PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	The DIADEMA Study- Prevalence of depression in patients with Type 2 Diabetes Mellitus in Spain: Results from the MADiabetes Cohort.
AUTHORS	SALINERO-FORT, MIGUEL; GÓMEZ-CAMPELO, PALOMA; San Andrés-Rebollo, Javier; Cárdenas-Valladolid, Juan; ABANADES-HERRANZ, JUAN; Carrillo de Santa Pau, Enrique; Chico-Moraleja, Rosa; Beamud-Victoria, Domingo; de Miguel-Yanes, Jose; Jimenez-Garcia, Rodrigo; Lopez-de-Andres, Ana; Ramallo-Fariña, Yolanda; DE BURGOS-LUNAR, CARMEN

VERSION 1 – REVIEW

REVIEWER	Ailish Hannigan Graduate Entry Medical School, University of Limerick Limerick, Ireland
REVIEW RETURNED	21-Dec-2017

GENERAL COMMENTS	This is a generally well presented paper on a commonly researched topic internationally. There are some issues to clarify in the Methods and Results before the conclusions could be supported. The patients have been recruited from multiple primary health care centres - have the authors considered whether there are any characteristics of the primary health care centres/physicians which could be important in understanding variation in prescribing practices for antidepressants/physician-diagnosed depression in patients? This dataset could be considered as clustered data with a multi-level analysis including both patient and health care centre/physician characteristics – it would be useful for the authors to discuss why this wasn't considered. It would also be useful to know if there was variation in the prevalence rate across the centres. Very little information is given on the measurement of physical activity – this is a self-report of patients on their level of activity? And how are patients classifying themselves into sedentary, moderate-intensity etc.? The relatively large sample will impact on finding statistically significant results even when the magnitude of these differences may be small. It would be useful to add effect sizes to Table 1 for the comparison between the two groups. The Statistical Analysis section suggests that variables will be included in the logistic regression if there were statistically significant

	in the bivariate analysis – this hasn't happened for, for example, family history of DM or is a higher cutoff for the p-value being used for variable selection? There is no measure of goodness of fit of the models given, which would be useful. These are multivariable rather than multivariate models. Have the authors considered using a comorbidity index as well as presence/absence of individual health conditions? Multimorbidity has been shown to be an important predictor of depression in those with diabetes. When trying to predict a new episode of depression in just over 1 in 10 patients (a relatively rare event in the context of fitting prediction models), there will be imprecision in the estimates and the models are more likely to correctly classify those without new depression than those with new depression. The imprecision can be seen in the width of some of the confidence intervals in Table 3. The authors should acknowledge the uncertainty in the estimates and also provide a measure of goodness of fit for the model. More care needs to be taken with some of the statements in the discussion e.g. 'moderate or vigorous physical activity significantly reduced by 59.1% the risk of depression' This estimate of the reduction in risk of depression (relative to those with sedentary behaviour) comes with considerable uncertainty - the confidence interval is from 0.24 to 0.69. It is also rare in either group (those with or without depression) to engage in moderate to vigorous activity (less than 10%) so any recommendations on physical activity need to reflect the reality that few people engage in moderate to vigorous activity in this cohort. Similarly the numbers of foreign born in the sample are very small (just 8 in those with depression). Without knowing the percentage in the population in Spain, this may be as a result of the exclusion criteria of not understanding Spanish. The very wide confidence intervals for foreign born reflect the small numbers and should be acknowledged. Also the statement that immigrant populations are at increased risk of depression comes from a paper which acknowledges that their sample is not representative of the general immigrant population. There looks to be no support for the statement 'limb amputation showed a tendency to be protective factor for incident depression' in either Table 2 or 3? Overall the paper has potential but more care and precision is required by the authors on their conclusions and analysis, looking in more detail at the data they have (and also where numbers are very small); acknowledging the uncertainty in their estimates rather than just focusing on the ORs themselves; and focusing on the results

	they can say with most certainty and with most relevance to practice.

