

PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (<http://bmjopen.bmj.com/site/about/resources/checklist.pdf>) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

TITLE (PROVISIONAL)	What information needs do people with recently diagnosed diabetes mellitus have and what are the associated factors? A cross-sectional study in Germany
AUTHORS	Grobosch, Sandra; Kuske, Silke; Linnenkamp, Ute; Ernstmann, Nicole; Stephan, Astrid; Genz, Jutta; Begun, Alexander; Haastert, Burkhard; Szendroedi, Julia; Müssig, Karsten; Burkart, Volker; Roden, Michael; Icks, Andrea

VERSION 1 – REVIEW

REVIEWER	Timothy Skinner Charles Darwin University, Australia
REVIEW RETURNED	18-Jul-2017

GENERAL COMMENTS	<p>The authors have attempted a critical question, for which there is very little published research to date. Although they do not seem to be aware of the most recent work in the area. There is a potentially a very good and clinically important paper here, but I have a number of significant concerns that prevent me from recommending it for publication.</p> <p>1. The authors report on a survey of people with type 1 and type 2 diabetes, and treat them the same. The aetiology, progression, treatment and impact of these two conditions is substantially different. The self-management demands and needs are completely different and these conditions generate very different profiles of psychological stress. As a result they need to be treated very different and reported separately.</p> <p>The authors have taken a large number of variables and collapsed them into simple dichotomous categories, such as "and quality of life was coded by 'low quality of life' and 'high quality of life'. First this means a lot of rich data is lost. Even if we accept the dichotomising of data, we are not provided with the cut offs for they have used to create this categories.</p> <p>3. I find the discussion of research as an information need, somewhat odd. A need is something that is required to achieve a goal. Information about research is not needed to manage diabetes, so I find its inclusion somewhat odd.</p> <p>4. The contradictory nature of the results regarding whether people feel they are well informed and whether they have need points to my own concerns about the nature of these questions and the validity of the approach, and this is not helped by the figures. For instance Figure 1 lists 99 individual saying they have need of information on treatment, whilst Figure 2 reports 103 say they are well informed. This makes no sense.</p> <p>5. The multivariate analysis is very vague and we do not have full reporting of these. The authors state "After discussion of the results of models using different groups of covariables, the following fixed</p>
-------------------------	---

	sets of independent variables including confounders were selected for the three main models:" this sounds like a fishing trip rather than a theoretically informed consistent analysis.
REVIEWER	Lisa Whitehead Edith Cowan University, Australia
REVIEW RETURNED	14-Aug-2017
GENERAL COMMENTS	Thank you for the opportunity to review this paper. Whilst the paper is technically correct in relation to methodology, I do not feel that the paper adds to our body of knowledge in this field. The paper is highly descriptive and the qualitative data fails to add the further insight one would hope to gain to expand on the meaning of the survey data.
REVIEWER	Carmine Gazzaruso Clinical Institute Beato Matteo - San Donato Hospital Group, Vigevano, Italy
REVIEW RETURNED	10-Feb-2018
GENERAL COMMENTS	The topic is of interest and the paper is well-written. I suggest to add a short comment on the potential clinical implications of the study in the discussion. Indeed, if confirmed, these results should be used to better define themes of a structured patient education program, as outlined by several recent papers, that may be used for the discussion.

VERSION 1 – AUTHOR RESPONSE

Reviewer #1

Comment 1: Although they do not seem to be aware of the most recent work in the area.

Answer 1: We performed a systematic review on the information needs of people with diabetes which was published after the submission of this manuscript (Biernatzki et al. 2018). We have added this information in the text (p. 6, l. 175; p. 19 l. 475–476). In that review as well as in the present study, we have tried to be specific and to strictly follow a published and common definition of information needs (Ormandy 2011) in order to work within a theoretically founded framework (Wilson 1997). We are aware that there are other thematic constructs which are closely related to information needs, e.g. needs in the context of self-management, education, and health literacy. To avoid misunderstandings, we have restructured and focussed the introduction.

Comment 2: The authors report on a survey of people with type 1 and type 2 diabetes, and treat them the same. The aetiology, progression, treatment and impact of these two conditions is substantially different. The self-management demands and needs are completely different and these conditions generate very different profiles of psychological stress. As a result, they need to be treated very different and reported separately.

