

Émilie Jestin¹, Jérôme De Launay²,
Philippe Castera¹, Yves Montariol¹,
Jean-Philippe Joseph¹, Nicolas
Rousselot¹

¹Département de Médecine Générale,
Université de Bordeaux, Case 148. 146 rue
Léo Saignat, 33076 Bordeaux cedex
nicolas.rousselot@u-bordeaux.fr

²Hôpitaux Universitaires de Strasbourg

Tirés à part : N. Rousselot

Résumé

Les médecins généralistes sont en première ligne pour la prise en charge de la dépression mais le repérage des patients concernés est parfois complexe. Le test à deux questions d'Arroll est recommandé pour ses très bonnes performances pour le repérage des épisodes dépressifs caractérisés (EDC) en soins premiers dans différents pays anglo-saxons. Mais ce test, dont l'utilisation était recommandée par l'ANAES dès 2002, n'a pas encore été validé en français. Son utilisation permettrait de mieux repérer les EDC, préalable à la mise en place de démarches diagnostique et thérapeutique adaptées. L'objectif de notre étude était de mesurer les qualités métrologiques du test à deux questions d'Arroll traduit en français, dans le repérage de l'EDC en médecine générale.

• Mots clés

dépistage ; épisode dépressif majeur ; médecine générale.

Abstract. Identifying the major depressive disorder (MDD) in primary care. Metrological quality of two screening questions according to Arroll

General practitioners are on the front line for the care of depressive patient but their identification can be difficult. The Arroll test in two screening-questions is recommended in many Anglo-Saxon countries for major depression screening, thanks to its good performances. In France, even though its use was recommended by ANAES in 2002, the Arroll test has not been validated in French language so far. Its use could help to identify major depression, which is later needed for adapted diagnosis and therapy. The purpose of this study was to measure metrological properties of the French

Repérer l'épisode dépressif caractérisé en médecine générale

Qualités métrologiques du test à deux questions d'Arroll en français

Introduction

En France, la dépression est un problème de santé publique : par sa fréquence, avec une prévalence annuelle estimée par l'INPES en 2010 à 7,5 % en population générale [1] ; par sa morbi-mortalité importante (première cause de suicide, troubles anxieux et addictions fréquemment associés, retentissement sur la qualité de vie...) ; et par les coûts qu'elle engendre de manière directe (traitements, hospitalisations...) ou indirecte (arrêts de travail, retentissement sur l'entourage...) [2-4].

Les médecins généralistes sont en première ligne pour sa prise en charge mais de nombreux patients restent mal diagnostiqués et traités en population générale [5, 6]. Les difficultés au repérage des épisodes dépressifs caractérisés (EDC), du fait de symptômes masqués par d'autres motifs de consultation ou des comorbidités, constituent un premier obstacle important dans des consultations en temps contraint. On estime que seuls 26 % des patients déprimés consultent explicitement pour une souffrance ou des difficultés psychologiques [7]. En 2017, la Haute Autorité de Santé (HAS) a fait de la meilleure identification des patients atteints d'un EDC isolé l'un de ses objectifs majeurs d'amélioration des pratiques [3].

Les classifications pour le diagnostic d'EDC (CIM-10, DSM-V) et les échelles d'évaluation de l'intensité des symptômes dépressifs (BDI, HDRS, MADRS...) sont peu utilisées en médecine générale du fait de leur longueur et de leur utilisation fastidieuse [3, 8]. Des questionnaires ultracourts de dépistage de la dépression à un, deux ou trois items ont ainsi été développés. Parmi eux, le test à deux questions d'Arroll¹ [8] a montré les meilleures performances à visée de repérage en soins premiers et en population générale dans plusieurs études anglo-saxonnes et suisse (*tableau 1*) [8-13]. Son utilisation permet donc de mieux repérer, parmi les patients pris en charge, ceux pour lesquels il est nécessaire de mettre en place une démarche diagnostique approfondie à la recherche d'un possible EDC. Son utilisation est ainsi notamment recommandée au Royaume-Uni, aux États-Unis et en Nouvelle-Zélande [14-16].

En France, l'Agence Nationale d'Accréditation et d'Évaluation en Santé (ANAES) évoquait dès 2002 l'intérêt de ce test à deux questions, malgré une traduction approximative, dans ses recommandations sur la prise en charge d'un épisode dépressif isolé de l'adulte en ambulatoire [17]. Mais aucune étude n'a validé son utilisation dans le dépistage de l'EDC, en soins premiers, en France.