REVIEWER	Dickens Akena Makerere University, Uganda
REVIEW RETURNED	02-Jan-2018

GENERAL COMMENTS	The authors conduct an important research in an area commonly neglected, and should be commended for that. My main concerns are in the methods that they authors used to come to their conclusions. 1. The author's definition of depression (on which this study hinges) is not very convincing. The authors would do well by describing in a bit more detail who assessed for depression. The authors report the use of the MINI 5.0, but are unclear who administered this instrument. Furthermore, the authors state that depression was diagnosed using clinical judgement; again, it is not clear who was responsible for this judgement. 2. The use of antidepressants as a proxy measure for depression in patients with DM is a major flaw. Antidepressants are used for all sorts of reasons including treating neuropathies, eating disorders, migraines, pain disorders etc etc. 3. It would help the reader if the authors stratified these results by stating how many of the 691 participants with depression were identified using the MINI, clinical judgement and use of antidepressants. It would make more scientific sense if the authors stuck with the MINI and clinical judgement, and not antidepressant use. 4. There are multiple variables including hypertension, heart failure, stroke etc that the authors report about. It is not clear how these parameters were assessed (heart failure for example), and by whom. 5. The authors also report that previous episodes of depression were associated with a current depressive disorder/episode. It would help the reader to know how these past episodes were assessed, and how the authors distinguished a past episode from chronic subclinical forms of depression or dysthymia. This finding further reiterates the need to stick to a clear diagnostic category. 6. The authors write about controlling for confounders and (known risk factors in the discussion section) while conducting multivariable analyses. It would help the reader if the authors stated which variables they controlled for, and why. 7. The authors state that 363 participants developed new episode of depression. It remains unclear what was used to assess these participants. 8. The authors report that a previous episode of depression was an indicator of an incidence of depression. If the participants already had suffered from a depression in the past, how can the new episode be called an incidence? The authors would do better by enlightening the readers about the new episode, and state whether it was distinctly different from the past episode; as it stands, it unclear whether these were simply recurrences/relapses, or chronic depressive illnesses. 9. Usually, risk ratios are reported for incident cases, and not odds ratios the way it has been stated. The authors need to double check with their statistician.

VERSION 1 – AUTHOR RESPONSE

Reviewer 1.

This is a generally well presented paper on a commonly researched topic internationally. There are some issues to clarify in the Methods and Results before the conclusions could be supported.

Q1. The patients have been recruited from multiple primary health care centres - have the authors considered whether there are any characteristics of the primary health care centres/physicians which could be important in understanding variation in prescribing practices for antidepressants/physician-diagnosed depression in patients? This dataset could be considered as clustered data with a multi-level analysis including both patient and health care centre/physician characteristics – it would be useful for the authors to discuss why this wasn't considered. It would also be useful to know if there was variation in the prevalence rate across the centres.

A1.

Thank you very much for this very pertinent comment.

Given the hierarchical structure of our data, we used a multilevel logistic regression analysis with two levels: level 1, patients, and level 2, health centers (our sampling unit). However, in the initial step (null model), the variation in the prevalence of depression between centers was not significant ($\sigma 2u0 = 0.02$, SE = 0.02, p = 0.115) with an intraclass correlation coefficient of 0.007; therefore, we did not consider it necessary to adjust for a hierarchical model.

Now, we include this explanation in the manuscript.

Q2. Very little information is given on the measurement of physical activity – this is a self-report of patients on their level of activity? And how are patients classifying themselves into sedentary, moderate-intensity etc.?

A2.

Thank you for highlighting this important point. Physical activity was measured using a short questionnaire based on the FAO/WHO/UNU Expert Consultation Report Energy and Protein Requirements (Geneva, 1985) and administered individually at a medical examination. The answers were coded from 1 to 3, with 1 representing inactivity or sedentary activity (remaining seated or at rest most of the time, sleeping, resting, sitting or standing, walking on flat ground, light housework, playing cards, sewing, cooking, studying, driving, typing, office duties, etc.), 2 representing low activity (walking at 5 km/h, heavy housework [cleaning windows, etc.], jobs such as carpenter, construction workers [except hard work], chemical industry, electrical, mechanized agricultural tasks, playing golf, child care, etc.), and 3, moderate or vigorous activity (non-mechanized agricultural tasks, mining, forestry, digging, chopping wood, hand mowing, climbing, mountaineering, playing football, tennis, jogging, dancing, skiing, etc.). 3. Drinking (0.1 through 4.9, or 5.0 or more g/d of alcohol).

Now, we include this explanation in the methods section.

Q3. The relatively large sample will impact on finding statistically significant results even when the magnitude of these differences may be small. It would be useful to add effect sizes to Table 1 for the comparison between the two groups.

A3.