Answer 2: We agree that both diabetes types are different and followed the reviewer's suggestion: (1) we have included the type of diabetes in our regression analyses; (2) we performed analyses stratified for diabetes type (see p. 12–13, l. 323–333). However, some analyses could not be calculated due to small sample size, and due to conditions, such as all type 1 patients being in need of information on diabetes research, meaning that only type 2 diabetes could be analysed for this model. We did indeed find differences, e.g. an association between information needs and mental component summary score of the SF-36 was significant only in people with type 1 diabetes (e.g. p. 16, l. 412–414, p. 20, l. 504–511).

Comment 3: The authors have taken a large number of variables and collapsed them into simple dichotomous categories, such as "and quality of life was coded by 'low quality of life' and 'high quality of life'". First this means a lot of rich data is lost. Even if we accept the dichotomising of data, we are not provided with the cut offs for they have used to create these categories.

Answer 3: We agree with the reviewer that information is lost when variables are dichotomised. However, it is a common procedure to do so for several constructs (German version II of the WHO-5 Index (Brähler et al. 2007), Problem Areas in Diabetes Scale (PAID) (Hermanns et al. 2006), ADS-L (Hautzinger et al. 2012)). We dichotomised these instruments accordingly to allow for comparison with other studies. We completely agree with the reviewer that the cut-offs should be defined and explained. We failed to do so for three instruments, thank you for this advice: WHO-5 Index (cut-off score >12), PAID (cut-off score ≥40), ADS-L (cut-off score >22). We have completed the definition of the cut-off values in the manuscript (p. 11, l. 284–290).

Comment 4: I find the discussion of research as an information need, somewhat odd. A need is something that is required to achieve a goal. Information about research is not needed to manage diabetes, so I find its inclusion somewhat odd.

Answer 4: This comment is very interesting. We believe that it is conceivable that the goal may not always refer to self-management itself. For example, the qualitative data showed that people need information regarding study participation (maybe due to the specific study population) as well as on study results and scientific progress. Potential aims could be to verify the knowledge of physicians or to get timely information on innovations. Other studies indicate these type of goals (St. Jean 2016). We have included a more detailed explanation in the discussion (p. 19, l. 473–476).

Comment 5: The contradictory nature of the results regarding whether people feel they are well informed and whether they have need points to my own concerns about the nature of these questions and the validity of the approach, and this is not helped by the figures. For instance, Figure 1 lists 99 individual saying they have need of information on treatment, whilst Figure 2 reports 103 say they are well informed. This makes no sense.

Answer 5: We assume that the subjectively perceived level of information does not match with the state of knowledge. Being well informed can also be expressed in the fact that people perceive that they have already received a lot of information, but have not yet achieved their actual goal. For example, the qualitative data shows that some people want more detailed information that is adapted to their level of knowledge. We have explained this in more detail in the manuscript (p. 19, l. 487–491).

Comment 6: The multivariate analysis is very vague and we do not have full reporting of these. The authors state "After discussion of the results of models using different groups of covariables, the following fixed sets of independent variables including confounders were selected for the three main models:" this sounds like a fishing trip rather than a theoretically informed consistent analysis.

Answer 6: Thank you for this comment. We agree that our selection of variables was not explained sufficiently and would like to describe our theoretically based approach in more detail below. We used two strategies to select the variables. Firstly, we considered the few variables for which an association with information needs has already been discussed in the literature (age, sex, education, type of diabetes, mode of treatment, diabetes-related comorbidity, see p. 9, l. 249–250). Secondly, we looked for further possible variables, considering five thematically relevant areas:

- 1) Socio-economic factors are very important in diabetes research and are associated with diabetes-related information-seeking behaviour, for example (Kuske et al. 2017). We therefore included further socio-economic factors: employment, school graduation, migration background.

- 2) Past diabetes experience is associated with information needs (Biernatzki et al. 2018). It can be assumed that diabetes-related and health-related factors may therefore have an impact on information needs (diabetes duration, HbA1c, number of overall drugs).
- 3) Regarding the need for information, some studies also reported on participation preferences (also in the sense of the 'informed patient' concept) and on the people's knowledge (Biernatzki et al. 2018).
- 4) People with diabetes have a higher prevalence of depression than people without diabetes mellitus (Hasan et al. 2015). Quality of life should play a major role in people with diabetes. Taking into account the different characteristics of the instruments, a disease-specific instrument (PAID) can focus on areas that are more relevant to those affected (Maurischat et al. 2008). The authors therefore recommend the instruments be combined.
- 5) Several studies have found that 'self-management' and 'lifestyle' are the main contents of the information needs of people with DM (Biernatzki et al. 2018), and thus the present study sought to identify a possible association (included: self-management, BMI, smoking behaviour, leisure time activity).

