Identifying Key Parent-Reported Symptoms For Detecting Depression in High Risk Adolescents.

Authors: Ajay K Thapar*[a,b], Kerenza Hood [c], Stephan Collishaw [b], Gemma Hammerton [b], Becky Mars [d], Ruth Sellers [b], Robert Potter [e], Nick Craddock [b], Anita Thapar [b], Frances Rice[b,f].

Institutional Affiliations:

[a]Taff Riverside Practice, Wellington Street, Cardiff CF11 9SH, Wales, UK

[b]Institute of Psychological Medicine and Clinical Neuroscience, MRC Centre for Neuropsychiatric Genetics and Genomics, Cardiff University School of Medicine, Hadyn Ellis Building, Maindy Road ,Cardiff CF24 4HX, Wales, UK

[c]Cardiff Centre for Trials Research, Neuadd Merionydd, University Hospital of Wales, Cardiff University; Cardiff CF14 4XN, Wales, UK

[d]School of Social and Community Medicine, University of Bristol, Canynge Hall, Whatley Road, Bristol BS8 2PS, England, UK

[e]Child and Adolescent Mental Health Services, Trehafod, Waunarlwydd Road, Swansea SA2 0GB;Wales, UK.

[f]Division of Psychology and Language Sciences, University College, Bedford Way, London WC1H 0AP, England UK (current address [b])

***Corresponding author:** Ajay Thapar, GP, Taff Riverside Practice, Riverside Health Centre, Wellington Street, Canton, Cardiff CF11 9SH. thaparak@cf.ac.uk Tel 20688490 Fax 02920803209

Abstract

The adolescent offspring of depressed parents are at heightened risk of developing early onset Major Depressive Disorder (MDD) yet are unlikely to access services. One solution involves asking parents about the adolescents' symptoms in order to identify those in need of additional assessment. We aimed to identify a parsimonious combination of parent-reported symptoms that accurately detected offspring MDD. We used a multi-sample study comprising a development sample of 335 offspring of adults with recurrent MDD assessed on three occasions (mean age 12.4-14.8 years) and an independent validation sub-sample of 807 adolescents drawn from a general population cohort (mean age 13.1 years). Parent ratings of psychiatric symptoms in adolescent offspring were assessed using established questionnaires and analysed using multivariate regression. The best performing combination of symptoms was identified. Accuracy in detecting concurrent DSM-IV MDD diagnosis, assessed by direct adolescent and parent interviews, was compared to the well-established 13-item short Moods and Feelings Questionnaire (sMFQ) using ROC curve analysis. We identified a symptom combination of four items (concentration problems, anhedonia, worrying excessively and feeling unloved) which performed equivalently to the sMFQ both in the development dataset (combination C-index (mean)= 0.83; sMFQ C-index(mean)=0.84) and in the validation dataset (combination C-index= 0.82; sMFQ C-index=0.83). We concluded that a combination of four parent-reported mental health items performs equivalently to an established, longer depression questionnaire measure in detecting a diagnosis of adolescent major depressive disorder among offspring of parents with recurrent MDD and needs further evaluation.

Key words : Screening; Major Depressive Disorder; ROC analysis; Adolescence; ALSPAC

Highlights:

- We focused on major depressive disorder (MDD) in adolescent offspring of adults with recurrent MDD
- We examined whether a few particular symptoms of psychopathology (from established questionnaires) would detect adolescent MDD at similar accuracy to a longer, established questionnaire measure.
- We identified that if at least two of four symptoms (anhedonia, concentration problems, worrying excessively and feeling unloved) were reported to be present in adolescent offspring by parents, then accuracy in detecting MDD was similar to that achieved by a longer, established questionnaire measure.
- We identified that if fewer than 2 parent- reported symptoms were present then MDD was highly unlikely to be present.

1. Introduction

Individual depressive symptoms are extremely common in adolescence (Kubik et al.,2003) and most resolve spontaneously (Rushton et al., 2002). However if symptoms are persistent and multiple this may indicate Major Depressive Disorder (MDD), a much more serious, debilitating problem with a substantially increased risk of serious sequelae, including suicide and long-term health risks (Lopez et al., 2006; Thapar et al., 2012). There are strict criteria for defining MDD using internationally applied classification systems (American Psychiatric Association , 2013; World Health Organization, 1992) based on psychiatric interview but for day-to-day practice, questionnaires to detect depression in adolescents using these criteria have been designed (Angold et al., 1995a).

Despite its importance, MDD in adolescence is still often unrecognised or untreated (Kramer et al., 1998; Coyle et al., 2003; Potter etal., 2012). The offspring of parents with recurrent MDD are at particular risk for developing more serious and impairing depression in adolescence but are less likely to go for treatment compared to offspring of non-depressed parents (Weissman et al., 1997). Key factors may include the reluctance of many adolescents to present psychological problems to a medical professional (Mauerhofer et al., 2009) and parental factors (Wu et al., 1999; Festen et al., 2014). Depressed parents are often well known to services (Hutton and Gunn, 2007; Institute of Medicine, 2009) and accurate detection of disorder is likely to be enhanced in this group of adolescents given that depression prevalence rates are elevated (Institute of Medicine, 2009; Kent et al., 1997; Wood et al., 1995). Moreover, the importance of early interventions in young people at high familial risk of depression has been highlighted

(Beardslee et al., 2013). Thus the question of how access to assessment and treatment services can be improved in adolescents at high risk for early-onset, impairing depressive disorder could lie in the parents of these adolescents: Whilst it is often regarded that depression is best rated by the adolescent themselves, particularly for older adolescents, there is strong evidence to support the validity of adolescent symptoms as reported by parents, most often mothers, (Rice et al., 2007; De Los Reyes and Kazdin, 2005) and both maternal and adolescent reports of adolescent depressive symptoms show association with longer-term functional outcome in adolescents (Rice et al., 2007). Thus, information provided by depressed parents about their offspring's depressive symptoms seems to be a reliable indicator of underlying depressive disorder (Lewis et al., 2012). There is also evidence to suggest that depressed parents may be especially sensitive at identifying depressive symptomatology in their children (Richters and Pelligrini, 1989; Weissman et al., 1987) suggesting that they may provide valuable information about their child's mental health

Given the importance of time constraints in both primary and secondary care (Linzer et al., 2000; Hutton and Gunn, 2007; Konrad et al., 2010) a strategy to reduce time outlay on detecting depression would seem to be worthwhile. Such strategies would need to lead to an accurate diagnosis given the concerns of parents and adolescents about stigma and other impacts of labelling (Fenten et al., 2014). Despite some promising results (Kroenke et al., 2003; Richardson et al., 2010) the accuracy of ultra-short (one/two item) question screens in detecting depression has been questioned (Mitchell and Coyne, 2007). Indeed, full questionnaires (9 or more item) generally seem to perform better both in adults (Thapar et al., 2014; although see Zimmerman et al., 2010) and adolescents (Allgaier et al., 2012; Rhew et al., 2010). Despite the fact that

parental reports of adolescent depressive symptoms may be useful in overcoming some of the barriers to accessing services for adolescents at high-risk of depression, the use of abbreviated screens to detect adolescent depression in this group with parent reports has yet to be evaluated.

The present study utilised a completed three-wave study of the high-risk adolescent offspring of parents with recurrent MDD to investigate the following primary aims:

1) Identify the most parsimonious combination of parent-reported mental health symptoms that accurately detects a diagnosis of major depressive disorder (MDD) in their high-risk offspring, assessed by gold standard direct interviews.