¹ « During the past month, have you often been bothered by feeling down, depressed or hopeless? » ; « During the past month, have you often been bothered by little interest or pleasure in doing things? ».



version of the two-question Arroll[®] test for MDD screening in primary care.

• **Key words**
screening; depressive disorder major; primary care.

DOI: 10.1684/med.2018.319

L'objectif de cette étude était de mesurer les qualités métrologiques du test à deux questions d'Arroll traduit en français pour le repérage de l'EDC en médecine générale.

Méthodes

Notre étude était prospective et observationnelle. Des médecins généralistes exerçant en Gironde ont proposé, durant leurs consulta-

tions, le test à deux questions d'Arroll à tous les patients consécutifs qu'ils recevaient sur une ou plusieurs demi-journées. La traduction utilisée, validée dans l'étude de Lombardo [11], était : « *Au cours du dernier mois, étiez-vous souvent triste, déprimé(e) ou désespéré(e) ?* » et « *Au cours du dernier mois, avez-vous ressenti un manque d'intérêt et de plaisir dans la plupart de vos activités ?* ». Le test était considéré comme positif quand le patient répondait oui à l'une des deux questions.

Le test de référence utilisé pour le diagnostic d'EDC était le questionnaire Patient Health Questionnaire 9 (PHQ9) [18] que les patients remplissaient en fin de consultation. Ce test d'auto-évaluation réalisable en moins de 5 minutes reprenait les critères du DSM-IV et avait été validé pour le diagnostic d'EDC chez les patients majeurs en soins premiers [3, 18, 19]. Le test était considéré comme positif pour un score supérieur ou égal à 10 [3, 18].

Étaient exclus les patients mineurs, souffrant de troubles cognitifs les empêchant de répondre aux tests, ayant une pathologie psychiatrique aiguë (manie, psychose) ou sous l'emprise d'une substance psychoactive au moment de la consultation, et ceux ne lisant pas le français ou déficients visuels.

Le nombre de sujets nécessaires était, pour une sensibilité attendue de 90 %, une prévalence d'EDC de 10 %, un risque de 5 % et une précision de 10 %, de 346 selon la formule de Buderer [20].

Les tests statistiques ont été réalisés avec le logiciel Stata/MP 14.0. Les sensibilité, spécificité, valeur prédictive positive (VPP) et négative (VPN) du test à deux questions d'Arroll ont été calculées en utilisant le questionnaire

PHQ9 comme référence pour le diagnostic d'EDC avec mesure d'un intervalle de confiance à 95 %.

Résultats

Population d'étude

Onze médecins généralistes girondins (d'âge, de sexe et de lieu d'exercice variés) ont inclus 348 patients du 04 mars au 01 août 2016 (*figure 1*), soit en moyenne 32 patients (± 33) par médecin.

Ces patients étaient à 61 % des femmes, ils avaient en moyenne 49 ans (± 18) et 19 % étaient considérés comme actuellement « déprimés » par les médecins avant la réalisation des tests (*tableau 2*). La dépression était l'un des motifs de consultation pour 38 patients (10,9 %). Parmi eux, 29 (76,3 %) avaient un test d'Arroll positif.

