Author's response to reviews

Title: Do gastrointestinal complaints increase the risk for subsequent medically certified long-term sickness absence? The HUSK study.

Authors:

Simon Øverland (Simon.Overland@iu.h.uib.no)
Marit Knapstad (Marit.Knapstad@hemil.uib.no)
Ingvard Wilhelmsen (Ingvard.Wilhelmsen@med.uib.no)
Arnstein Mykletun (Arnstein.Mykletun@uib.no)
Nick Glozier (Nick.Glozier@sydney.edu.au)

Version: 2 Date: 28 April 2011

Author's response to reviews: see over
Reviewer 1 presented a range of relevant and constructive comments enabling us to make needed clarifications and increase the overall precision of the manuscript. In the following, we present our responses to the comments in the order presented in the reviewer’s report.

1. The first comment addresses a general need for increased precision and clarity throughout the manuscript. We have revised the manuscript carefully to improve the language. We have attempted to refine statements, definitions and information that we agree may have been lacking in precision in the previous version. As this comment related to the manuscript in general, it would be excessive to point to all changes made to meet this, but we have had this in mind through the entire revision and large parts of the manuscript are changed in order to improve clarity. As for the initial specific example offered by the reviewer, we have added data from one general population based study in Norway suggesting that the prevalence is high there as well (Haug et.al, 2002).

2. The reviewer comments upon a need for increased precision regarding the statement “…growing concern over increasing proportions of the work force absent from work…”, and a need for further contextualization for Norway and relevance for generalizability of findings. We have added a statement to bring in the Norwegian context and have added comments on the generalizability of the findings in the discussion.

3. In comment #3, the reviewer asks us to provide a better qualification for our choice of design over and above previous cross-sectional studies with self-report of sick leave as an outcome. We agree this is integral to presenting our design and have made considerable changes to the manuscript as a result. First, we have added a further qualification of the choice of design and explicitly state our opinion on the added value of prospective objective data e.g. reducing the likelihood of reverse causality being an explanation. Furthermore the use of objective outcome data reduces random misclassification, potentially reducing strength of associations. More importantly objective outcome data reduces non-random information bias common with self-report of both exposure and outcome. The prospective design is important for two main reasons: firstly, both the exposures and perhaps particularly some of the co-occurring conditions like anxiety and depression may be exacerbated by an ongoing sickness absence period, due to distress, behavioural responses in line with “sick-roles”, and loss of role-functioning. If gastrointestinal complaints in part are a response to a sickness absence (through any of the candidate mechanisms), this may inflate any true associations. Even more relevant for anxiety and depression, adjustment for these factors in a cross-sectional design, may leave an over-adjusted model. A prospective design, based on incident sickness absences after health information is provided, reduces these potential sources of error. Secondly, a prospective association is by definition studying something else than a cross-sectional association. A prospective analysis of influenza as an exposure for sickness absence months later would surely be of limited interest, as influenza generally is a transient condition. However, many who have gastrointestinal complaints and conditions report these over prolonged time periods making a prospective design more useful. Studying gastrointestinal complaints as a prospective risk factor for the same outcome, and with a positive association, points to an association lasting beyond the transient period of the health survey. A sickness absence when ill is the very point of having a sickness absence benefit available. However, sickness absences occurring in the period after the exposure, provides some scope to suggest reducing the exposure may improve the chances of preventing the outcome. With long term sickness absences as the outcome, the outcome is of great importance to individuals and society.
4. Why is extended sick leave more important to examine than possible frequent short-term sick days? Was this selection due to convenience (availability of data) or a theoretical rationale?

Reply: Both are important and the selection of extended absence is both our research focus and determined by the nature of the data linkage we made. Movement into long term dislocation from the workplace in the form of disability pensions, which in many OECD countries often last until retirement age, is a transition from extended sick leave, not frequent short term absence. The latter is generally a cost to an organisation, the former a cost to public sector finances. We have added this argument in the introduction.

In the Norwegian benefit system, the employer pays for the first sixteen days, while the government takes over from day 17 (was changed from first 14 days in 1998). Thus, the official statistics on benefits we are able to link the health data to only provide information on sickness absences exceeding 17 days. We have tried to make this clearer in the methods section where we present the outcome measurement.

5. The reviewer commented on a need for more precise presentation of the principal study questions need to be clarified using more precise language. We agree and have used the suggestions from the reviewer to inspire the reformulation of these. We did not feel comfortable framing these as hypothesis as they would then be formulated after having seen and analysed the data. We therefore present them as research questions in line with the previous presentation, only hopefully a lot clearer.