We agree with the reviewer that adding the effect size allows to move away from the simple identification of statistical significance and toward a more generally interpretable, quantitative description of the magnitude of an effect.
Effect sizes describe the observed effects; effects that are large but non-significant may suggest further research with greater power, whereas effects that are trivially small but nevertheless significant because of large sample sizes can warn researchers against possibly overvaluing the observed effect.
When examining the difference between two conditions, effect sizes based on standardized differences between the means are commonly recommended. These include Cohen's *d*, Hedges's *g,* and Glass's *d.* We have preferred to use *Cohen's d* in this manuscript. Also, the measure of association that is deemed appropriate for cross-tabulation where one or both variables have more than two categories is *Cramer's V.* For this reason, we have used this statistic.

New text:
Quantitative variables were expressed as mean and standard deviation; qualitative variables were expressed as frequency distribution. Normally distributed continuous variables were compared using the t test, non-normally distributed variables were compared using the Mann-Whitney test, and categorical variables were compared using the chi-square test. ***Effect sizes were calculated using Cohen's d for continuous measures and Cramer's V for categorical variables****.*

Q4. The Statistical Analysis section suggests that variables will be included in the logistic regression if there were statistically significant in the bivariate analysis – this hasn't happened for, for example, family history of DM or is a higher cutoff for the p-value being used for variable selection?

A4.
Following the recommendations of the second reviewer, we have modified the definition of depression, in such a way that it now includes the diagnosis through the MINI 5.0 and the concurrent diagnosis by its usual doctor. This circumstance has changed baseline characteristics (Table 1). Now, we have not used a p-value cut-off to include variables in the logistic regression, because we prefer a more explicative model than the previous one.

New text:
The primary outcome variable was depression. The diagnosis of depression was considered a combined variable, as suggested by other authors (21), consisting of a diagnosis based on the module of major depressive disorder of the International Neuro-psychiatric Interview (MINI 5.0.0) (22), The interview was applied by a trained psychologist, and the diagnosis was made with the patient's general practitioner clinical who used his/her clinical judgment to determine whether the patient's symptoms and signs were compatible with a depressive disorder.
The MINI is a short and efficient diagnostic interview to diagnose mental disorders, which was used in its Spanish version (23).

Q5. There is no measure of goodness of fit of the models given, which would be useful.

A5.

Thank you for your suggestion. Now, we have included the value of Hosmer-Lemeshow goodness of fit test in results. A non-significant Hosmer-Lemeshow test result, while consistent with the null hypothesis that the model is correctly specified / fits the data well. Our Hosmer-

Lemeshow test result is chi-square=5.132; df=5; p=0.743.

Tabla de contingencias para la prueba de Hosmer y Lemeshow

		Without depression		With depression		
		Observed	Expected	Observed	Expected	Total
Paso 1	1	283	279,688	6	9,312	289
	2	275	274,281	14	14,719	289
	3	272	269,789	17	19,211	289
	4	267	264,497	22	24,503	289
	5	258	257,821	31	31,179	289
	6	245	249,619	44	39,381	289
	7	235	236,919	54	52,081	289
	8	205	213,093	84	75,907	289
	9	166	169,377	123	119,623	289
	10	108	98,915	177	186,085	285

The Hosmer-Lemeshow goodness of fit test is based on dividing the sample up according to their predicted probabilities, or risks. Specifically, based on the estimated parameter values β_0, β_1,..,β_p, for each observation in the sample the probability that Y=1 is calculated, based on each observation's covariate values:

$\pi = \exp(\beta_0 + \beta_1 X_1 + .. + \beta_p X_p) \,/\, 1+\exp(\beta_0 + \beta_1 X_1 + .. + \beta_p X_p)$

The observations in the sample are then split into g groups -we will later come back to the choice of g- according to their predicted probabilities. Suppose (as is commonly done) that g=10. Then the first group consists of the observations with the lowest 10% predicted probabilities. The second group consists of the 10% of the sample whose predicted probabilities are the smallest next ones, and so forth.

Suppose for the moment, artificially, that all of the observations in the first group have a predicted probability of 0.1. Then, if our model is correctly specified, we would expect the proportion of these observations who have Y=1 to be 10%. Of course, even if the model is correctly specified, the observed proportion will deviate to some extent from 10%, but not by too much. If the proportion of observations with Y=1 in the group were 90% instead, this would suggest that our model is not accurately predicting probability (risk), i.e. an indication that our model would not be fitting the data right.

Q6. These are multivariable rather than multivariate models.

A6.

The terms multivariate and multivariable are often used interchangeably in the public health literature. However, these terms actually represent 2 very distinct types of analyses. Statistically speaking, multivariate analysis refers to statistical models that have 2 or more dependent or outcome variables, and multivariable analysis refers to statistical models in which there are multiple independent variables.

