Author’s response to reviews

Title: Methodological limitations in experimental studies on symptom development in individuals with idiopathic environmental intolerance attributed to electromagnetic fields (IEI-EMF) – A systematic review

Authors:

Kristina Schmiedchen (schmiedchen@femu.rwth-aachen.de)

Sarah Driessen (driessen@femu.rwth-aachen.de)

Gunnhild Oftedal (gunnhild.oftedal@ntnu.no)

Version: 2 Date: 15 Aug 2019

Author’s response to reviews:

Reviewer #1: Overall this is a much stronger manuscript than the previous submission and the authors have carefully taken into consideration the reviewers comments. The following comments are based on the new changes to the manuscript and are suggestions on how to make these section clearer to the reader.

Response: Thank you very much for this positive overall statement. Again, your feedback on the revised version was valuable and very constructive and we have made every attempt to fully address your comments. Please find below our specific responses.

P. 12 lines 9 - 14 makes it sound as if some analyses were conducted using t-tests and others Mann-Whitney-U tests; however, given that all of the data was not normally distributed (p. 16 lines 1 - 3) it appears no t-tests were conducted. See comments below on how to help make this clear to the reader.

Response: We agree with you that our previous description suggested the use of Student’s test for some analyses which was the option if the assumptions for applying parametric tests had been fulfilled. However, as we later state that none of the data were normally distributed, we understand that wording may have been misleading and we have modified the respective paragraphs accordingly.

P. 12 line 54: specify which online tool was used to calculate the power analysis.

Response: We have added that the ClinCalc online tool was used to estimate the power of the conducted analyses and added some more information on the selected options. We modified the
sentence as follows: The power estimate was calculated using the ClinCalc online tool [50], by selecting the options “Two independent study groups” and by assuming Student’s t-test, i.e., “Continuous (means)” was selected as primary endpoint.

P. 12 line 54: states that power analysis was conducted on Student's t-test; however, in none of the resulting analyses were t-tests utilized. Instead given that non-parametric tests were used for all analysis and non-parametric tests are less powerful than parametric tests, the power analyses should be based on Mann-Whitney-U tests.

Response: We agree with you that estimating the power for Mann-Whitney U-test would have been ideal. However, this cannot be done analytically (similar to the case for Student t-test) and thus, no on-line tool is available for doing so, at least to our knowledge. While the power could have been estimated by applying Monte Carlo simulation, the estimated value would anyway only be an indication of the true power. When estimating the power, we had to make assumptions about the distribution of the data based on the current dataset and about the difference in means between the two study groups, both assumptions include uncertainties. Since both the variation of data (standard deviation in case of Student’s t-test) and the difference in mean between the two study groups strongly influence the power estimate, we regard it as sufficient to estimate the power roughly by assuming a Student t-test. As was already stated in the discussion (previous version: p. 21, lines 57-59), the power for the Mann-Whitney-U-test is slightly lower than for Student’s t-test. Therefore the provided estimate of 0.45 shows that the statistical power is low, which is the main point. We acknowledge that it would be useful for the reader to provide the relevant information about statistical power for Mann-Whitney-U-test compared to Student’s-test already under in the method part. We have therefore added to section 2.6 that the power would be somewhat lower for the Mann-Whitney-U-test than the estimated value based on Student’s t-test, and have removed this information from discussion where we now only stress that the power of the conducted analyses was very low.

P. 16 lines 1 - 7: this should be moved to section 2.6 (statistical analysis) in which you describe the data analysis. It would be helpful for the reader to know going into the results section that although parametric tests were planned given that the data were not normally distributed non-parametric tests were conducted.

Response: Thank you for this suggestion. We have now removed that t-test was used and included in section 2.6. the following sentence:

Student’s t-test would have been applied if assumptions for this parametric test had been fulfilled, which was not the case: all data distributions differed statistically significantly from normality (Kolmogorov-Smirnov test) and one of the sample sizes was low.
P. 16 lines 1 - 3 - specify which statistical analyses were used to test for normality of the data.

Response: We have specified in the methods (section 2.6.) that Kolmogorov-Smirnov-test was applied to test normality (see response to previous comment).

P. 16 lines 40 - 50 move to section 2.6 (statistical analysis) as this would fit in better in that section rather than the results section.

Response: Agreed. We have moved –as you suggested – this paragraph to section 2.6. Thus, in the revised version, all information regarding the power estimate (parameters, result) is now provided in the methods.

P. 16 line 50 states that the power was estimated to be .45. This is very low and should be noted in the test along with the power analysis.