Valeurs métrologiques du test d'Arroll

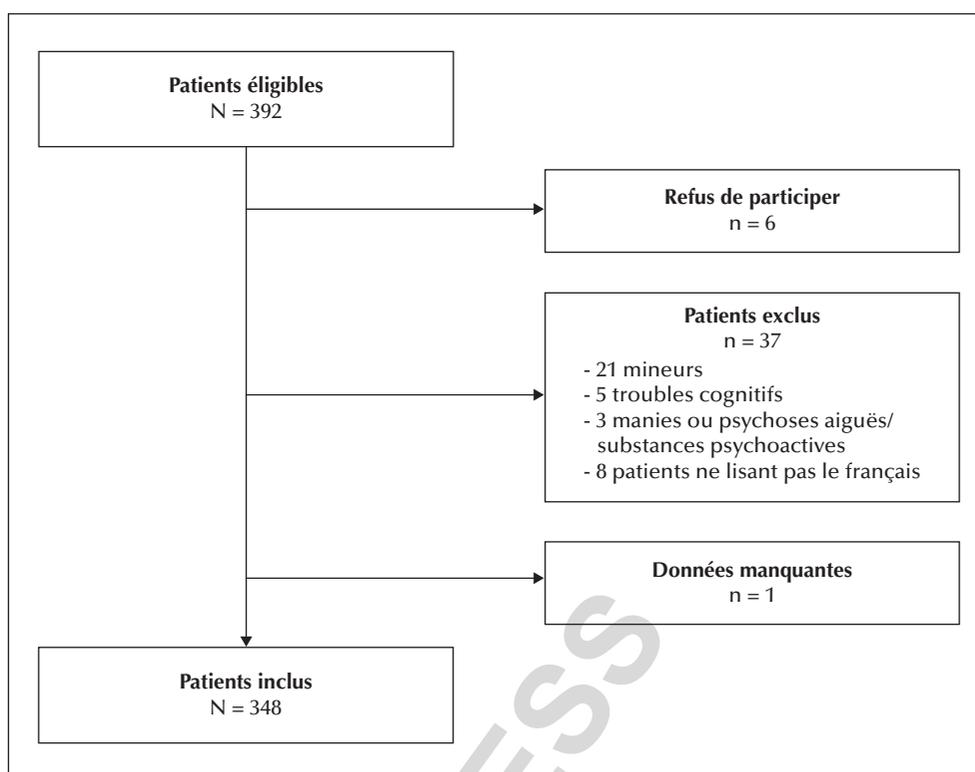
La prévalence d'EDC dans la population d'étude d'après les résultats du test PHQ9 était de 24,7 %. Pour le dépistage d'EDC par le test d'Arroll, la sensibilité était de 84,9 % (75,5-91,7), la spécificité de 75,6 % (69,9-80,7), la VPP de 53,3 % (44,6-61,9) et la VPN de 93,8 % (89,7-96,7) (*tableau 3*).

En faisant une analyse dans les différents sous-groupes de la population incluse, la sensibilité et la VPN semblaient moins bonnes chez les patients de plus de 65 ans et

• **Tableau 1.** Caractéristiques et résultats des principales études ayant évalué le test d'Arroll.

Étude	Sen (%)	VPN (%)	Effectif (n)	Gold standard	Prév dép. (%)
Arroll 2003 (NZ)	97	99	476	CIDI	6,9
Goodyear Smith 2008 (NZ)	96	100	995	PHQ9	5,3
Lombardo 2011 (Suisse)	91	99	724	PHQ9	9,5
Mohd Sidik 2011 (Malaisie)	99	-	146	CIDI	21,2
Bosanquet et al 2015 (RU)*	96	-	-	DSM-IV	-

Sen = sensibilité, VPN = valeur prédictive négative, Prév dép. = prévalence dépression. NZ = Nouvelle-Zélande, RU = Royaume-Uni. *Méta-analyse.



• Figure 1. Diagramme de flux.

• Tableau 2. Caractéristiques des 348 patients inclus dans l'étude.

Âge moyen en années (ET)*	49	(18)
Sexe, n (%)		
Hommes	137	(39,4)
Femmes	211	(60,6)
Pathologies, n (%)		
Trouble dépressif actuel	66	(18,9)
Antécédent de dépression	88	(25,2)
Trouble anxieux actuel	94	(26,9)
Trouble bipolaire	7	(2,0)
Trouble psychotique	5	(1,4)
Traitements en cours, n (%)		
Antidépresseur	67	(19,2)
Benzodiazépine	71	(20,3)
Thymorégulateur	13	(3,7)
Neuroleptique	27	(3,3)
Toxiques, n (%)		
Tabac	77	(22,1)
Alcoolisme	20	(5,7)
Cannabis chronique	20	(5,7)

*ET = écart-type.

meilleures chez les patients de moins de 35 ans, ceux ayant un antécédent de dépression, ceux considérés comme ayant un trouble anxieux et les fumeurs (tableau 4). L'exclusion des patients déjà considérés par leur médecin généraliste comme actuellement dépressifs ne modifiait pas les valeurs métrologiques du test (tableau 4). Plusieurs médecins investigateurs ont signalé le fait que la proposition de ce test sur ces consultations leur avait permis de diagnostiquer puis prendre en charge plusieurs patients souffrant d'EDC qu'ils n'auraient pas identifiés.