Multivariate, by contrast, refers to the modeling of data that are often derived from longitudinal studies, wherein an outcome is measured for the same individual at multiple time points (repeated measures), or the modeling of nested/clustered data, wherein there are multiple individuals in each cluster (1).

For this reason, we agree with the reviewer and, now, we use “multivariable” model.

(1). Hidalgo B, Goodman M. Multivariate or multivariable regression? Am J Public Health. 2013;103(1):39-40

Q7. Have the authors considered using a comorbidity index as well as presence/absence of individual health conditions?. Multimorbidity has been shown to be an important predictor of depression in those with diabetes.

A7.

Unfortunately, our database does not include a comorbidity index as the Charlson’s index. However, we have used the combined variable cardiovascular event that includes non-fatal myocardial infarction, non-fatal stroke and peripheral arterial disease. We have also included heart failure. We think that is most relevant to know the magnitude of association with depression of each independent variable.

Q8. When trying to predict a new episode of depression in just over 1 in 10 patients (a relatively rare event in the context of fitting prediction models), there will be imprecision in the estimates and the models are more likely to correctly classify those without new depression than those with new depression. The imprecision can be seen in the width of some of the confidence intervals in Table 3. The authors should acknowledge the uncertainty in the estimates and also provide a measure of goodness of fit for the model.

A8.

We have detected an error in the calculation of incidents cases of depression after one year of follow-up and only 28 cases (1) were new diagnoses in patients without previous depressive disorder. Now, we have derived a new predictive model for the incidence of depression. Backwards LR binary logistic regression was performed to determine which factors were predictive of depression. This predictive model included in the saturated model variables as cardiovascular event, and heart failure. The final model only included four variables (gender, exercise, social support and diastolic blood pressure) and the confidence interval for female gender was wide (1.129 to 6.083). The Hosmer-Lemeshow goodness of fit test has been incorporated (p value=0.881) (2).

(1)

CROSSTABS

			DEPRESSION AT FINAL FOLLOW-UP		
			NO	YES	Total
DEPRESSION AT BASELINE	NO	Count	2335	28	2363
		% among DEPRESSION AT BASELINE	98.8%	1.2%	100.0%
	YES	Count	394	198	592
		% among DEPRESSION AT BASELINE	66.6%	33.4%	100.0%
Total		Count	2729	226	2955
		% among DEPRESSION AT BASELINE	92.4%	7.6%	100.0%

(2)

Hosmer y Lemeshow Test

Paso	Chi square	df	Sig.
1	3,062	8	,930
2	3,065	8	,930
3	2,238	8	,973
4	3,534	8	,897
5	3,524	8	,897

6	3,521	8	,898
7	4,364	8	,823
8	4,662	8	,793
9	3,699	8	,883
10	2,127	8	,977
11	4,791	8	,780
12	7,732	8	,460
13	7,638	8	,470
14	7,708	8	,462
15	11,678	8	,166
16	11,573	8	,171
17	7,136	8	,522
18	4,890	8	,769
19	13,482	8	,096
20	8,960	8	,346
21	7,621	8	,471
22	2,493	8	,962
23	3,776	8	,877
24	3,728	8	,881

Q9. More care needs to be taken with some of the statements in the discussion e.g. 'moderate or vigorous physical activity significantly reduced by 59.1% the risk of depression' This estimate of the reduction in risk of depression (relative to those with sedentary behaviour) comes with considerable uncertainty - the confidence interval is from 0.24 to 0.69. It is also rare in either group (those with or without depression) to engage in moderate to vigorous activity (less than 10%) so any recommendations on physical activity need to reflect the reality that few people engage in moderate to vigorous activity in this cohort.

A9. We agree with this commentary. Now, we have changed the sentence by this: "the association with depression could be reduced by physical activity, as we found in patients with low physical activity compared with a sedentary lifestyle (OR, 0.552; 95%CI, 0.408 to 0.746;

p=<0.001). However, we did not demonstrate a similar benefit in those who undertake moderate or vigorous physical activity. This phenomenon could be explained by the fact that very few had a high level of activity"

Q10. Similarly the numbers of foreign born in the sample are very small (just 8 in those with depression). Without knowing the percentage in the population in Spain, this may be as a result of the exclusion.

A10. The vast majority of immigration people in Spain are young, and very few of them suffer T2DM. Following the suggestion of the reviewer, now, we only adjust the results for country of origin. The immigrant population registered in Spain is nearly to 12.5%. At least, ten percentage points more than the sample studied.

