Author’s response to reviews

Title: Predictors of physicians' stress related to information systems: A nine-year follow-up survey study

Authors:

Tarja Heponiemi (tarja.heponiemi@thl.fi)
Hannele Hyppönen (hannele.hypponen@thl.fi)
Sari Kujala (sari.kujala@aalto.fi)
Anna-Mari Aalto (anna-mari.aalto@thl.fi)
Tuulikki Vehko (tuulikki.vehko@thl.fi)
Jukka Vänskä (jukka.vanska@laakariliitto.fi)
Marko Elovainio (marko.elovainio@thl.fi)

Version: 1 Date: 23 Jan 2018

Author’s response to reviews:

22.1.2018

Dear Prof. Pfaff,

Enclosed, please find our revised manuscript "Predictors of physicians' stress related to information systems: A nine-year follow-up survey study." We highly appreciate reviewers´ valuable comments. We have now revised the manuscript following these comments. We are willing to do additional changes if necessary. All the changes are made with the track changes mode.

A detailed description of the revisions is included on the following pages. We hope that our revised manuscript is now acceptable for publication in BMC Health Services Research.

With Best Wishes,

Tarja Heponiemi
COMMENT: 1. On p.8, line 16ff. authors indicate a sample loss that is quite substantial. In the Discussion section a potential selection bias may be discussed (e.g. underestimation of effects as highly stressed physicians had a higher probability of leaving the hospital).

RESPONSE: As suggested, we have added discussion about this issue to our limitations section:

(from page 16, discussion) “In addition, our sample is not totally representative of the present physician population in Finland. Our sample included a higher percentage of women than the mean percentage for the eligible population. Moreover, because our sample was gathered as a random sample in 2006 our sample also included older physicians and more specialists in 2015 compared to the eligible population in 2015. Due to sample loss our sample may be biased and this may have led to an underestimation of effects for a number of reasons, for example, because highly stressed physicians may have a higher probability of dropping from the sample.”

COMMENT: 2. On p. 9, line 7: the Cronbach's Alpha of this scale has to be inserted!

RESPONSE: The alpha has been added.

COMMENT: Throughout the text, authors' arguments are closely related to the Finnish context. One wonders whether they might comment on the generalization of their findings to other high income countries, at least in the Discussion section.

RESPONSE: As suggested, we have added discussion about this issue to our discussion:

(page 17, discussion) “In Finland tax-financed universal health care is provided for all residents, therefore generalizing our findings to countries with other types of health care systems or IT systems should be done with caution. However, digitalisation is increasing at a high pace in most developed countries, thus we may assume that SRIS is also a problem in other developed countries given that physicians from all these countries have to face new challenges coming from IS.”
#Reviewer 2

COMMENT: General: The manuscript needs a thorough proof reading from an English native speaker, since it is not smooth and easy to read and has some major grammatical errors.

RESPONSE: As suggested, the manuscript is now edited by a native English speaker, and we hope that the language has improved.

COMMENT: General: The authors use a lot of abbreviations and terms for somehow the same thing: HER, IS, HIS, ICT… could the authors be clearer about the differences or stick to only one term?

RESPONSE: As suggested, we have gone through the manuscript and unified our terms; we have replaced HIS with IS. However, when referring specifically to EHRs we have wanted to keep that term because it is a more specific than broader IS.

COMMENT: Background: unfortunately, to me the manuscript makes the impression of someone finding data and making the best out of it by simply trying out statistical models and not referring to conceptual models and theory. A conceptual model of physician stress factors for example or a stress theory from which the study would be derived, would be very useful. But I am not sure that can be done retrospectively now …

RESPONSE: The theoretical ground for the present study comes from Karasek’s Job Strain model, Kirsch’s Work Load theory and Information chaos theory. Kirsch’s theory already existed in the original manuscript and now we have added information about Karasek’s Job Strain model and Information chaos theory to the background section, it now says:

(page 5, background) “Karasek’s [17] job strain model assumes that high-strain jobs that involve high levels of demands and low levels of control lead to higher psychological strain and physical illness. It is possible that a lack of time to learn to use all the new systems and functions leads physicians to regard IS as being extremely complicated and stressful.”

