item scale containing apparent and reported sadness, inability to concentrate, lassitude and inability to feel, with a final coefficient α of 0.77.</p> <p>IC</p> <p>Three items (reduced sleep, reduced appetite and pessimistic</p>		

thoughts) failed to correlate at 0.30 or above with any other items.		
6.	CrV – Sen – Sp Agrell and Dehlin (1989) found the MADRS to have excellent criterion validity ($r = 0.86$) against a global clinical rating of depression as well as showing a high sensitivity (88%) and specificity (70%).	
7. IC Cronbach's alpha : 0.82 S The test-retest correlation over the two days was 0.90.	CrV The clinician MADRS scores correlated 0.88 and 0.61 with the CGI-S (Clinical Global Impression scale for severity) and PGI-S (Patient version of the Clinical Global Impression scale for severity), respectively.	
8. E intraclass correlation ICC : 0.98 E interrater reliability Spearman's ρ correlations : 0.91 Spearman's ρ correlations (Davidson <i>et al.</i> , 1986) : 0.76	CrV MADRS – HAM-D ₁₇ : 0.90 MADRS – HAM-D ₃₁ : 0.88 MADRS – HAM-D ₆ subscale : 0.89 HMI (HAM-D/MADRS Interview) – MADRS : 0.96	

Betrouwbaarheid/ fiabilité: Stability (S), Internal Consistency (IC), Equivalence (E)

Validiteit/ validité: Face Validity (FV), Content Validity (CtV), Criterion Validity (CrV), Construct Validity (CsV)

Sensitivity (Sen), Specificity (Sp), Positive Predictive Value (PPV), Negative Predictive Value (NPV), Receiver Operating Curve (ROC), Likelihood Ratio (LR), Odds Ratio (OR),

Area Under the Curve (AUC)