We have added this information to the manuscript (p.10–12, l. 257–302). Furthermore, we analysed these five groups separately in the regression analysis. The decision to exclude the variables from the groups was made because there was an insufficient number of cases (for instance clinically relevant depression with the ADS-L) or because of their low impact in the regression analysis. Methodologically, there was no difference in the use of various instruments to measure quality of life. We therefore decided to use the results of the SF-36 so as not to lose any rich data. Final models including independent variables of the groups were fitted so that each thematic group is represented by at least one variable. We have added the information in the manuscript (p. 12, l. 312–322).

Reviewer #2

Comment 1: Whilst the paper is technically correct in relation to methodology, I do not feel that the paper adds to our body of knowledge in this field. The paper is highly descriptive and the qualitative data fails to add the further insight one would hope to gain to expand on the meaning of the survey data.

Answer 1:

We would like to clarify the particularities of our results. According to the results of a recent systematic review (Biernatzki et al. 2018), there is a research gap on the information needs of people with diabetes mellitus and potential associated factors. The results can be used to develop information strategies for the target group. In our data, we saw that some people want more detailed information which is adapted to their knowledge level, for example. Therefore, level of knowledge should be taken into account at the beginning of the disease (we have added this information in the manuscript (e.g. p. 5, l. 143–145, p. 19, l. 490–491, p. 22, l. 548–549)). In addition, there does not appear to be a preference for information on clinical topics. The qualitative data clearly shows that the explanation of clinical topics frequently includes management-related information. In reference to the topic 'mental strain', for example, participants explain that they want more information on stress management. Resource-oriented provision of information at this point in time is therefore more likely to meet the needs of people with recently diagnosed diabetes (we have added this explanation in the manuscript (p. 19, l. 478–481)). Several questions remain unanswered, such as whether there are differences between patient subgroups, and which factors are associated with information needs.

Reviewer #3

Comment 1: I suggest to add a short comment on the potential clinical implications of the study in the discussion. Indeed, if confirmed, these results should be used to better define themes of a structured patient education program, as outlined by several recent papers, that may be used for the discussion.

Answer 1: Thank you for this advice. Regarding the clinical implications of the study, results may contribute to an adjustment of the design of communication strategies and education programs at an early stage of the disease. In some cases, people with diabetes felt that they received enough information about diabetes to not have to attend self-management education programmes (Horigan et al. 2017). An individual and patient-centred approach to designing programmes can increase participation (Horigan et al. 2017). We have added the clinical implications in the discussion section as suggested (p. 20, l. 514–518).

VERSION 2 – REVIEW

REVIEWER	Joseph Nolan Northern Kentucky University (USA)
REVIEW RETURNED	16-May-2018

GENERAL COMMENTS	<p>Please note that my field is statistics/biostatistics and I am reviewing only the statistical aspects of the manuscript. N/A in the review table indicates that I did not review that item.</p> <p>I have the following general concerns:</p> <ol style="list-style-type: none"> 1. It appears that you may have run 20+ models and are reporting only the significant results from these. If that is accurate, then the inflated false positive rate is probably not appropriate and should be controlled. 2. Running both the stratified and non-stratified models simultaneously does not seem appropriate. If there is a known difference between the two types of diabetes as indicated in the previous review, then the stratified models are appropriate and the combined model is not. 3. The manuscript should make better use of tables and include more comprehensive results for the logistic regression models. Further, when reporting on the basics, use percentages rather than counts. 4. You have reported on several odds ratios, but they do not appear to be well connected to your discussion. Consider enhancing the discussion to better incorporate. <p>The reviewer also provided a marked copy with additional comments. Please contact the publisher for full details.</p>
-------------------------	---

VERSION 2 – AUTHOR RESPONSE

Comments of Reviewer #4
Reviewer #4

General comments:

Comment 1: It appears that you may have run 20+ models and are reporting only the significant results from these. If that is accurate, then the inflated false positive rate is probably not appropriate and should be controlled.