Reviewer 2

The authors conduct an important research in an area commonly neglected, and should be commended for that. My main concerns are in the methods that they authors used to come to their conclusions.

Q1. The author's definition of depression (on which this study hinges) is not very convincing. The authors would do well by describing in a bit more detail who assessed for depression. The authors report the use of the MINI 5.0, but are unclear who administered this instrument. Furthermore, the authors state that depression was diagnosed using clinical judgement; again, it is not clear who was responsible for this judgement.

A1.

Thank you very much for this very pertinent comment. We have modified the definition of depression, in such a way that it now includes the diagnosis through the MINI 5.0 applied by a trainer psychologist and the concurrent diagnosis by the usual doctor who takes care of the patient and used his clinical judgment. Clinical judgment was applied to determine if symptoms and signs expressed by a patient were compatible with a depressive disorder.

This circumstance has changed the baseline characteristics (Table 1) and the results of the multivariable analysis.

Q2. The use of antidepressants as a proxy measure for depression in patients with DM is a major flaw. Antidepressants are used for all sorts of reasons including treating neuropathies, eating disorders, migraines, pain disorders etc etc.

A2.

Following your suggestion, we have not considered the use of antidepressant drugs in the definition of depression. The methods, results, tables and discussion have been modified accordingly

Q3. It would help the reader if the authors stratified these results by stating how many of the 691 participants with depression were identified using the MINI, clinical judgement and use of antidepressants. It would make more scientific sense if the authors stuck with the MINI and clinical judgement, and not antidepressant use.

A3.

Following your suggestion, we have included in results the proportion of patients who suffer depression diagnosed by MINI 5.0 and by clinical judgment.

Q4. There are multiple variables including hypertension, heart failure, stroke etc that the authors report about. It is not clear how these parameters were assessed (heart failure for example), and by whom.

A4.

Following your suggestion, we have included in the methods section a short definition of variables such as hypertension, heart failure and stroke.

New text:

Comorbidity variables: hypertension, defined as systolic blood pressure ≥140 mmHg, and/or diastolic blood pressure ≥90 mmHg; heart failure, which was defined as symptoms of dyspnea or edema associated with bilateral rales, elevated venous pressure, or interstitial or alveolar edema on chest X-ray, and required the addition of diuretics or inotropic medications; myocardial infarction, defined as a history of chest pain/discomfort associated with elevation of ST segment in electrocardiographic in two or more contiguous leads and elevation of myocardial enzymes; stroke, defined as a rapidly developing clinical syndrome of focal disturbance of cerebral function that lasted more than 24 hours; peripheral artery disease, defined as a symptomatic and documented obstruction of the distal arteries of the leg; low limb amputations, defined as the complete loss in the transverse anatomical plane of any part of the lower limb; erectile dysfunction, defined as the consistent inability to achieve or maintain an erection sufficient for satisfactory sexual performance; retinopathy, defined as a documented diagnosis by an ophthalmologist of non-proliferative retinopathy, proliferative retinopathy or macular edema; nephropathy, defined as a history of renal disease due to diabetes mellitus or requiring dialysis; neuropathy, defined as diminished or lack of perception of touch or pain stimuli and loss of joint position sense and vibration sense, and renal failure, defined as an estimated glomerular filtration rate below 30 mL/1.73 m^2. Cardiovascular disease (CVD) was defined as one or more of the following: myocardial infarction, stroke or peripheral vascular disease.

Q5. The authors also report that previous episodes of depression were associated with a current depressive disorder/episode. It would help the reader to know how these past episodes were assessed, and how the authors distinguished a past episode from chronic subclinical forms of depression or dysthymia. This finding further reiterates the need to stick to a clear diagnostic category.

A5.

Sorry if this was unclear. The previous episodes of depression were reported by the usual doctor of patient, after consulting the patient's clinical records. In Spain the most care for depression is delivered by general practitioners (GPs) and individually many GPs have considerable expertise in managing depression. Usually the GPs have easy access to consult patients who offer diagnostic doubts with the mental health team in the area.

Q6. The authors write about controlling for confounders and (known risk factors in the discussion section) while conducting multivariable analyses. It would help the reader if the authors stated which variables they controlled for, and why.

A6.

The objective was to obtain an explicative model and for this reason the final model was saturated. In other words, the model included all variables by method “enter”.

Q7. The authors state that 363 participants developed new episode of depression. It remains unclear what was used to assess these participants.