Discussion

Dans notre population, le test à deux questions d'Arroll possédait de bonnes performances pour le dépistage de l'EDC avec une sensibilité à 84,9 % (75,5-91,7) et une VPN à 93,8 % (89,7-96,7) par rapport au questionnaire PHQ9.

• Tableau 3. Tableau de contingence des résultats des tests d'Arroll et PHQ9 dans la population d'étude. N = 348.

	PHQ9 positif	PHQ9 négatif	Total
Arroll positif	73	64	137
Arroll négatif	13	198	211
Total	86	262	348

• **Tableau 4.** Qualités métrologiques du test d'Arroll selon les principales caractéristiques de la population d'étude.

Caractéristiques	Sensibilité % (IC95 %)	VPN % (IC95 %)
Ensemble des patients	84,9 (75,5-91,7)	93,8 (89,7-96,7)
Sexe féminin	84,4 (73,1-92,2)	93,9 (89,6-96,4)
Âge (en années)		
< 35	90,0 (68,3-98,8)	95,7 (85,5-98,8)
35-65	85,7 (72,8-94,1)	94,3 (89,2-97,1)
> 65	76,5 (50,1-93,2)	90,8 (80,6-95,9)
Patients non actuellement dépressifs*	84,1 (69,9-93,4)	93,7 (88,3-96,7)
Antécédent de dépression	95,2 (76,2-99,9)	98,1 (88,2-99,7)
Trouble anxieux actuel	88,9 (73,9-96,9)	94,1 (86,2-97,6)
Traitement		
Antidépresseur	80,0 (63,1-91,6)	91,3 (83,8-95,5)
Benzodiazépine	82,4 (65,5-93,2)	90,7 (81,8-95,5)
Tabagisme	92,0 (74,0-99,0)	96,1 (86,7-99,0)

Ces résultats sont un peu moins bons que ceux des études anglo-saxonnes qui retrouvaient une sensibilité entre 91 % et 99 % et une VPN à 99 % (*tableau 1*) [8, 10-13]. Ces études présentaient cependant des différences notables en termes de conditions de réalisation, de populations considérées (pays, sexe exclusivement féminin [12], âge plus élevé [11], prévalence retrouvée de dépression très inférieure [8, 10, 11]) et de tests de référence utilisés (CIDI [8, 12], DSM [13]).

Population d'étude

Le choix de recruter des patients consécutifs, comme dans la plupart des autres études sur le sujet [8, 10, 12], permettait de limiter un biais de sélection des patients par les médecins investigateurs. Les patients recrutés étaient majoritairement des femmes, ce qui correspond à la pratique clinique en médecine générale [21]. Leur répartition en âge était globalement équilibrée avec cependant une sous-représentation des patients de plus de 70 ans [21]. Ceci s'explique probablement par le fait que les patients âgés sont souvent vus à leur domicile et non au cabinet.

Malgré tout, les proportions importantes de patients ayant un antécédent de dépression (25 % versus 18 % selon l'enquête ANADEP 2005 [22]), considérés comme souffrant actuellement d'un trouble dépressif ou anxieux (respectivement 19 % et 27 % versus 7,5 % et 15 % selon l'INPES [1, 23]) suggèrent un possible biais de sélection et/ou d'information dans notre étude. La sélection était faite par un petit nombre de médecins investigateurs, qui malgré leur diversité, étaient volontaires et donc probablement plus sensibles à cette problématique de la dépression, et donc avec des patientèles plus souvent déprimées. Toutefois, la prévalence de la dépression est souvent sous-estimée et les données chiffrées, peu nombreuses, concernent plus souvent la population générale française. Or celle-ci n'est pas superposable à la population consultant en médecine générale ; dans les années 2000, la prévalence d'EDC était estimée à 6 % en

population générale et entre 9 et 14 % en médecine générale [24]. Enfin, la prévalence des syndromes dépressifs, relevée pour les caractéristiques des patients dans notre étude, est également supérieure à la prévalence d'épisodes dépressifs caractérisés.

Test de référence

Le choix du test de référence pour le diagnostic de certitude d'un EDC en médecine générale est complexe du fait de la validité imparfaite des outils existants, mais également d'un problème de faisabilité, notamment pour l'usage du DSM-V couplé à un entretien psychiatrique.