Answer 1: Variable selection was performed on the basis of a content-based investigation of variable groups. The results of the selection process were interpreted in an explorative manner. The limitations of multiple testing and possible bias by variable selection are mentioned. Adjustment for multiple testing is not usual in such multifactorial association analyses of observational trials. The presentation

of the results focuses mainly on descriptive aspects. However, we have added tables of the final models including p-values (Appendix 2) and the following additional remark considering possible Bonferroni adjustment at the end of 'Associated factors' (p. 15, l. 392–394):

'If a Bonferroni adjustment for multiple testing for the number of independent variables were to be considered, only the association of need for diabetes research and diabetes-related comorbidity would remain significant in subjects with type 2 diabetes.'

Comment 2: Running both the stratified and non-stratified models simultaneously does not seem appropriate. If there is a known difference between the two types of diabetes as indicated in the previous review, then the stratified models are appropriate and the combined model is not.

Answer 2: We have deleted the non-stratified model (including diabetes type as an independent variable), because it did not contain relevant additional information (p.11, l. 298; p. 12, l.305–306; p. 14–15, l. 370–394; p. 17–19, l. 440–499).

Comment 3: The manuscript should make better use of tables and include more comprehensive results for the logistic regression models. Further, when reporting on the basics, use percentages rather than counts.

Answer 3: For a better presentation of the results, we have added a figure in the manuscript, and two tables and the questionnaire as appendices (Appendix 1, 2). Furthermore, we use percentages rather than counts (manuscript and figures). In particular, the results of the multiple models are presented in an additional table.

Comment 4: You have reported on several odds ratios, but they do not appear to be well connected to your discussion. Consider enhancing the discussion to better incorporate.

Answer 4: We have added tables in the appendix presenting the results of the multiple logistic regression models in detail, stratified only for diabetes type as suggested. Furthermore, as explained in the footnote of the table, the odds ratio of an independent variable corresponds to one unit change of this variable (Appendix 2; p.11, l. 289). In case of a binary variable, the odds ratio corresponds to 'yes' versus 'no'=reference. Age and two SF36 subscores are the only continuous variables with the units 'years' and 'points'.

Additional comments in the manuscript:

Comment 1: This is not qualitative. Please separate your comments on quantitative / qualitative (perhaps make it three sentences)

Answer 1: We have separated these sentences (p. 4, l. 119–120).

Comment 2: I find dichotomization of these variables to be a good strategy – effectively allowing for the use of logistic regression. An alternate strategy would be to use ordinal logistic regression, however it is not certain that doing that would be somehow more effective.

Answer 2: Thank you for pointing that out. We might consider this option for a larger study population.

Comment 3: Descriptive summaries were obtained (this action, while standard, does not deserve the term "analysis")

Answer 3: We have adopted the wording (p. 11, l. 283).

Comment 4: Was any control of false positives implemented? (e.g. Bonferroni corrected alpha level). If so, describe. If not, why not?

Answer 4: Please see our answer to comment 1 (general comments).

Comment 5: So you are effectively fitting at least 18 models? Some sort of control over the false-

positive rate should be implemented.

Answer 5: Please see our answer to comment 1 (general comments).

Comment 6: More information about this model selection is needed. Based upon what I'm reading here, it sounds like you are looking at MANY possible models – in which case even if there is nothing important you are likely to find something “statistically significant”.

Answer 6: As explained under ‘Quantitative analysis’, variables were selected based on a combination of contentual and statistical aspects. Groups of independent variables were included in multiple models and estimates. After fitting the model, confidence intervals and p values were discussed and a final multiple model was presented. Limitations of this process due to overestimated associations are referred in the text.

Comment 7: This is not logical. If you believe there is a difference between the two, you run stratified models. If not, you run an overall model. If there is a difference, the stratified models make sense and the overall model should not be considered (unless it includes diabetes type as an independent variable).

Answer 7: Please see answer 2 (general comments).

Comment 8: Please review SAS website for proper citation.

Answer 8: Thank you for this advice. We have adjusted the information according to the website (p. 12, l. 308–309).

Comment 9: The word “data” is plural. Through the manuscript fix this to “data were”.

Answer 9: We have adapted the wording throughout the manuscript.

Comment 10: What is meant by this, beyond what has already been explained above?

Answer 10: We have deleted this sentence (p. 12, l. 318–319).

Comment 11: It is a waste of space to detail these things in long text when you can simply refer to the summary table.