2) Examine whether this parent-reported combination of symptoms is as effective in detecting depression as a well validated, widely used 13-item adolescent depression questionnaire (the short MFQ).

A secondary aim was to replicate findings in an independent dataset, for a similarly defined subgroup of adolescents with a recurrently depressed parent.

2.Methods

2.1. Subjects

Two samples were used in this analysis: one to identify the combination of items and a second, independent sample to validate the findings.

1) *Development dataset*: The Early Prediction of Adolescent Depression (EPAD) study. This was a prospective, longitudinal study of 337 high-risk offspring of recurrently depressed parents

identified mainly from UK Primary Care (Mars et al., 2012). Parent recurrent depression was defined as at least two episodes of DSM-IV defined MDD, later confirmed using a diagnostic interview. Information from parents and children were simultaneously collected prospectively at three time points approximately 12-16 months apart.

2) *Validation dataset*: The Avon Longitudinal Study of Parents and Children (ALSPAC) is a birth-cohort set up to examine genetic and environmental determinants of health and development (Boyd et al., 2013). The initial cohort consisted of 14,541 pregnant women resident in the former county of Avon, United Kingdom who had an expected date of delivery between 1st April 1991 and 31st December 1992 (www.alspac.bris.ac.uk). For information on all available ALSPAC data see the fully searchable data dictionary (http://www.bris.ac.uk/alspac/researchers/data-access/data-dictionary). A sub-sample was selected comprising offspring of mothers who had reported recurrent depression in the past on at least 2 separate time points during the lifetime of the child with at least one of these being self-rated as severe with complete data on the relevant assessments. Ethical approval for the study was obtained from the ALSPAC Ethics and Law Committee and the Local Research Ethics Committees and the Wales Multicentre Research Ethics Committee.

2.2. Measures

2.2.1. Parent-report symptom screens

We used parent rated information on offspring mental health from the 25-item *Strengths and Difficulties Questionnaire (SDQ)* (Goodman, 1997)and the 13-item *Short Moods and Feelings Questionnaire (sMFQ)* (Angold et al., 1995a) at the three data points in the development dataset

(concurrent interview data available) and at the time point in the validation dataset closest to when the diagnostic assessment was performed.

The sMFQ has been shown to be a valid measure to detect depression (Angold et al., 1995; Thapar and McGuffin, 1998; Kuo et al., 2005). Parents rated whether their child had a particular symptom either over the previous three months (development dataset) or over the previous 2 weeks (validation dataset) using a three point scale ("not true", "somewhat true" and "definitely true").

The SDQ is a more general measure of child psychopathology (with items on mood, behaviour and social difficulties). Parents rated whether or not their child had had these symptoms, behaviours or difficulties over the previous 6 months using a three-point scale ("not true", "somewhat true" and "certainly true").

For all items on both questionnaires both "true" responses were merged. This was done in order to simplify the questions, reduce respondent burden, align with standard clinical and diagnostic practice where symptoms are usually viewed as present or absent (Angold et al., 2000) as well as to aid interpretation of odds ratios.

2.2.2. Childhood research diagnostic interviews

2.2.2.1. The Child and Adolescent Psychiatric Assessment (CAPA) (Angold et al., 2000) is a semi-structured psychiatric diagnostic interview and used in the development dataset. Both child and parent were separately interviewed. For this analysis, symptoms endorsed by either child or parent were used to generate DSM-IV diagnoses that were further checked by clinical consensus (reviewed by two consultant child and adolescent psychiatrists).

2.2.2.2. The Development and Well-Being Assessment (DAWBA (Goodman et al., 2000) is a semi- structured diagnostic instrument completed by parents and was used in the validation dataset when the child was aged 13 years. The results are expressed as a probability of the child having MDD and, as recommended, all those with a >50% probability were coded as having MDD (Goodman et al., 2011).

2.3. Analysis

2.3.1. Development dataset

Taken together, the parent-reported SDQ and sMFQ included a total of 38 potential mental health items. After initial univariate analysis (to identify a sub-sample of items significantly associated with offspring MDD diagnosis at every wave- (see Supplementary Table 3)), multivariate analysis (forward stepwise regression) was then used to identify significant predictors of MDD at each wave. Finally the best performing 4 items were selected ("combination") based on the consistency of associations across all or most waves and the highest odds ratio if only predictive at a single wave. The performance of this 4-item combination for detecting concurrent MDD (as the outcome), was compared to that for the sMFQ at each time point using ROC curve analysis.. Comparative **predictive validity** of the combination to the sMFQ was also examined, using MDD at a future time point (irrespective of baseline status) as the outcome.

2.3.2. Validation dataset

To examine whether the pattern of findings from the development dataset could be **replicated** the performance of this combination score (derived from the same items as in the development dataset) was compared to the sMFQ score for detecting concurrent MDD (derived from the DAWBA assessment) using ROC curve analysis.

The Youden index was used to determine the optimal cut-off from ROC curve analysis. The areas under the curve (AUC, C-index) for the combination and for sMFQ were compared using the STATA "roccomp" option. SPSS Version 20 and STATA version 13 were used for analyses.

3. Results

3.1. Development dataset

For figures detailing recruitment and retention see Supplementary Material. The eligible sample at baseline consisted of 335 high-risk offspring of recurrently depressed parents (Lewis et al., 2012, Mars et al., 2012). 313 of these parents were mothers (93.4%) and 22 were fathers (6.6%). Three hundred and twenty six individuals (97.3%) provided complete interview and questionnaire data at least at one time point. Of these 326 individuals, 287 (88.0%) provided complete interview and questionnaire data for at least 2 time points. The mean age at baseline (Wave1) was 12.4 years (SD 1.9 years) with a range of 9-17 years, 56.7% were female and 27 adolescents (21 girls and 6 boys) met criteria for MDD. At Wave 2 the mean age was 13.7 years

(SD 2.0 years), 59.9% were female and 26 (20 girls and 6 boys) had MDD. At Wave 3 the mean age was 14.8 years (SD 2.0 years), 58.2% were female and 28 (22 girls and 6 boys) had MDD. In univariate tests of association between individual parent-reported mental health items and MDD diagnosis, 17 (6 from the SDQ and 11 from the MFQ) of the 38 items on both questionnaires were significantly associated with interview-confirmed MDD diagnosis at all three waves (Supplementary Material-Table 3). These 17 items were then entered into a forward stepwise multivariate regression. Only 7 of these items were associated with MDD as diagnosed by clinical interview at any interview phase (see Table 1). The incremental improvement in Nagelkerke R^2 resulting from the addition of each item for each wave in the development dataset is detailed in Supplementary Material 4.

The four items selected to be included in the final combination were on the basis of consistency of associations across all or most waves ("hard to think properly or concentrate" (from sMFQ) and "didn't enjoy anything at all" (from sMFQ)) or from the magnitude of the odds ratio at a single wave ("many worries…often seems worried" (from SDQ-first wave) and "thought nobody really loved him/her" (from sMFQ-third wave)).

Using ROC curve analysis, the combination was compared to the sMFQ with concurrent adolescent MDD as the outcome variable. The results for a range of scores are presented in Table 2 and graphically in Figure 1.

3.2. Validation dataset

In this sample of 807 offspring of mothers who had recurrent depression with complete data on sMFQ and DAWBA (53% female, 47% male, mean age 13.1 years), 24 adolescents (12 boys and 12 girls) met criteria for MDD (depression point prevalence of 2.7%).