Le PHQ9, utilisé dans de nombreuses autres études pour ses bonnes performances diagnostiques, était simple d'utilisation, mais présentait les limites d'être un auto-questionnaire, reprenant les deux questions du test d'Arroll parmi ses neuf items (malgré des modalités d'évaluation différentes) et n'évaluant pas l'ensemble des critères du DSM [3]. Sa sensibilité, de l'ordre de 85 % selon les études, peut varier en fonction du seuil choisi [3, 18].

L'imperfection de notre test de référence constitue un potentiel biais de mesure qui pourrait aussi expliquer la prévalence importante de patients souffrant d'un EDC dans notre échantillon (25 %).

Place dans la stratégie de dépistage

La stratégie de dépistage de l'EDC n'est actuellement pas claire. Certains pays recommandent un dépistage systématique régulier (États-Unis), d'autres un dépistage orienté en fonction des facteurs de risque (Angleterre, Nouvelle-Zélande). En France, le dépistage de l'EDC est évoqué dans les dernières recommandations HAS de 2017 mais non clairement recommandé et ses modalités ne sont pas abordées [3]. Le PHQ-2 est cité parmi les outils utiles pour aider à repérer les EDC [3]. Cet outil, non validé en France, est très proche du test à deux questions d'Arroll que nous avons utilisé dans notre étude.

Il comprend les deux mêmes questions, administrées en auto-questionnaire avec des réponses cotées en quatre niveaux de fréquence de symptômes sur les deux dernières semaines, alors que notre test était un hétéro-questionnaire avec des réponses portant sur l'existence ou non des symptômes lors du dernier mois. Dans tous les cas, ces outils ne se substituent pas au sens clinique et à la nécessaire démarche diagnostique qui doit être mise en œuvre en cas de test positif [3].

Les résultats dans les sous-groupes de notre étude, malgré un manque de puissance, suggèrent que le test d'Arroll présente de bonnes performances pour un repérage de l'EDC, ce d'autant que l'exclusion des patients déjà considérés comme dépressifs ne modifie pas ses valeurs métrologiques. Ils suggèrent aussi qu'il pourrait être intéressant pour un dépistage orienté vers les patients ayant certains facteurs de risque connus de dépression [1-3, 7] : patients jeunes (ici les 18-35 ans), patients avec un antécédent de dépression, patients ayant un trouble anxieux ou patients fumeurs. En revanche, son utilisation semble moins pertinente chez les patients âgés de plus de 65 ans du fait de moindres performances et de l'existence d'autres outils de dépistages validés.

Conclusion

Malgré des limites en termes de représentativité et de puissance, notre étude suggère que le test à deux questions d'Arroll fait preuve de bonnes performances pour le repérage de l'EDC en médecine générale en France. Il serait intéressant de confirmer ces résultats par



En pratique

- Test à 2 questions d'Arroll :
 - « Au cours du dernier mois, étiez-vous souvent triste, déprimé(e) ou désespéré(e) ? »
 - « Au cours du dernier mois, avez-vous ressenti un manque d'intérêt et de plaisir dans la plupart de vos activités ? »
- Positif en cas de réponse affirmative à l'une des deux questions.
- Ce test simple et rapide présente de bonnes performances pour le repérage de l'EDC en médecine générale avec une sensibilité à 84,9 % (75,5-91,7) et une VPN à 93,8 (89,7-96,7).
- Sa place dans un programme de dépistage de l'EDC en médecine générale reste à définir. Un résultat positif nécessite ensuite une démarche diagnostique approfondie.

une étude plus large et de définir des stratégies de dépistage les plus adaptées en termes de santé publique.

Quoi qu'il en soit, ce test est d'une grande simplicité, rapide et acceptable, utilisable en consultation pour le repérage des EDC.

~ **Liens d'intérêts** : les auteurs déclarent n'avoir aucun lien d'intérêt en rapport avec cet article.