A7.

Unfortunately, we have detected an error in the calculation of incidents cases of depression after one year of follow-up and only 28 cases (1) were new diagnoses in patients without previous depressive disorder at baseline. Now, we have derived a new predictive model for the incidence of depression. Backwards LR binary logistic regression was performed to determine which factors were predictive of depression. This predictive model included in the saturated model variables as cardiovascular event, and heart failure. The final model only included four variables (gender, exercise, social support and diastolic blood pressure).

(1)

CROSSTABS

			DEPRESSION AT FINAL FOLLOW-UP		
			NO	YES	Total
DEPRESSION AT BASELINE	NO	Count	2335	28	2363
		% among DEPRESSION AT BASELINE	98.8%	1.2%	100.0%
	YES	Count	394	198	592
		% among DEPRESSION AT BASELINE	66.6%	33.4%	100.0%
Total		Count	2729	226	2955
		% among DEPRESSION AT BASELINE	92.4%	7.6%	100.0%

Q8. The authors report that a previous episode of depression was an indicator of an incidence of depression. If the participants already had suffered from a depression in

the past, how can the new episode be called an incidence? The authors would do better by enlightening the readers about the new episode, and state whether it was distinctly different from the past episode; as it stands, it unclear whether these were simply recurrences/relapses, or chronic depressive illnesses.

A8.

We deeply regret having had an error in calculating the incidence of new cases of depression, as we did not exclude patients who were already depressed a year earlier. Now, we have excluded patients with depression at baseline.

Q9. Usually, risk ratios are reported for incident cases, and not odds ratios the way it has been stated. The authors need to double check with their statistician.

A9.

Much clinical research is concerned with the extent to which one or more factors affect the occurrence of an outcome. The factor may be dichotomous, in which case there is only one increment, or continuous, with multiple increments. Results of a study may be expressed as the comparative risk for occurrence of the outcome with incremental change in the factor. The most common expressions of comparative risk in the medical literature are the risk ratio (RR) and the odds ratio (OR). The RR is a ratio of probabilities, which are themselves ratios. The numerator of a probability is the number of cases with the outcome, and the denominator is the total number of cases. The RR lends itself to direct intuitive interpretation. For example, if the RR equals X, then the outcome is X-fold more likely to occur in the group with the factor compared with group lacking the factor. The OR is a ratio of odds, which are also, themselves, ratios. Odds have a numerator the same as a probability, the number of cases with the outcome. However, the denominator differs; it is the number of cases without the outcome, not the total cases. There is no simple quantitative interpretation for the OR, except to the extent that it approximates the RR.

Despite the intuitive difficulty of the OR, it frequently appears as a measure of risk in multivariable analysis because of convenient mathematical properties of odds (ranging from 0 to +∞) compared with probabilities (limited to the interval between 0 and 1) (1). For the reader trying to understand the magnitude of an effect, the divergence between the OR and the RR can be important. It can be shown that this divergence is particularly large when the outcome is common in the study population. There are methods to estimate RR from OR reported in cross-sectional, cohort, and randomized studies (2,3). However, many readers are unfamiliar with these methods and may be led to an exaggerated impression of the risk.

So long as the risks of the disease are low, OR will approximate RR (4). Davies et al. (5) note that this approximate relationship breaks down when the risk in either group rises above 20%, with OR and RR becoming increasingly disparate.

Therefore, given the low incidence of depression after 1-year of follow-up found in our study, the OR is nearly equivalent to the RR as an estimator of the strength of association between the predictive variables and depression disorder.

1. D Hosmer, S Lemeshow. Applied logistic regression, Wiley, New York (1989).
2. C Zocchetti, D Consonni, P Bertazzi. Relationship between prevalence rate ratios and odds ratios in cross-sectional studies. Int J Epidemiol 1997; 26: 220-223.

3. J Zhang, K Yu. What's the relative risk? A method of correcting the odds ratio in cohort studies of common outcomes. JAMA 1998; 280: 1690-1691.
4. Bewick, V., Cheek, L., & Ball, J. (2004). Statistics review 11: Assessing risk. Critical Care 2004; 8:287–291.
5. Davies, H., Crombie, I., & Tavakoli, M. When can odds ratios mislead? British Medical Journal 1999; 316:989–991.

VERSION 2 – REVIEW

REVIEWER	Dickens Akena Makerere University
REVIEW RETURNED	05-Apr-2018

GENERAL COMMENTS	The authors have addressed all my comments