ROC curve analysis was again used to compare the combination derived from the validation dataset to the sMFQ using a diagnosis of MDD (derived from the DAWBA interview) as the outcome variable (Table 3). Similar results were obtained for the combination and sMFQ.

3.3. Predictive validity.

Using the development dataset, the combination was compared to sMFQ for performance in detecting all those individuals with depression at a future time point (irrespective of baseline status (Tables 4 and 5). The combination and the sMFQ were able to detect later MDD with similar levels of sensitivity and specificity.

4.Discussion

A combination of four parent-reported, mental health items performed as well as a well-established, widely used depression screening questionnaire (sMFQ) in detecting both concurrent and future adolescent MDD amongst offspring of parents with recurrent depression. These findings were replicated in an external dataset using similar criteria for selection of parents.

We found that, for adolescent offspring of depressed parents, if at least two out of four parent-reported mental health symptoms were present then MDD could be detected as accurately as by using an established longer depression questionnaire. Conversely if less than two of these four symptoms were present then MDD was highly unlikely. There have been no previous studies reporting on briefer methods for detecting a likely depression diagnosis in adolescents when it is their parents who report on symptoms. In adolescents from the general population some studies have reported on the use of shorter depression questionnaires, such as the PHQ-2 (Kroenke et al., 2003), using self-report in either hospital or out-patient settings (Richardson et al., 2010; Allgaier et al., 2012) with adequate performance as a depression screen but lower accuracy than longer measures (Allgaier et al., 2012; Rhew et al., 2010) for detecting current depression.

The four items we found in the best performing combination were anhedonia, poor concentration, excessive worrying and feeling unloved. Some of these items map well onto findings from neurobiological research highlighting the importance of reward- and fear-related brain circuitry in depression as well as cognitive deficits (Thapar et al., 2012; Hasler et al., 2004) but yet are under-recognised in their role in detecting depression. Interestingly, a study focusing on the mothers of the children in our development dataset found that anhedonia and poor concentration (along with low mood and restlessness) accurately detected adult depressive disorder relapse (Thapar et al., 2014). The most consistent findings from other studies on depression symptom clusters also highlight concentration problems (Sund et al., 2001; Cole et al., 2011). Many studies which have examined this topic are either wholly questionnaire based or have used a non-validated psychiatric interview. The heterogeneity of depression (Hasler et al., 2004) and age, gender and cultural case mix may also be relevant (Fu-I andWang, 2008).

Thus it may be that this combination of items performed well in detecting MDD in part because it assesses several disparate aspects of depressive symptomatology.

In this study we used parent ratings of mental health in their offspring to detect depression. Parents with recurrent depression seem to provide accurate ratings of mental health symptoms in their offspring and these ratings seem at least as accurate as self-ratings by the child themselves in detecting depression (Rice et al., 2007; Lewis et al., 2012) although child reports are needed to assess suicidality (Rice et al., 2007). Moreover these offspring are an easily identifiable risk group and preventive forms of psychological therapy and psycho-education have shown efficacy in preventing the onset of MDD and in improving understanding about depression and reducing symptoms of distress respectively (Garber et al., 2009; Beardslee et al., 2003). Either asking the parent about their offspring's mental health or making parents aware of key symptoms to look out for in their adolescent offspring would seem to offer the advantage of working in partnership with the parent to determine the most appropriate method of engaging with and helping their child (Wu et al., 1999; Festen et al., 2014). Working sensitively in partnership with parents is likely to be important given that parents can sometimes act as a barrier to children with depression accessing services (Wu et al., 1999). Conversely if major depressive disorder can be confidently excluded this may again ensure care is appropriate.

4.1. Strengths and limitations

The original study was a community based high-risk longitudinal study with comprehensive questionnaire assessment of adolescent health and functioning as well as a detailed, validated psychiatric interview so that accurate diagnoses of psychopathology could be made. High

completion rates were noted with over 97% of the eligible sample providing at least one set of complete interview and questionnaire data and over 85% of the eligible sample provided complete information from at least two time points. The findings of the analysis were consistent across all three waves of this dataset at different mean ages of the adolescent. A major strength of the analysis reported in this paper was that the findings were replicated using data from a second, large independent sample. Moreover the validation dataset used a different psychiatric assessment to generate child/adolescent diagnosis of MDD giving additional evidence of validity of the findings. High values for the negative predictive value for scores below the cut-off also highlight the value of this measure in excluding depression.

However several limitations also need to be noted. Only modest values for positive predictive value (PPV) for our combination were obtained. However these results are not only similar to those for the sMFQ, but also to those from other studies (Fu-I andWang, 2008; Katon et al., 2008). The prevalence of MDD was modest for both datasets (hence the PPV values) but is likely to relate to the age of the samples as many adolescents had not passed the age of maximum risk for depression (Weissman et al., 1997). Moreover the rates are higher than the prevalence of depression at this age found in large community surveys (Ford et al., 2003). Moreover we used point prevalence rates which may underestimate burden as depressive disorder in adolescence is marked by relapses and remissions (Thapar et al., 2012). The lower prevalence of adolescent depression in the validation dataset than in development dataset could be related to differences in parental depression severity, selective attrition as well as difference in diagnostic methods. However the consistent pattern of results across the different samples reinforces the validity of the findings. As with all longitudinal studies there was also some loss to follow-up in the development dataset and baseline depression scores were slightly higher amongst subsequent

non-responders. However the findings were consistent across waves and loss to any follow up was small as noted above. There was also a nine month gap between questionnaires and the DAWBA interview in the validation dataset. However predictive validity was demonstrated over longer time periods (see Table 4) suggesting that this was not a major issue. Finally the 4 items selected for the scale were embedded in longer questionnaires so one cannot be certain about the performance of these items if administered in a separate scale.

In conclusion, a combination of four parent reported mental health items appears to accurately detect concurrent and future MDD in the adolescent offspring of parents with recurrent depression. These mental health items perform as well as an established, longer depression questionnaire in correctly identifying MDD. These findings require further replication to establish it is these mental health items rather than the measures from which they were derived that are crucial. If this is the case, parents can be asked or advised about these items verbally rather than by using "paper and pen" measures. Such an approach may in future offer the valuable opportunity of incorporating rapid parent-based screening for adolescent depressive disorder and allow more time for a more comprehensive mental health screen with adolescents that could be used by those responsible for providing mental health care for families where the parent has a history of recurrent, severe depression.

Acknowledgments

We thank all the families who took part in the EPAD study, the GP surgeries for their help in recruiting them and the whole EPAD study team.

We thank the Sir Jules Thorn Charitable Trust for funding the original cohort study, NISCHR AHSC who funded research time for AKT and Waterloo Foundation for part-funding SC.

We thank Michael Owen (Cardiff University), Gordon Harold (Sussex University), Danny Smith (Glasgow University) and all the research team members for their contribution to the original study and data collection.

We are extremely grateful to all the families who took part in the ALSPAC study, the midwives for their help in recruiting them, and the whole ALSPAC team, which includes interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists and nurses. The UK Medical Research Council and the Wellcome Trust (Grant ref: 092731) and the University of Bristol provide core support for ALSPAC.

Conflict of interest

None declared for all authors.