(page 5, background) “According to information chaos theory [6], information overload and information scatter are sources of information chaos and EHRs may make the situation even worse by encouraging electronic copying and pasting, adding irrelevant information and mixing data.”

(page 5, background) “Support at work has later been added to Karasek’s model as a possible buffer for the effect of high work strain on stress-related illness [22].”

COMMENT: Background and discussion: since there is no conceptual framework, the background and discussion do not have a clear focus, but go from one topic or result to the next without aligning them. Also, causal directions of associations between the many factors in the
models are unclear for the reader. Please try to give it a better and more connected/coherent structure.

RESPONSE: We have added some theoretical background to the study and done some rewriting, therefore we hope that the manuscript is clearer and more understandable now.

COMMENT: Background: in the first paragraph workload and cognitive demands are introduced as consequences of IS, but in the statistical models they are treated as predictors of IS stress. Please comment on that.

RESPONSE: Because there was a discrepancy, we have changed this sentence, it now says:

(page 4, background)“The problems associated with IS may have negative ramifications for patient care, such as problems in clinical performance and patient safety [6].”

COMMENT: Methods: Since the authors do not take up private vs. public care in their models, I think the section on the Finnish healthcare system can be skipped. The section on the EHR adoption can be interesting in this context, but does belong to the background section in my view.

RESPONSE: As suggested, we have deleted the section about Finnish health care and removed the EHR adoption section to background section (page 6).

COMMENT: Methods: I assume that all measures except the one from Oldham and Hackman were self-developed in 2006. I must say that I have concerns about the reliability and validity of the measures. Has there been a validation of these measures beyond the calculation of internal consistency? Have factor analyses or structural equation modelling tested the structure of the scales? Since many items encompass several statements and concepts (e.g. distracted, worried or stressed), I doubt that the respondents had a common understanding of them and a common response pattern. One of the central rules of survey design is not to include more than one concept in one item. Why did the authors not make use of internationally established validated measures? At least, I would disagree that the measures are well-known and validated (as written in the discussion), right?

RESPONSE: We apologize that we did not explain our measures better. All other our measures except cognitive load and staffing problems have previously been widely used. We have now added some examples of previous findings regarding other measures used here. In addition, we have changed the sentence in limitations section, it now says:

(page 16, discussion) “However, to minimise problems with self-reports we used measures that showed good reliability.”
COMMENT: Methods: Also, is there a reasoning for not using all items from 2006 in 2015 again? This is of course a limitation.

RESPONSE: We have added discussion about this to our limitations section, it now says:

(page 16, discussion) “One limitation of our study is that we did not measure all the items from both years, for example, we did not assess all the items from cognitive workload and staffing levels in 2015 that were assessed in 2006. Therefore, our finding that associations of increases in cognitive workload and staffing level with SRIS were not significant should be taken cautiously. Future studies should examine this in more detail.”

COMMENT: Methods: the information on the structural equation model are very limited and require more detail.

RESPONSE: We have added more information on the structural equation modeling in the Methods section. In the text:

(page 11, methods, statistical analyses)“We tested the reciprocal associations between SRIS and other psychosocial factors using cross-lagged structural equation modelling (SEM). In all SEM models, the contribution of the potential confounding factors in the relationships between SRIS and other psychosocial factors were taken into account by using adjusted values (adjusted for age, gender and employment sector), predicted by the linear regression models. The cross-lagged analyses were applied to all available data for individuals who responded during either of the data collection phases using maximum likelihood (ML). The direction for the associations was evaluated by a) evaluating the significance of the associations and b) comparing the fit of the models where either direction of the associations were dropped to the saturated model (where both directions were present). The goodness-of-fit of the models was evaluated using the following fit indices: chi-square, the root mean square error of approximation (RMSEA), the Akaike information criterion (AIC), the Bayesian information criterion (BIC), the comparative fit index (CFI), and the Tucker-Lewis Fit Index (TFI). A non-significant chi-square value indicates that the model fits the data. However, chi-square is highly sensitive to sample size. RMSEA values of less than 0.05 and 0.10 represent a good and acceptable fit, whereas CFI values above 0.90 and 0.95 indicate an acceptable and good fit [31]. In comparing alternative models, a statistically significant improvement in the chi-square value indicated a better fit of the model. We used the complete scales (not the items) in the SEM models. These analyses were performed using the Lavaan R-package (version 0.5-23.1097). “

COMMENT: Methods: What type of regressions were performed? Linear? What about the R² of the models? Why stepwise models? What is the conceptual reasoning behind the steps?

RESPONSE: We have added to the text the information that we used linear regression analyses. Moreover, we have added R²s of the models to tables. Stepwise models were chosen to get more detailed information whether, for example, time pressure attenuates the effect of cognitive load and whether job satisfaction attenuates the effects of time pressure and cognitive load. The
effects of cognitive load might be accounted for simply by high time pressure and the effects of
cognitive load and time pressure might be accounted for simply because respondent is not
satisfied with ones job and therefore stresses more about information systems. This has been
added to statistical analyses section:

(page 11, methods, statistical analyses) “The analyses were conducted in these steps to
find out a) whether time pressure or problems in teamwork would partly account for the possible
effects of cognitive workload or staffing problems on SRIS and b) whether job satisfaction
would partly account for the possible effects of all previously mentioned variables on SRIS. “

COMMENT: Results: How many missings did the variables in table 1 have? How did the
authors treat missings? Listwise deletion? Imputation?

RESPONSE: Because there were missing data in some questions the n varied in analyses
between 1109 and 1009, this information has now been added to sample section, it now says:

(page 8, methods, sample)“Due to missing information for some variables, n varied between
1109 and 1009 in analyses.”

COMMENT: Results: the results of the SEM are not sufficiently presented. A figure of the
model with the loadings etc. is needed and also a table with the standard fit indices. In the SEM
the authors must have had a look on the factor structure of the scales. Doesn't that provide
information on the scales? (see 7.)

RESPONSE: As suggested, we have now provided a figure of the saturated SEM cross-lagged
SEM model (Figure 1) and a table with the fit indices (new Table 2) and reported the SEM
results more fully. It now says in the text:

(page 12, results)“We additionally tested the reciprocity of this association with cross-lagged
structural equation models using the complete scales. The repeated cross-lagged SEM was only
conducted for cognitive workload. As can be seen in Figure 1, cognitive workload in 2006
predicted SRIS in 2015 (βeta = 0.13; p = 0.001), but SRIS in 2006 did not predict cognitive
workload in 2015 (βeta = 0.04; p = 0.155) in addition to autoregressive associations (the model
tested both directions simultaneously). Comparing the model fit indexes (Table 3) also showed
that it was possible to drop the pathway from SRIS 2015 to cognitive workload in 2015 without
significantly reducing the fit of the model (Δχ2 = 2.32; p = 0.128) compared to the saturated
model. In contrast, it was not possible to drop the pathway from cognitive work load to SRIS
2006 without significantly reducing the fit of the model (Δχ2 = 7.89; p = 0.005).”

COMMENT: Discussion: my concern is that the associations between the stress factors as
predictors and the SRIS do mostly reflect common method bias or the general tendency of
increasing stress in physicians throughout the last years. Also, the use of IS in healthcare and the
associated burden has changed from 2006 until 2015. This will definitely affect the results, but
has not really been mentioned or discussed. The changes in IS use for example could have been assessed in the survey as well in order to control for it.