Response: We comment on the very low power in the discussion in section 4.1.2 as we believe this would better fit there. However, along with the result of the power estimate in section 2.6., we have added the following sentence: “A power estimate of 0.8 or greater is commonly regarded as sufficient to detect a true effect.”, which implicates that our estimated power is comparatively low.

P. 17 line 47 - p. 18 line 28: It would be helpful to separate this text into two separate paragraphs for clarity. One paragraph should focus on homogeneity of participant (p. 17 line 47 - p. 18 line 5) and the other should focus on medical or mental disorders that explain symptoms (p. 18 line 5 - line 28). As the text currently reads the argument does not make sense. It would be helpful to restructure the argument to first talk about how the majority of studies used heterogeneous sample; however, those that have use homogeneous sample has also not found an effect. Even though heterogeneous samples may be one reason for different results the data from studies that use homogeneous sample do not support this as such and this should be made clear to the reader.

You might want to restructure the paragraph along the following lines:

In 23 (82%) studies, the selection of study participants may have introduced substantial bias in the results. These studies did not consider a careful assessment of potential participants before enrollment in the study (i.e., at least one key question within the domain selection bias was judged to be at high risk of bias). Heterogeneous study groups may have been the result of such selection procedures and the applied exposure parameters were probably not appropriate for all participants. However, five (18%) studies had applied strict inclusion criteria in an effort to
identify individuals with symptoms that could be related to EMF exposure and ensured that the
participants and the design of the experiments were matched [22,52,58-60]. It is important to
note that four [22,58-60] of these five studies were unable to find an effect of exposure on well-
being or symptom development.

Response: Thank you for this advice. We have restructured this paragraph as you suggested, i.e.,
we started with heterogeneity and then point out the five studies that used homogeneous samples.
With this first paragraph we actually intended to give a short summary on selection bias before
 going on with the discussion of specific key question that were often judged to be at high risk of
bias (the exclusion of individuals with somatic diseases and mental disorders that might explain
the EMF-related symptoms is just one of the key questions that we discuss in more detail.). To
make this clear, we have included the following transition between the two paragraphs: “One
specific limitation in the selection procedure relates to a lack of screening for somatic diseases or
mental disorders that might explain the EMF-attributed symptoms”

P. 21 lines 57 - 59 - see comment above referring to recalculating the power estimate for a
Mann-Whitney-U test.

Response: Please see above for response to your comment regarding power analysis based on
Mann-Whitney-U-test.

p. 17 line 62 - enrollment is misspelled

Response: In our manuscript we use British spelling, i.e., “enrolment” is the correct spelling

Reviewer #2: Dear authors,

Thank you for your detailed answers and the comprehensive revision of your manuscript. I
appreciate your efforts and am mostly convinced, apart from a few quibbles detailed below.
Stylistically, I feel the text has become somewhat clumsy at places and would benefit from a
global streamlining (especially, to shorten some sentences and avoid repetitions). You will find a
few suggestions below.

Response: We are very thankful for your time and interest you spend on re-reading and
commenting on the revised version of our manuscript and we are happy to hear that you mostly
agree with the changes made to the text. We understand that the text was unbalanced in places
and included several repetitions. We have therefore shortened or rephrased some of the sentences
and tried to avoid repetitions where possible (with the latter referring in particular to the
descriptions and results of the power analysis which have been moved to the methods, Section
2.6.).

1. Regarding the exclusion of studies investigating physiological parameters

All right, I understand your position and find it better exposed in the manuscript. I still regret no
way could be found to analyse objective and subjective results jointly, but maybe that is simply
not rigorously possible. The title and abstract are coherent with your approach now.

Response: We also regret that no way could be found to jointly analyse subjective and objective
outcomes, but as you say, such an approach could not be implemented straightforwardly. Good
to see that you were satisfied with the new title and the explanation for why we did not include
objective outcomes in our analysis.

2. Regarding the exclusion of individuals with other diagnoses

I follow your reasoning and appreciate your efforts to present it in §4.1.1, even though I am not
totally satisfied by the wording. If I understand correctly, what is at stake is the risk of including
individuals who are misattributing symptoms to EMF exposure in EHS groups. A
straightforward way to reduce this risk is to exclude individuals whose symptoms could be
explained independently from EMF exposure, by an objectively diagnosed somatic disease or
mental disorder. However, as you remark in your manuscript, it can prove quite challenging to
determine whether such diseases or disorders constitute definitive medical explanations of the
symptoms experienced by any given EHS individual. Consequently, in practice, individuals with
any medical diagnosis tend to be excluded from experimental studies of IEI-EMF, whether or not
their diagnosis could adequately explain their EHS symptoms. This might induce a secondary
bias against truly EHS individuals who happen to suffer from

additional diseases or disorders. Given that such individuals are presumably very few, and that
potential EMF physical effects are likely of low magnitude, this secondary bias can be regarded
as less important than the risk of including individuals who are misattributing symptoms to EMF
exposure in EHS groups. However, as you also state, other and maybe preferable ways or
reducing this risk could be to question carefully potential participants about their symptoms
when and when not exposed, or to observe their reactions during habituation sessions.
If that makes sense you, the wording should be amended throughout the paper to make clear that individuals are excluded because their symptoms might be, instead of have been, explained independently from EMF exposure.