References

Allgaier, A.-K., Pietsch, K., Frühe, B., Sigl-Glöckner, J., Schulte-Körne, G., 2012. Screening for depression in adolescents: validity of the patient health questionnaire in pediatric care. Depress. Anxiety 29, 906–13. doi:10.1002/da.21971

American Psychiatric Association, 2013. Diagnostic and Statistical Manual of Mental Disorders, 4th. Edition, Arlington VA.

Angold, A., Costello, E. J., Messer, S. C., Pickles, A., Winder, F., & Silver, D.1995.The development of a short questionnaire for use in epidemiological studies of depression in children and adolescents. Int J Methods Psychiatr Res. 5, 237 - 249.

Angold, A., Costello, E.J., 2000. The Child and Adolescent Psychiatric Assessment (CAPA). J. Am. Acad. Child Adolesc. Psychiatry 39, 39–48. doi:10.1097/00004583-200001000-00015

Beardslee, W.R., Brent, D.A., Weersing, V.R., Clarke, G.N., Porta, G., Hollon, S.D., Gladstone, T.R.G., Gallop, R., Lynch, F.L., Iyengar, S., DeBar, L., Garber, J., 2013. Prevention of depression in at-risk adolescents: longer-term effects. JAMA psychiatry 70, 1161–70. doi:10.1001/jamapsychiatry.2013.295

Boyd, A., Golding, J., Macleod, J., Lawlor, D.A., Fraser, A., Henderson, J., Molloy, L., Ness, A., Ring, S., Davey Smith, G., 2013. Cohort Profile: the “children of the 90s”--the index offspring of the Avon Longitudinal Study of Parents and Children. Int. J. Epidemiol. 42, 111–27. doi:10.1093/ije/dys064

Cole, D.A., Cai, L., Martin, N.C., Findling, R.L., Youngstrom, E.A., Garber, J., Curry, J.F., Hyde, J.S., Essex, M.J., Compas, B.E., Goodyer, I.M., Rohde, P., Stark, K.D., Slattery, M.J., Forehand, R., 2011. Structure and measurement of depression in youths: applying item response theory to clinical data. Psychol. Assess. 23, 819–33. doi:10.1037/a0023518

Coyle, J.T., Pine, D.S., Charney, D.S., Lewis, L., Nemeroff, C.B., Carlson, G.A., Joshi, P.T., Reiss, D., Todd, R.D., Hellander, M., Depression and Bipolar Support Alliance Consensus Development Panel, 2003. Depression and bipolar support alliance consensus statement on the unmet needs in diagnosis and treatment of mood disorders in children and adolescents. J. Am. Acad. Child Adolesc. Psychiatry 42, 1494–503. doi:10.1097/01.chi.0000091945.28938.8f

De Los Reyes, A., Kazdin, A.E., 2005. Informant discrepancies in the assessment of childhood psychopathology: a critical review, theoretical framework, and recommendations for further study. Psychol. Bull. 131, 483–509. doi:10.1037/0033-2909.131.4.483

Festen, H., Schipper, K., de Vries, S.O., Reichart, C.G., Abma, T.A., Nauta, M.H., 2014. Parents' perceptions on offspring risk and prevention of anxiety and depression: a qualitative study. BMC Psychol. 2, 17. doi:10.1186/2050-7283-2-17

Ford, T., Goodman, R., Meltzer, H., 2003. The British Child and Adolescent Mental Health Survey 1999: the prevalence of DSM-IV disorders. J. Am. Acad. Child Adolesc. Psychiatry 42, 1203–11. doi:10.1097/00004583-200310000-00011

Fu-I, L., Wang, Y.P., 2008. Comparison of demographic and clinical characteristics between children and adolescents with major depressive disorder. Rev. Bras. Psiquiatr. (São Paulo, Brazil 1999) 30, 124–31.

Goodman, A., Heiervang, E., Collishaw, S., Goodman, R., 2011. The "DAWBA bands" as an ordered-categorical measure of child mental health: description and validation in British and Norwegian samples. Soc. Psychiatry Psychiatr. Epidemiol. 46, 521–32. doi:10.1007/s00127-010-0219-x

Goodman, R., 1997. The Strengths and Difficulties Questionnaire: a research note. J. Child Psychol. Psychiatry. 38, 581–6.

Goodman, R., Ford, T., Richards, H., Gatward, R., Meltzer, H., 2000. The Development and Well-Being Assessment: description and initial validation of an integrated assessment of child and adolescent psychopathology. J. Child Psychol. Psychiatry. 41, 645–55.

Hasler, G., Drevets, W.C., Manji, H.K., Charney, D.S., 2004. Discovering endophenotypes for major depression. Neuropsychopharmacology 29, 1765–81. doi:10.1038/sj.npp.1300506

Hutton, C., Gunn, J., 2007. Do longer consultations improve the management of psychological problems in general practice? A systematic literature review. BMC Health Serv. Res. 7, 71. doi:10.1186/1472-6963-7-71

Institute of Medicine., 2009. Depression in Parents, Parenting, and Children: Opportunities to Improve Identification, Treatment, and Prevention, Consensus report.

Katon, W., Russo, J., Richardson, L., McCauley, E., Lozano, P., n.d. Anxiety and depression screening for youth in a primary care population. Ambul. Pediatr. 8, 182–8. doi:10.1016/j.ambp.2008.01.003

Kent, L., Vostanis, P., Feehan, C., 1997. Detection of major and minor depression in children and adolescents: evaluation of the Mood and Feelings Questionnaire. J. Child Psychol. Psychiatry. 38, 565–73.

Konrad, T.R., Link, C.L., Shackelton, R.J., Marceau, L.D., von dem Knesebeck, O., Siegrist, J., Arber, S., Adams, A., McKinlay, J.B., 2010. It's about time: physicians' perceptions of time constraints in primary care medical practice in three national healthcare systems. Med. Care 48, 95–100. doi:10.1097/MLR.0b013e3181c12e6a

Kramer, T., Garralda, M.E., 1998. Psychiatric disorders in adolescents in primary care. Br. J. Psychiatry 173, 508–13.

Kroenke, K., Spitzer, R.L., Williams, J.B.W., 2003. The Patient Health Questionnaire-2: validity of a two-item depression screener. Med. Care 41, 1284–92. doi:10.1097/01.MLR.0000093487.78664.3C

Kubik, M.Y., Lytle, L.A., Birnbaum, A.S., Murray, D.M., Perry, C.L., n.d. Prevalence and correlates of depressive symptoms in young adolescents. Am. J. Health Behav. 27, 546–53.

Kuo, E.S., Stoep, A. Vander, Stewart, D.G., 2005. Using the short mood and feelings questionnaire to detect depression in detained adolescents. Assessment 12, 374–83. doi:10.1177/1073191105279984

Lewis, K.J.S., Mars, B., Lewis, G., Rice, F., Sellers, R., Thapar, A.K., Craddock, N., Collishaw, S., Thapar, A., 2012. Do parents know best? Parent-reported vs. child-reported depression symptoms as predictors of future child mood disorder in a high-risk sample. J. Affect. Disord. 141, 233–6. doi:10.1016/j.jad.2012.03.008

Leydon, G.M., Dowrick, C.F., McBride, A.S., Burgess, H.J., Howe, A.C., Clarke, P.D., Maisey, S.P., Kendrick, T., QOF Depression Study Team, 2011. Questionnaire severity measures for depression: a threat to the doctor-patient relationship? Br. J. Gen. Pract. 61, 117–23. doi:10.3399/bjgp11X556236

Linzer, M., Konrad, T.R., Douglas, J., McMurray, J.E., Pathman, D.E., Williams, E.S., Schwartz, M.D., Gerrity, M., Scheckler, W., Bigby, J.A., Rhodes, E., 2000. Managed care, time pressure, and physician job satisfaction: results from the physician worklife study. J. Gen. Intern. Med. 15, 441–50.