Answer 11: We suggest retaining this information because a short summary is a service for the reader.

Comment 12: I don't believe these fit the category of “participant characteristics”. Rather, these seem to be related to your outcome variables / results.

Answer 12: We have added another heading ‘Current level of information’ (p. 13, l. 329).

Comment 13: This definitely is not “patient characteristics” but rather appears to be a result. Also, rationale and intent for using McNemar test was not discussed in the methods section. Further, reporting the conditional percentages would be more appropriate than reporting the N's.

Answer 13: We have added another heading ‘Current level of information’ (p. 13, l. 329). Furthermore, we have added McNemar's test in the methods “Quantitative analysis” (p. 11, l. 284–286). Finally, N's were replaced by percentages (p. 13, l. 330–341).

Comment 14: In general, provide percentages rather than counts. Counts mean nothing without the memory of the sample size and converting to percentage. Please reflect this change across the paper. A table that was formatted with a column that contains both count and percentage would be most effective: Variable Count (%)

XXXXX 103 (75%)

YYYYY 80 (58%)

Answer 14: Please see our answer to comment 3 (general comments). Nevertheless, we suggest using only a limited number of tables as some descriptive results are summarized more briefly in the

text.

Comment 15: Unless I'm missing something, this is less than 40% of your sample meaning that, combined, the other two categories received greater priority. Why not simply present a table that summarizes these things (and those above) via percentages.

Answer 15: We have now reported on percentages instead of N's, have added a figure (see Figure 3) and have made adjustments to the manuscript to clarify the question (p. 8, l. 206–212; p. 13-14, l. 344–367). The description refers to the individual topics of the questionnaire and not to the categories (Appendix 1).

Comment 16: Be specific. (multivariable logistic regression?)

Answer 16: We have adopted the wording (p. 15, l. 380).

Comment 17: For any model on which you are reporting you should include a table that contains this sort of information for ALL variables that were considered in the model. As you say, low power might be a consideration. The only reasonable way to assess that is by reviewing the confidence intervals for odds ratios.

Answer 17: We have created a table for each model stratified by diabetes type in order to better present our results (see Appendix 2).

Comment 18: It is unclear what these numbers mean. Use tables in such a way as to make it clear. See comment above.

Answer 18: Please see comment 17 (additional comments in the manuscript).

Comment 19: May be. Not "is". Your CI for the odds ratio nearly includes the value 1. Therefore it seems much too strong to state clearly that we know of an important association here.

Answer 19: All p-values and significances should be interpreted with caution and in a descriptive manner. On the other hand, the size of the p-value itself is not a measure of the association. But as usual, 0.05 is used as a border for significance. We changed the wording to 'is significantly associated' (p. 15, l. 384).

Comment 20: Again this formatting of numbers is not clear what they mean. As a statistician I assume they are OR and CI. But you need to put this information into tables, and include it for the non-significant variables as well.

Answer 20: We think you're right. Please see comment 17 (additional comments in the manuscript).

Comment 21: How much lower? Is it relevant? You don't seem to have related any of this to your calculated OR's. Further, OR's may not be effective in answering the question. For example an odds ratio of 2 would mean that the likelihood for group 1 was twice that for group 2. But if the likelihood for group 2 is very small, is any of it even relevant? If your final model is univariate, you might be better to simply estimate the percentage for each group, so that a more clear comparison can be made. For multivariate models you could do the same thing by fixing covariates at their averages.

Answer 21: We have added an explanation that the OR's correspond to unit changes of the independent variables (only age and both SF36 subscores) (see Appendix 2, p. 11, l. 289). We present only multiple models, now including details in tables. It is difficult to interpret the magnitude of the OR's and to compare OR's between different independent variables. No results from other trials were available.

Comment 22: Meaning what?? Explain better

Answer 22: We have changed the wording and hope it is clearer now (p. 20, l. 515).

Comment 23: I believe this is a very big limitation. If the paper is to be accepted without adjustment

for multiplicity, you need to be clear on how many models / tests were actually performed, keeping in mind that if we do 100 tests where there is no effect, we would expect to “see” around 5 statistically significant results.

Answer 23: We agree that this is a limitation. This has been added in the discussion section (p. 20, l. 517–520). Furthermore, we now describe the change of results when Bonferroni adjustment is performed (please see also answer 1, general comments).