6. The reviewer comments on demographic variables included in the model and arguments for including them. It is also argued that a more careful selection of which variables are included would be useful as the statistical power in the study makes minor and irrelevant associations strongly significant. We agree with the comment and have provided arguments for including each of the potential covariates in the introduction. We have added more candidate variables, but applied criteria for including them in the multivariate models: a significant univariate association to both the exposure or the outcome variable, and an effect size (Cohen’s $\geq 0.10$) with at least one of them. We have added a statement on this in the methods, and this has led us to much more condensed multivariate models.

7. What was the rationale for tallying total number X frequency of complaints in determining a total score for GI complaints and, further, for dichotomizing this continuous scale using a cut off at the 80th percentile? Is this the 80th percentile for this group? Is this the 80th percentile based on a norm group? Is this methodology validated or was it created for this study by the authors for a reason that can be specified? If any reliability and/or validity information can be provided for this measure of GI complaints, that would be important to interpretation of results.

Reply: The main rationale was to use a variable indicating if participants had “low vs high” levels of gastrointestinal complaints. The main reason for doing so was our lack of a validated measurement as exposure variable. We would certainly have preferred if we had variables that would allow us to approach the “Rome-criteria” or other agreed upon standards.

Unfortunately, this was not available and is listed as a main limitation of the current study. In this situation, our view is that approaching the data analysis using total symptom load, and then dichotomising reflecting “low vs high” serves as a crude, but easily understandable indicator. Further, the variable summing up the total gastro-complaints was far from a normal distribution, which further points towards applying a cut off. We have added comments on this to the methods section. The strength of the study is that we have carried it out on a large
sample from the general population and the strong, prospective and objective outcome data. In
the revised discussion, we have tried to address this better and explicitly communicate what
our study adds and how this could inform future, more specific studies.

8. The reviewer asks whether all participants were followed up for an equivalent amount of
time. The health survey was completed over a period of more than one year. We had the exact
date of the health survey participation for all participants. The sickness absence registry also
has exact information on the date the sickness absence starts. We were therefore able to
ensure that each participant was at risk for exactly four years to the day. However,
reconsidering our approach, with exact information on time between exposure and event we
have re-analysed using Cox proportional hazard models to allow us to use the maximal follow
up time for all participants, overriding issues of different lengths of follow-up.

9. What was the rationale for serially dichotomizing extended sick leave into present/absent,
0/1(+) period of > 55 consecutive days, and 0/1(+) period of >100 consecutive days? Why did
the authors decide on this approach versus counting total number of sick days as a continuous
variable, or some other approach?

Reply: There are differing definitions of what constitutes extended or long term sick leave
which ranges from 4 weeks to 3 months depending upon organisation or jurisdiction. We
chose 8 weeks as the cutpoint as in Norway the patients GP must carry out a new and more
thorough examination of the patients health, as well as options for reintegration to work at 8
weeks. As part of the latter, the patient, the GP, the employer and representatives from the
social security organization are to have meetings to examine the case. As a consequence, there
are incentives to keep sickness absences shorter than 56 days, if possible. The next cut-off of
90 days was chosen as it corresponds to previously suggested cut offs, but also as it falls close
to the next “decision point” of 85 days in the Norwegian system, which calls for an extended
version of the “56 day procedure”. In Norway sickness absence over one year becomes
reclassified into other forms of benefit or different systems engage and so there are fewer
people in this longer group than may be observed in other countries that allow longer periods
of extended sick leave. We have added comments on this in the methods.

Regarding using binary indicators in the first place: As we have access to number of sick
days, an immediate thought would be to model this linearly with number of days sickness
absent as the outcome. There were however major challenges using that approach: First,
sickness absence days are not normally distributed. Furthermore, as above the maximum
duration of sickness absence (in the Norwegian social security system) is one calendar year.
After this, one must return to work, enter other disability arrangements if qualified or any
other option than receiving sickness absence benefits. This means that on the right tail of the
skewed distribution, there is another aggregation around 365 days, and this further
complicates using a linear model and any attempts to log-transform the distribution. Another
issue is that half the sample have 0 days of sickness absence. The latter renders Poisson
regression, the default approach for “count” data, inappropriate. Models using zero-inflated or
zero-inflated negative binominal regressions were considered, but in our opinion, the cost of
using such models is high in terms of a clear and lucid presentation of the results balanced
against a relatively small gain. To keep the results accessible for a larger proportion of
readers, and since more complicated models had their own challenges, we chose to employ
well-known models based on a categorical approach.
10. Were any diagnoses (physical or mental health) used to exclude an extended sick leave directly related to that diagnosis given the focus of this study on relatively “unexplained” somatic or GI symptoms?