Lopez AD, Mathers CD, Ezzati M, Jamison DT, Murray CJL, editors. 2006. Global burden of disease and risk factors. Washington: Oxford University Press and the World Bank.

Mars, B., Collishaw, S., Smith, D., Thapar, A., Potter, R., Sellers, R., Harold, G.T., Craddock, N., Rice, F., Thapar, A., 2012. Offspring of parents with recurrent depression: which features of parent depression index risk for offspring psychopathology? J. Affect. Disord. 136, 44–53. doi:10.1016/j.jad.2011.09.002

Mauerhofer, A., Berchtold, A., Michaud, P.-A., Suris, J.-C., 2009. GPs' role in the detection of psychological problems of young people: a population-based study. Br. J. Gen. Pract. 59, e308–14. doi:10.3399/bjgp09X454115

Mitchell, A.J., Coyne, J.C., 2007. Do ultra-short screening instruments accurately detect depression in primary care? A pooled analysis and meta-analysis of 22 studies. Br. J. Gen. Pract. 57, 144–51.

Potter, R., Mars, B., Eyre, O., Legge, S., Ford, T., Sellers, R., Craddock, N., Rice, F., Collishaw, S., Thapar, A., Thapar, A.K., 2012. Missed opportunities: mental disorder in children of parents with depression. Br. J. Gen. Pract. 62, e487–93. doi:10.3399/bjgp12X652355

Rhew, I.C., Simpson, K., Tracy, M., Lymp, J., McCauley, E., Tsuang, D., Stoep, A. Vander, 2010. Criterion validity of the Short Mood and Feelings Questionnaire and one- and two-item depression screens in young adolescents. Child Adolesc. Psychiatry Ment. Health 4, 8. doi:10.1186/1753-2000-4-8

Rice, F., Lifford, K.J., Thomas, H. V, Thapar, A., 2007. Mental health and functional outcomes of maternal and adolescent reports of adolescent depressive symptoms. J. Am. Acad. Child Adolesc. Psychiatry 46, 1162–70. doi:10.1097/chi.0b013e3180cc255f

Richardson, L.P., Rockhill, C., Russo, J.E., Grossman, D.C., Richards, J., McCarty, C., McCauley, E., Katon, W., 2010. Evaluation of the PHQ-2 as a brief screen for detecting major depression among adolescents. Pediatrics 125, e1097–103. doi:10.1542/peds.2009-2712

Rushton, J.L., Forcier, M., Schectman, R.M., 2002. Epidemiology of depressive symptoms in the National Longitudinal Study of Adolescent Health. J. Am. Acad. Child Adolesc. Psychiatry 41, 199–205. doi:10.1097/00004583-200202000-00014

Sund, A.M., Larsson, B., Wichstrom, L., 2001. Depressive symptoms among young Norwegian adolescents as measured by the Mood and Feelings Questionnaire (MFQ). Eur. Child Adolesc. Psychiatry 10, 222–9.

Thapar, A., McGuffin, P., 1998. Validity of the shortened Mood and Feelings Questionnaire in a community sample of children and adolescents: a preliminary research note. Psychiatry Res. 81, 259–68.

Thapar, A., Hammerton, G., Collishaw, S., Potter, R., Rice, F., Harold, G., Craddock, N., Thapar, A., Smith, D.J., 2014. Detecting recurrent major depressive disorder within primary care rapidly and reliably using short questionnaire measures. Br. J. Gen. Pract. 64, e31–7. doi:10.3399/bjgp14X676438

Thapar, A., Collishaw, S., Pine, D.S., Thapar, A.K., 2012. Depression in adolescence. Lancet (London, England) 379, 1056–67. doi:10.1016/S0140-6736(11)60871-4

von Elm, E., Altman, D.G., Egger, M., Pocock, S.J., Gøtzsche, P.C., Vandenbroucke, J.P., STROBE Initiative, 2007. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Ann. Intern. Med. 147, 573–7.

Weissman, M.M., Warner, V., Wickramaratne, P., Moreau, D., Olfson, M., 1997. Offspring of depressed parents. 10 Years later. Arch. Gen. Psychiatry 54, 932–40.

Wood, A., Kroll, L., Moore, A., Harrington, R., 1995. Properties of the mood and feelings questionnaire in adolescent psychiatric outpatients: a research note. J. Child Psychol. Psychiatry. 36, 327–34.

World Health Organization., 1992. The ICD-10 classification of mental and behavioural disorders: Clinical descriptions and diagnostic guidelines. Geneva: World Health Organization.

Wu, P., Hoven, C.W., Bird, H.R., Moore, R.E., Cohen, P., Alegria, M., Dulcan, M.K., Goodman, S.H., Horwitz, S.M., Lichtman, J.H., Narrow, W.E., Rae, D.S., Regier, D.A., Roper, M.T., 1999. Depressive and disruptive disorders and mental health service utilization in children and adolescents. J. Am. Acad. Child Adolesc. Psychiatry 38, 1081–90; discussion 1090–2. doi:10.1097/00004583-199909000-00010

Zimmerman, M., Galione, J.N., Chelminski, I., McGlinchey, J.B., Young, D., Dalrymple, K., Ruggero, C.J., Witt, C.F., 2010. A simpler definition of major depressive disorder. Psychol. Med. 40, 451–7. doi:10.1017/S0033291709990572

Table 1 : Results of multivariate(MV) forward stepwise logistic regression using concurrent adolescent MDD as the outcome at each Wave.

	Wave 1(*n=316*) R^2 =0.33		Wave 2(*n=253*) R^2 =0.32		Wave 3(*n=268*) R^2 =0.39	
Significant symptoms from forwards conditional stepwise regression	Main MV Predictor in this wave?	Odds ratio (95% confidence intervals	Main MV Predictor in this wave?	Odds ratio (95% confidence intervals	Main MV Predictor in this wave?	Odds ratio (95% confidence intervals)
Lack of enjoyment	Yes	*3.1 (1.2-8.0)*	Yes	*3.4 (1.3-9.1)*	No	
Problems with concentration	Yes	*3.1 (1.1-8.4)*	Yes	*6.6 (2.0-21.6)*	Yes	*3.8 (1.3-10.9)*
Problems with excessive worrying	Yes	*9.4 (1.2-73.8)*	No		No	
Viewed themselves as being not as good as others	No		Yes	*3.2 (1.2-8.4)*	No	
Tiredness	Yes	*3.3 (1.1-9.9)*	No		No	
Loneliness	No		No		Yes	*2.9 (1.1-8.2)*
Feeling unloved	No		No		Yes	*8.4 (3.0- 23.1)*

Table 2: Comparison of the four-item combination of symptoms to the established short Moods and Feelings Questionnaire (sMFQ) at three waves using different cut off points for the detection of Major Depressive Disorder in the development dataset.