RESPONSE: This issue is now discussed in our limitations section, it now says:

(page 17, discussion) “A lot has happened regarding digitalisation in Finland between the study years 2006 and 2015. For example, Kanta, the national digital repository for electronic patient data, has been launched, almost all physicians now use EHRs and e-prescription is mandatory. Thus, the burden associated with IS may have changed substantially, which may have had an effect on our results.”

COMMENT: Discussion: burnout and other health consequences are mentioned here, but have not been included in the model. That would be very interesting. Does the survey cover these topics?

RESPONSE: We agree that burnout would be interesting in this context, but unfortunately, the survey did not include burnout.

Reviewer 3

COMMENT: Abstract: Description of results could be clearer. Please consider integration of numbers/significances to quantify associations

RESPONSE: As suggested, we have rewritten the results section of the abstract and associated β values for the findings.

COMMENT: Background: A clear research question is missing.

RESPONSE: As suggested, we have added a clear research question to background section: it now says:

(page 6, last sentence of background) “The present nine-year longitudinal study aimed to examine the predictors of physicians’ stress related to information systems (SRIS) among Finnish physicians. The examined predictors were cognitive workload, staffing problems, time pressure, problems in teamwork and job satisfaction, adjusted for baseline levels of SRIS, age, gender, and employment sector. We questioned a) whether the levels of these predictor variables were associated with the levels of SRIS nine years later and b) whether changes in these predictor variables from the year 2006 to 2015 were associated with SRIS in 2015.”

COMMENT: How did you select the predictors?
RESPONSE: The theoretical ground for the present study came from Karasek’s Job Strain model, Kirsch’s Work Load theory and Information chaos theory from which the other predictors came from except job satisfaction which has previously been found important in so many contexts that it was thought relevant also here.

COMMENT: Is there a theoretical foundation?

RESPONSE: As mentioned above the theoretical ground for the present study was from Karasek’s Job Strain model, Kirsch’s Work Load theory and Information chaos theory. In addition to Kirsch’s theory which already existed in the previous version of the manuscript, we have added Karasek’s Job Strain model and Information chaos theory to the revised version, it now says:

(page 5, background) “Karasek's [17] job strain model assumes that high-strain jobs that involve high levels of demands and low levels of control lead to higher psychological strain and physical illness. It is possible that a lack of time to learn to use all the new systems and functions leads physicians to regard IS as being extremely complicated and stressful.”

(page 5, background) “According to information chaos theory [6], information overload and information scatter are sources of information chaos and EHRs may make the situation even worse by encouraging electronic copying and pasting, adding irrelevant information and mixing data.”

(page 5, background) “Support at work has later been added to Karasek’s model as a possible buffer for the effect of high work strain on stress-related illness [22].”

COMMENT: Please comment on a potential conceptual overlap between some predictors as well between predictors and outcome.

RESPONSE: We have added discussion about this issue to our limitation section, it now says:

(page 16, discussion) Our results may have been affected because there was some overlap in our predictor variables and outcome, given that many of our variables dealt with the experience of stress (coming from different bases).”

COMMENT: Methods: There is also an overlap when you look at the measures. What does that mean for the analyses and the results?

RESPONSE: Please, see response above

COMMENT: Why did you chose a stepwise approach? A hypothesis/research question for this is missing (see above: research question).
RESPONSE: Stepwise models were chosen to get more detailed information whether, for example, time pressure attenuates the effect of cognitive load and whether job satisfaction attenuates the effects of previously used predictors. The effects of cognitive load might be accounted for simply through high time pressure and the effects of cognitive load and time pressure might be accounted for simply because respondent is not satisfied with his or hers job and therefore stresses more about information systems. This is now more clearly indicated in the statistical analyses section:

(Page 11, methods, statistical analyses) “The analyses were conducted in these steps to find out a) whether time pressure or problems in teamwork would partly account for the possible effects of cognitive workload or staffing problems on SRIS and b) whether job satisfaction would partly account for the possible effects of all previously mentioned variables on SRIS. “