Reply: We have not attempted do any such selections/exclusions of outcomes. It could be interesting, but we would not a priori know where to draw the line between excluded and included diagnoses. We also fear that, albeit tempting, this would also be a contagious strategy. Sickness absence is a highly complex issue with both medical and societal factors being highly relevant. Excluding sickness absences exclusively due to medical labels would make sense for many cases, but not for others as there may be strong extra-medical factors that interplay with the risk for the outcome. We have therefore decided not to attempt this approach in this study, but may certainly be something worth looking into on its own merits.

11. We agree that our previous presentation of the statistics was unclear and have changed the first sentence of the statistical analysis paragraph, to state that we present means/SD’s/frequencies and specific tests to examine the univariate associations to the exposure and outcome variables.

12. We have presented our reasons for using odds ratios/hazard ratios, dichotomized exposures and outcomes and use of logistic regression versus hierarchical multiple regression. Further, we have presented why we do separate models to make adjustments. In short; the main arguments for using dichotomies is that we did not have access to an established measurement of gastrointestinal complaints, and that the total symptom count variable is non-normally distributed. A “high/low” binary indicator is intuitive and “common-sense”, and hence accessible. We defined this based on a percentile defined cut-off of total symptom load. Our exposure variable thus represents a dichotomous “high/low” indicator of problems. When having a dichotomous variable as exposure, it would be wrong to use continuous symptom indicators as confounders or co-occurring conditions, as this would increase risk of type-II error (there being more “information” in the confounders than the exposure).

13. The issue of statistical versus statistical significance seems very relevant in the current study. With such a large sample, the actual group differences appear quite small and, yet, are statistically significant indicating a possible problem with overpowering the study. It may be helpful to include either effect size measures to provide greater interpretive information about the actual magnitude of the effects found and/or to discuss clinical meaningfulness in terms of the actual means for each group. For example, the baseline differences between groups (cut for 80th percentile of GI complaints) on BMI yields an extremely small effect size despite the significant p value. Taking effect size into consideration in selecting variables for entry into the model may help to streamline the analyses, as well as aid in more accurate interpretation of results. Alternately, the authors may wish to split the group randomly, using one half of the participants to develop the model and the remaining half to confirm the model to decrease overfitting to the data and to better improve generalizability.

Reply: This point is well taken, and we agree with the comment on statistical vs clinical significance. The comment, and our reply, relates to the reviewers comment #6. We have chosen to select covariates based on the magnitude of univariate associations, with use of a conventional cut-off for effect sizes based on chi-square’s (Cohen’s w).

14. Given the large amount of participants and data, the authors also may wish to consider using a more sophisticated statistical approach, such as structural equation modelling that best
capitalizes on the study’s power while preventing overfitting of the data. This would require the authors to make their hypotheses about relationships among variables more clear, but would allow for adding or pruning paths in an empiric fashion, as well. This approach also would allow for more information about possible mediation or moderation associated with specific variables (e.g., genders impact on the relationship between GI complaints and LTSA).

Reply: The use of an alternative strategy for data analysis is certainly interesting and also one that we are undertaking: We are currently carrying out a Latent Class Analysis (LCA in Mplus) of 17 somatic symptoms. As this is an ongoing study, the final results are not clear, but so far there seems to be support for a number of latent classes (5), where one class is multisymptomatic across all symptom domains, one class is best labelled as “asymptomatic” and the three remaining groups are characterized by being more likely to present symptoms on specific domains – whereof one of the domains are gastrointestinal symptoms. We then relate these groups to cross-sectional anxiety and depression and prospective awards of disability pension (permanent work incapacity benefit). In other words, we share the reviewer’s point that alternative data-analytic strategies are interesting and relevant. However, the current study addresses another issue in much more specificity, and we think the specific associations between gastrointestinal complaints and later sickness absence should be studied in detail by its own right. Clinicians, especially in primary care, are often presented with a small subset of the patients more pressing symptoms and rarely fully explore all other systems. Results from similar LCA approaches to common (psychological and physical) symptoms in primary care have proven difficult to translate into clinical practice. The results of the present study also suggest that there is further scope to understand how clinicians identify and deal with gastrointestinal complaints and how these relate and might contribute to the societal challenge of sickness absence.