Measure (no.) Cut off	Sensitivity %	Specificity %	PPV %	NPV %	LR+	LR-	AUC (95%CI) Comparing ROC curves
WAVE 1	**(% MDD=8.2%, n=26)**						**AUC(sMFQ)=AUC(combination) *X^2 = 0.18 p=0.67***
13 item sMFQ (*n=316*)							***0.83 (0.76- 0.91)***
≥4	92.3	56.6	16.0	98.8	2.12	0.14	
≥5#	88.5	62.8	17.6	98.4	2.38	0.18	
≥6*	76.9	69.7	18.5	97.1	2.54	0.33	
≥7	76.9	74.1	21.0	97.3	2.97	0.31	
≥8⌘	61.5	81.0	22.5	95.9	3.24	0.47	
≥9	57.7	83.8	24.2	95.7	3.56	0.50	
≥10	53.8	87.9	28.5	95.5	4.46	0.52	
≥11	50.0	90.0	31.0	95.2	5.00	0.56	
4 item combination(*n=316*)							***0.84 (0.77—0.92)***
≥1	96.2	33.4	11.5	99.0	1.44	0.11	
≥2#*	92.3	62.4	18.0	99.0	2.46	0.12	
≥3	65. 4	85.5	28.8	96.5	4.51	0.40	
≥4	38.5	95.5	43.4	94.5	8.58	0.64	
WAVE 2	**(% MDD=9.4%, n=24)**						**AUC(sMFQ)=AUC(combination) *X^2 = 0.24 p=0.73***
13 item sMFQ (*n=256*)							***0.81 (0.73- 0.89)***
≥4	87.5	63.8	20.0	98.0	2.42	0.20	
≥5	79.2	69.8	21.3	97.0	2.62	0.30	
≥6#*	75.0	75.4	24.0	96.7	3.05	0.33	
≥7	62.5	78.9	23.5	95.3	2.96	0.48	
≥8⌘	58.3	85.3	29.1	95.2	3.99	0.49	
≥9	41.7	88.4	27.1	93.6	3.59	0.66	
≥10	41.7	91.4	33.4	93.8	4.83	0.64	
≥11	33.3	92.2	30.6	93.0	4.30	0.72	
4 item combination (*n=256*)							***0.80 (0.71- 0.90)***
≥1	91.7	37.5	13.2	97.8	1.47	0.22	
≥2*#	83.3	69.8	22.2	97.6	2.76	0.24	
≥3	50.0	86.2	27.3	94.3	3.62	0.58	
≥4	33.3	97.4	57.0	93.4	12.89	0.68	
WAVE 3	**(% MDD=9.7%, n =26)**						**AUC(sMFQ)=AUC(combination) *X^2 = 0.92 p=0.34***
13 item sMFQ (*n=268*)							***0.89 (0.81- 0.95)***
≥4	88.5	67.4	22.6	98.2	2.71	0.17	
≥5	84.6	77.7	29.0	97.9	3.79	0.20	
≥6*#	80.8	83.2	34.1	97.6	4.89	0.23	
≥7	73.1	85.5	35.1	96.7	5.05	0.31	
≥8⌘	73.1	88.0	39.6	96.8	6.10	0.31	
≥9	69.2	90.9	44.0	96.5	7.62	0.34	
≥10	61.5	93.0	48.6	95.7	8.76	0.41	
≥11	61.5	94.6	55.0	95.8	11.46	0.41	
4 item combination (*n=268*)							***0.86 (0.80- 0.94)***
≥1	100.0	45.4	16.4	100	1.83	0.00	
≥2*	80.8	71.1	25.0	98.5	2.79	0.27	
≥3#	65.4	88.8	38.5	96.0	5.86	0.39	
≥4	34.6	97.1	56.2	93.3	11.97	0.67	

* overall cut point for defining a DSM diagnosis of depressive disorder based on overall highest Youden index (sensitivity + specificity -1) across all 3 waves
#highest Youden index for each wave ⌘ cut-off recommended in literature [7]

Figure 1: ROC curve analysis comparing the four-item combination of symptoms to the established Short Moods and Feelings Questionnaire (sMFQ) at Wave 1 for the detection of Major Depressive Disorder in the development dataset.

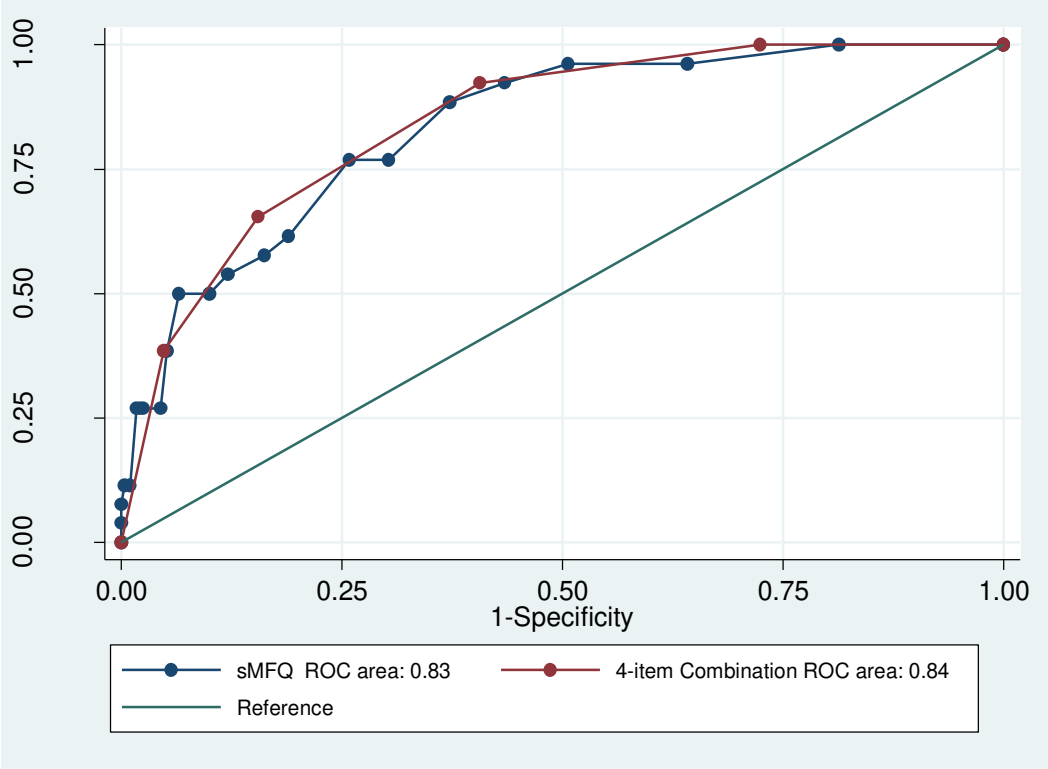


Table 3 : Comparison of the four-item combination of symptoms to the established Short Moods and Feelings questionnaire (sMFQ) using different cut off points for the detection of Major Depressive Disorder in the validation dataset.