RÉFÉRENCES

1. Beck F, Guignard R. La dépression en France (2005-2010) : prévalence, recours au soin et sentiment d'information de la population. *La Santé de l'homme* 2012 ; (421) : 43-5.
2. Direction de la recherche des études de l'évaluation et des statistiques, Morin T. Épisodes dépressifs : des situations multiples. *Études et Résultats*. 2008 ; (661).
3. HAS. Épisode dépressif caractérisé de l'adulte : prise en charge en soins de premier recours. Méthode Recommandation pour la pratique clinique. 2017.
4. Caisse nationale d'assurance maladie des travailleurs salariés. Améliorer la qualité du système de santé et maîtriser les dépenses. Propositions de l'Assurance Maladie pour 2015. 2014.
5. Direction de la recherche des études de l'évaluation et des statistiques, Coldefy M, Nestrigue C. La prise en charge de la dépression dans les établissements de santé. *Études et Résultats* 2013 ; (860).
6. Dumesnil H, Cortaredona S, Cavillon M, et al. La prise en charge de la dépression en médecine générale de ville. Direction de la recherche des études de l'évaluation et des statistiques. *Études et Résultats* 2012 ; (810).
7. Magalon-Bingenheimer K, Magalon D, Zendjidjian X, Boyer L, Griguer Y, Lançon C. Dépression en médecine générale. *Presse Med* 2013 ; 42 (4) : 419-28.
8. Arroll B, Khin N, Kerse N. Screening for Depression in Primary Care with Two Verbally Asked Questions : Cross Sectional Study. *Brit Med J* 2003 ; 327 (7424) : 1144-6.
9. Mitchell AJ, Coyne JC. Do ultra-short screening instruments accurately detect depression in primary care? *Brit J Gen Pract* 2007 ; 57 (535) : 144-51.
10. Goodyear-Smith F, Coupe NM, Arroll B, Elley CR, Sullivan S, McGill A-T. Case finding of lifestyle and mental health disorders in primary care : validation of the "CHAT" tool. *Brit J Gen Pract* 2008 ; 58 (546) : 26-31.
11. Lombardo P, Vaucher P, Haftgoli N, et al. The "Help" Question Doesn't Help When Screening for Major Depression : External Validation of the Three-Question Screening Test for Primary Care Patients Managed for Physical Complaints. *BMC Med* 2011 ; 9 : 114.
12. Mohd-Sidik S, Arroll B, Goodyear-Smith F, Zain AM. Screening for Depression with a Brief Questionnaire in a Primary Care Setting : Validation of the Two Questions with Help Question (Malay Version). *Int J Psychiat Med* 2011 ; 41 (2) : 143-54.
13. Bosanquet K, Bailey D, Gilbody S, et al. Diagnostic accuracy of the Whooley questions for the identification of depression : a diagnostic meta-analysis. *Brit Med J Open* 2015 ; 5 (12) : e008913.
14. National Institute for Health and Care Excellence. The treatment and management of depression in adults. NICE clinical guideline 90. London : NICE ; 2009.
15. US. Preventive Services Task Force. Screening for Depression in Adults : US. Preventive Services Task Force Recommendation Statement. *Ann Intern Med* 2009 ; 151 (11) : 784-92.
16. New Zealand Guidelines Group. Identification of Common Mental Disorders and Management of Depression in Primary Care. Evidence-based best practice guideline. Wellington : Ministry of Health ; 2008.
17. Agence nationale d'accréditation et d'évaluation en santé. Prise en charge d'un épisode dépressif isolé de l'adulte en ambulatoire. Recommandations pour la pratique clinique. Paris : ANAES, 2002.
18. Kroenke K, Spitzer RL, Williams JB. The PHQ-9 : Validity of a Brief Depression Severity Measure. *J Gen Intern Med* 2001 ; 16 (9) : 606-13.
19. Gilbody S, Richards D, Barkham M. Diagnosing depression in primary care using self-completed instruments?: UK validation of PHQ-9 and CORE-OM. *Brit J Gen Pract* 2007 ; 57 (541) : 650-2.
20. Buderer NM. Statistical Methodology : I. Incorporating the Prevalence of Disease into the Sample Size Calculation for Sensitivity and Specificity. *Acad Emerg Med* 1996 ; 3 (9) : 895-900.
21. Direction de la recherche des études de l'évaluation et des statistiques, Labarthe G. Les consultations et visites des médecins généralistes. Un essai de typologie. *Études et Résultats* 2004 ; (315).
22. Institut National de Prévention et d'Éducation pour la Santé, Chan Chee C, Beck F, Sapinho D, Guilbert P. La dépression en France. Enquête Anadep 2005. *Études santé*. 2009.
23. HAS. Affections psychiatriques de longue durée : troubles anxieux graves. Guide affection longue durée. 2007.
24. Gallais JL, Alby ML. *Psychiatrie, souffrance psychique et médecine générale*. Encycl Méd Chir Psychiatrie. Paris : Editions Scientifiques et Médicales Elsevier ; 2002.