15. The point made about there being a general tendency of symptom reporting driving the association between GI complaints and extended sick leave is an interesting one that may be testable using a structural equation model or hierarchical multiple regression and looking at several variables simultaneously for contribution to LTSA as an outcome. As alluded to above, this manuscript may be better served by a focus on larger scale model-building and testing, rather than on examining individual relationships in a somewhat separate manner.

Reply: As stated above, we agree but do not see this as an “either or”, but rather that both (and more) studies are needed. A thorough systematic review from 2004 clearly identified a major lack of studies on sickness absence causes, with prospective, objective data explicitly mentioned as called for. On the issue of somatic symptoms vs anxiety/depression, and how much of the explanation they each explain, we have calculated the “percentage of adjustment” as a means to provide some quantified information on the relative impact.

16. The Discussion should include some interpretation of study results from the point of view of social and economic cost, given the initial rationale provided about why the study questions are important. In addition, it would be helpful to include some discussion of potential clinical targets for intervention to ameliorate the identified problem, if available, or future directions to clarify what might be appropriate targets for future intervention.

Reply: We have tried to be more accurate on these issues and now present some thoughts on how our results could inform interventions and not at least that these results suggest future,
more clinically accurate studies could benefit from adding sickness absences as an outcome of interest.

17. The authors should verify that AMA style is used throughout the manuscript, including for inline citations, etc.

Reply: We have used the template for BMC Gastroenterology in Endnote, and hope the style is now conforming to the standards.

18. Comments 18 through 20 related to the abstract. Due to the changes in the manuscript in general, the abstract is now revised as well.

19. See pt 18

20. See pt 18

21. Greater specificity is warranted in language related to organicity. Recent literature increasingly has identified biological (or organic) contributors to the types of GI complaints asked about in the current study (historically considered “not to have a medical explanation”, using the words of the current authors). While an “organic disease” or “a clear organic cause” may not be found to fully explain the GI complaints endorsed, the authors should be cautious in their language to reflect current thinking in this area.

Reply: We agree that these statements should be made with caution and have revised this to be more circumspect on several occasions.

22. The participation rate given of 63% is higher than the actual percentage included in the final analysis. The actual final participation rate should be given after accounting for all of those who did not participate, including those who did not complete the questionnaires or clinical exams, deletions due to preceding disability pensions, lack of paid employment, and casewise deletion due to missing variables. Rationale for casewise deletion should be presented here, rather than in the discussion. In addition, these decisions (and the relatively large proportion of the population who were not included for some reason), along with the selection of a Norwegian population and the limited age range involved, should be discussed in terms of generalizability to other groups.

Reply: We have added the final response rate after accounting for all with partial missing responses.

23. Was this a secondary analysis of existing data or an actual prospective longitudinal study? Clarification on this point would be helpful, as well as citation to the original HUSK study manuscript which would provide further details on participant selection and data collection methodology, if available.

Reply: The proper term describing this design is a “historical cohort study”. It is not a true longitudinal study as the participants were approached only once for a health survey with primary aims not including sickness and disability, while we gather outcome information using external registries that cover events occurring in the years after the health study. We have added references to some previous studies using this exact data linkage in the revised
section where we present the data resources (Methods section; population and data material and outcome variables). We hope this now is a lot clearer.

24. The concept of absences less than 14 days in length providing a “wash out” period, or how the “wash out” period would be helpful, was not clear to this reader.

Reply: We can fully understand that this was not clear to the reader. Reviewer two also commented on this, and we fully agree that our previous statement was unclear (at best…). We have rephrased this accordingly.

25. Organization of the Method could be enhanced by labeling variables as predictor variables and target or outcome variables and reordering them to appear together as a block (e.g., GI complaints, somatic symptoms, physical conditions, psychological conditions, and demographic variables before presenting extended sick leave, the target).

Reply: We have reorganised the methods to become more in line with how the variables are used in the analyses.

26. Table 1 should be broken down into two tables to better present baseline differences related to LTSA. As currently organized, this half of the table is difficult to follow and does not provide mean/SD information for those not experiencing an LTSA during the study period.

Reply: We agree, and hope the editor does not object to us adding another table to accommodate this.

27. Please explain more fully the sentence in the second to last paragraph of the results which states “Those with more GI complaints more often had an ICPC-diagnosis from chapter “D-Digestive” warranting the sickness absence”.

Reply: After the initial submission, we have learned that the quality of the coding of sickness absence diagnoses was not optimal and as this was a minor, supplemental point in the previous version, we have chosen to omit this altogether.