Measure (no.) Cut off	Sensitivity %	Specificity %	PPV %	NPV %	LR+	LR-	**AUC (95%CI) Comparing ROC curves**
(% MDD=2.73% n=24)							**AUC(sMFQ)=AUC(combination)** ***X^2= 0.17 p=0.7***
13 item sMFQ (*n=807*)							***0.83 (0.76- 0.91)***
≥4	86.4	66.5	6.7	99.4	2.58	0.21	
≥5#	86.4	73.9	8.5	99.5	3.31	0.18	
≥6*	68.2	79.2	8.4	98.9	3.28	0.40	
≥7	63.6	83.2	9.6	98.8	3.78	0.44	
≥8⌘	59.1	86.4	10.9	98.7	4.34	0.47	
≥9	50.0	88.9	11.2	98.4	4.51	0.56	
≥10	50.0	90.3	12.6	98.5	5.16	0.55	
≥11	40.9	92.4	13.1	98.2	5.35	0.64	
4 item combination (*n=798*)							***0.82 (0.74- 0.90)***
≥1	95.2	48.5	4.9	99.7	1.85	0.1	
≥2#*	76.2	73.9	7.6	99.1	2.92	0.32	
≥3	52.4	86.6	9.9	98.5	3.91	0.55	
≥4	38.1	94.3	15.6	98.2	6.73	0.66	

*optimal cut point for defining a DSM diagnosis of depressive disorder based on highest Youden index in original sample (sensitivity + specificity -1) and specificity ≥70%.⌘ cut-off recommended in literature

#highest Youden index for each wave

Table 4: Comparison of the four-item combination of symptoms to the established Short Moods and Feelings Questionnaire (sMFQ) to predict later (persistent or new onset) Major Depressive Disorder in future waves in the development dataset.

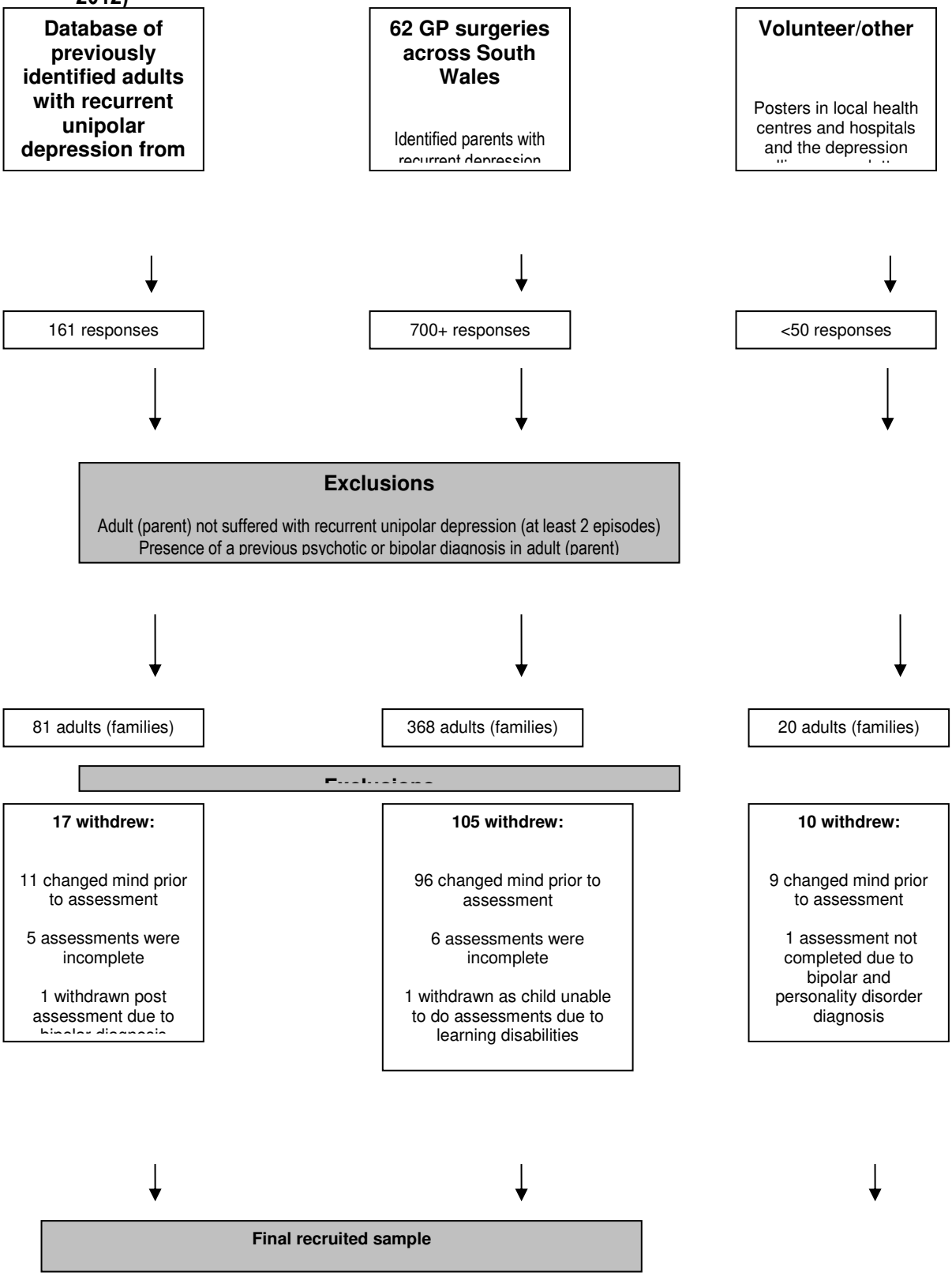
	Wave 2 MDD			**Wave 3 MDD**		
	Area under curve	**95% CI**		**Area under curve**	**95% CI**	
Wave 1 4-item combination	0.73	0.64	0.82	0.76	0.66	0.86
Wave 1 sMFQ	0.75	0.66	0.83	0.81	0.73	0.89
AUC sMFQ= AUC combination	*n=279, X^2 = 0.45 p=0.50*			*N=273, X^2 = 3.45 p=0.06*		
Wave 2 4-item combination				0.71	0.60	0.82
Wave 2 sMFQ				0.71	0.61	0.82
AUC sMFQ= AUC combination				*N=253, X^2 = 0.01 p=0.93*		

Table 5: Cut-off points for four-item combination of symptoms and Short Moods and Feelings Questionnaire (sMFQ) measured at Wave 1 for detecting Wave 3 Major Depressive Disorder in the development dataset.

Measure (no.)	Sensitivity %	Specificity %	PPV %	NPV %	LR+	LR-	**AUC (95%CI) Comparing ROC curves**
WAVE 3	**(% MDD=9.70%)**						**AUC(sMFQ)=AUC(combination)** ***$X^2 = 3.45$ p=0.06***
Wave 1 13 item sMFQ ***(n=273)***							***0.81 (0.73- 0.89)***
≥4	81.5	59.4	17.7	96.8	2.00	0.31	
≥5	74.1	65.0	18.5	95.9	2.12	0.40	
≥6*	70.4	72.0	21.3	95.8	2.51	0.41	
≥7#	70.4	76.8	24.6	96.0	3.04	0.39	
≥8⌘	51.8	83.3	25.0	94.1	3.11	0.58	
≥9	51.8	85.8	28.2	94.3	3.64	0.56	
≥10	44.4	89.4	31.0	93.7	4.21	0.62	
≥11	40.7	91.5	34.0	93.5	4.77	0.65	
Wave 1 4 item combination ***(n=273)***							***0.76 (0.66- 0.86)***
≥1	88.9	35.8	12.9	96.8	1.38	0.31	
≥2#*	77.8	64.6	19.1	96.4	2.2	0.34	
≥3	48.2	87.0	28.5	94.0	3.70	0.60	
≥4	25.9	96.3	42.9	92.4	7.08	0.77	

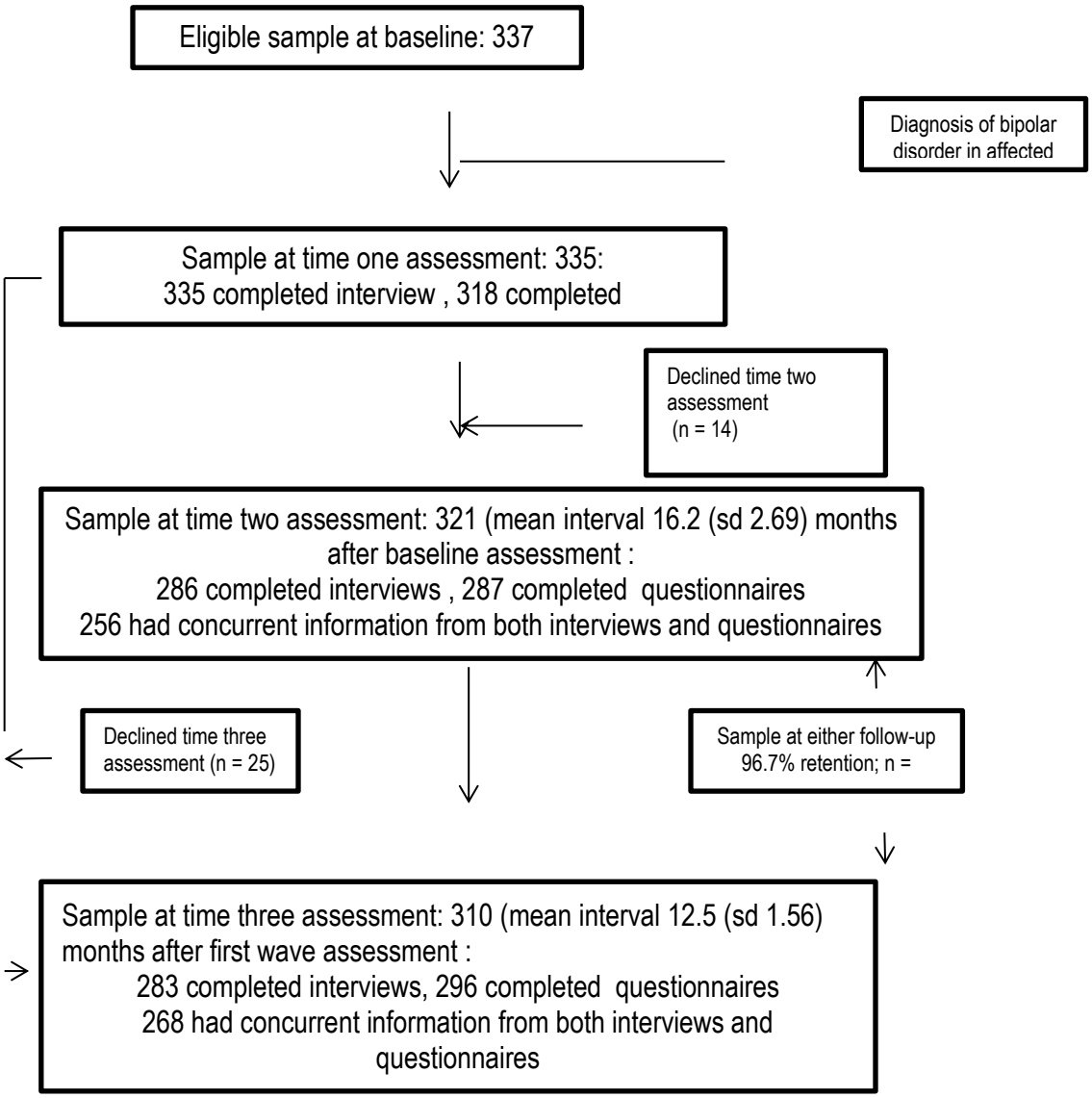
*optimal cut point for defining a DSM diagnosis of depressive disorder based on highest Youden index in original sample (sensitivity + specificity -1) and specificity ≥70%.⌘ cut-off recommended in literature
#highest Youden index for each wave

Supplementary Material 1
STROBE (von Elm et al.,2007) diagram: Recruitment In Development Dataset (based on Mars et al., 2012)



Supplementary Material 2

Flow Chart of Retention At Each Assessment in Development Dataset(based on Mars et al., 2012)



Supplementary Material 3
Results Of Univariate Analysis For Individual SDQ And MFQ Items With Concurrent MDD As Assessed By CAPA Interview As Outcome In The Development Dataset

SDQ items				
Item	*Description*	*Wave 1 p value*	*Wave 2 p value*	*Wave 3 p value*
3	**..headaches, stomach aches**	**.013**	**.030**	**.0010**
8	**Many Worries…..**	**.000**	**.002**	**.002**
13	**Often unhappy….**	**.000**	**.001**	**.000**
16	Nervous or clingy…	.007	.004	.06
24	**Many fears……**	**.009**	**.012**	**.000**
5	Often has temper tantrums…	.228	.132	.50
7	Generally obedient…..	.108	.0.26	.40
12	Often fights with other children ..	.317	.42	.32
18	Often lies or cheats	.80	.30	.80
22	Steals from home,school…/	.40	.07	.10
2	Restless, overactive…..	1.0	.03	.20
10	Constantly fidgeting….	1.0	0.1	0.1
15	Easily distracted…	.677	.10	.20
21	Thanks things out before acting	1.0	0.010	0.10
25	See tasks through to the end …	.30	0.02	0.07
6	**Rather solitary….**	**.03**	**.02**	**.002**
11	Has at least one good friend	.582	.01	.30
14	Generally liked by other children	1.0	.50	.10
19	Picked on..bullied by other children	.80	.08	**.005**
23	**Gets on better with adults .**	**.007**	**.017**	**.018**
1	Considerate..other people's feelings	1.0	1.0	.03
4	Shares readily…..	.30	.20	.60
9	Helpful if someone is hurt ….	.08	.30	.60
17	Kind to younger children	1.0	1.0	.20
20	Often volunteers to help others	0.015	.70	10
sMFQ items				
	Description	*Wave 1 p value*	*Wave 2 p value*	*Wave 3 p value*
	Miserable	**.000**	**.006**	**.000**
	Anhedonia	**.000**	**.000**	**.000**
	Tired	**.000**	**.009**	**.003**
	Restlessness	**.012**	.086	**.000**
	No good	**.000**	**.000**	**.000**
	Cried	**.001**	**.006**	**.000**
	Concentration	**.000**	**.000**	**.000**
	Self hate	**.000**	**.003**	**.000**
	Bad person	**.000**	.30	**.000**
	Lonely	**.000**	**.001**	**.000**
	Nobody loved them	**.006**	**.03**	**.000**
	Not as good	**.002**	**.000**	**.000**
	Wrong	**.001**	**.003**	**.000**

Those items with significant results (p≤0.05) at all three waves highlighted in bold. Used Chi-square test.

Supplementary material 4

Multivariate forward stepwise logistic regression using concurrent adolescent MDD as the outcome at each Wave: Incremental validity of each item-for each Wave in the Development dataset (using Nagelkerke R^2):

- For Wave 1 , R^2 increased from 0.180 (“Lack of enjoyment” alone) to 0.262 (adding “excessive worrying”) to 0.298 (adding “Problems with concentration”) to 0.329 (adding “Tiredness”)
- For Wave 2 , R^2 increased from 0.228 (“Problems with concentration” alone) to 0.285 (adding “Lack of enjoyment”) to 0.325 (adding “Not as good as others”).
- For Wave 3 , R^2 increased from 0.309 (”Feeling unloved” alone) to 0.360 (adding “Problems with concentration”) to 0.389 (adding “loneliness”).

+