28. All tables should be referenced in appropriate areas of the text and all statistical information mentioned, even in the Discussion, should be present in a table or within the text itself (e.g., information related to “joint pain and stiffness”).

Reply: We certainly agree that statistical information mentioned should be backed by data presented in tables or text and have tried to make sure that now is the case throughout the manuscript.

29. Table 4 is difficult to read, as it is the reverse of an earlier table, with presentation of the models across the top, rather than the side. Some degree of consistency in presentation would be helpful to the reader.

Reply: Changed according to the comment. We apologize for the inconsistency.

30. The authors state that there “was no attrition during follow up”, yet this is a bit misleading in that numerous potential participants were dropped from analyses due to incomplete data, or
other issues, limiting the pool of participants on the front end of the study. This should be clarified in the strengths and limitations more clearly. In addition, it is not clear whether any participants could have moved out of the healthcare system being studied and, thus, could have looked like no LTSA when, in fact, they were no longer in the pool.

Reply: By attrition, we mean loss of participants during follow-up. Participants excluded due to missing health information are in our opinion better described as non-participators or incomplete responders. The outcome data are taken from the official registries on benefits, based on personal identification numbers for all Norwegian citizens or those entitled to benefits. The base population was those who lived in the catchment area in 1997-99, within the defined age cohorts (born 1953-1957) and were registered in the population registry with their personal identification number. The only exit from this system is either permanent removal of citizenship or death. Moving to other parts of the country or even abroad does not incur loss to follow up in the registry. We therefore think the statement “no attrition” is correct, but have described the data sources better in the methods to avoid further misunderstandings on this point.

31. The Discussion may benefit from reorganization, with the initial focus on addressing the study’s aims and hypotheses (clarified as discussed in the Background), moving on to interpretation and clinical implications, followed by strengths and weaknesses/future directions. As it currently reads, it is difficult to follow the authors’ logical flow and identify the main findings and implications of this study. Further specifics cannot be given at this time, as the content of the Discussion may change considerably if the Background becomes more focused and the statistic plan is altered.

Reply: We have made some changes to the discussion, in line with the reviewer’s suggestions. We have tried to make it much clearer what this study adds, and issues of generalization.

32. The authors may wish to consider changing the title to better reflect main study findings.

Reply: We have reconsidered the title, but chose to keep it as was which we think states that the paper describes the link between gastrointestinal complaints and sickness absence.

33. Please consider eliminating the use of acronyms in the Abstract, unless they are defined on first appearance in the Abstract.

Reply: This was unintentional and we thank the reviewer for pointing this out, and we have changed this accordingly.
Reviewer 2, has presented specific comments that we agree needed attention. We thus agree with all the comments and in the following, we present how we have dealt with these and corresponding changes in the manuscript.

1) Reviewer 2 comments on our unclear statement regarding the 14 days “wash out” period. The idea behind a “wash out” period is to enhance the prospective design, where the outcome (sickness absence) does not contaminate the exposure (health variables). As we have added in the paper (introduction) an ongoing sickness absence may have an impact on self-reported health status. Numerous processes can contribute to this, described within the theoretical framework of “sick roles”, but also through empirically supported health consequences from being without a job. Thus, being on sickness absence may in itself increase or even cause health problems. Using a wash out period, a period immediately after the health measurement where we do not observe outcomes, reduces the risk of such reversed causality. We agree that this was very unclear in the previous version of the manuscript. We have therefore changed this to state that we only had information about sickness absences lasting more than sixteen days, and therefore started registering incident sickness absences 17 days after the health survey to make sure the sickness absence was started after the health measurement was completed.

2) The term "somatisation disorder" which has a specific meaning that may not be familiar to most gastroenterologists occurs on page 14 for the first time, although some of the features of it are discussed also in the introduction.

Reply: We agree the term “somatisation disorder” warrants a further explanation when presented in the context of gastrointestinal medicine. We have tried to be more descriptive instead of using this form of labels to convey our point.

3) The sentence (on page 15) referring to animal work and inflammation is a bit "off label" - there are many more human studies that link stress to UC or CD.

Reply: We absolutely agree and have replaced the animal studies with those conducted on human subjects.

4) In the Conclusion section, the sentence starting "The results ..." has an incorrect wording: it appears to me the authors would like to say that "... to explain work related outcomes of of functional GI complaints" rather than "...work related functional outcomes ..."

Reply: We agree and have revised accordingly.