Response: We understand that the wording was not concise throughout the manuscript (including the figures) such that it now reads: “individuals with somatic diseases or mental disorders that may explain their EMF-attributed symptoms”. Also, it is correct that what matters is to reduce the risk that individuals with other reasons for the symptoms which they misattribute to EMF-exposure are included in the study.

Furthermore, is there any somatic disease or mental disorder directly related to EMF exposure? I believe not. Therefore, you could lighten your wording by simply alluding to somatic diseases or mental disorders, without adding "unrelated to EMF exposure" every time.

Response: You are right. We therefore removed “unrelated to EMF exposure” where we talk about somatic diseases and mental disorders.

p.18, l.33: Are [54] and [55] the only studies where results were calculated for IIEI-EMF and control groups as a whole? If so, you could make it clearer, to give more weight to your argument. If not, why focusing on these two studies?

Response: What was specific to these two studies is that they are the only which concluded about an effect of exposure only based on the analysis of the combined data from the groups of individuals with IIEI-EMF and healthy controls. We have added the specification to this paragraph.

Also Hillert et al. (2008) analysed the effect of exposure for the combined groups of participants, but did not base their conclusions about an effect solely on this analysis. They used ANOVA with groups of participants and sessions as factors and thus also looked at the main effect of session.

3. Regarding the nocebo effect

I agree with the corrections you made. Your assessment in §4.1.3 and §4.3 is sufficiently clear indeed. To make it even clearer, you could maybe reformulate p.24, l.47-59.

In your text: "Thus, the nocebo effect could either overshadow physical effects or may otherwise be the only reason why symptoms are experienced in everyday life. Alternatively, in previous studies, only a few of the individuals believing that their symptoms are caused by EMF exposure
were right, and in most cases, factors unrelated to EMF exposure may explain the symptoms (e.g., somatic diseases, psychosomatic disorders, or imbalance in the autonomic nervous system, as specified in the Introduction)."

Suggested wording: Thus, the nocebo effect could either overshadow physical effects, add to symptoms resulting from somatic diseases or mental disorders unrelated to EMF-exposure, or be the only reason why symptoms are experienced in everyday life.

Response: We are fine with your suggestion to reformulate this paragraph. We therefore included your wording, but replaced “resulting from” with “provoked by”.

4. Regarding the content of reviewed studies

Thank you for adding details on the outcome measures in Table 2 and a short discussion in §4.1.1. It is quite interesting and give more weight to your manuscript. Do you feel that the lack of standardization is a problem here, and that further studies should preferably use validated tools?

Response: Yes, the lack of standardization can be problematic in terms of comparability of studies. We have added this to the text, but also indicate that standardization should not be used on the expense of a match with experienced symptoms in everyday exposure situation.

I strongly dispute your assertion that cordless phones are not a relevant source of exposure because the tend to fall out of use. Even if mobile telephony is indeed becoming the favoured technology, many people, especially among older generations, still possess cordless phones and use them at home. Moreover, in France at least, all Internet service providers supply their clients with residential gateways that include DECT base stations. These stations have been observed to generate EMF continuously, even when not in communication. Consequently, virtually anyone having a landline Internet connection is exposed to DECT signals, whether or not owing a cordless phone. The same reasoning goes for Wi-Fi signals.

I understand that when assessing the quality of each study, what only matters if the adequation between the signals used for experimental exposure and the subjects' sensitivity. But when assessing the scientific literature as a whole, one might regret that certain types of signals have never been studied. Don't you think this is worth discussing among research needs (§4.3)? Why not suggesting to study technologies that have not been investigated to date?

Response: We agree with you the DECT system is still a relevant source of exposure, but its use may differ among countries. However, this is not the main reason for why we decided not to comment on testing every common type of signal. As you also mention, what matters is the relevance of the source and type of exposure and that is what we have chosen to stress. Other types of signals that have not been tested yet can of course be applied, but since a number of
different types of sources have been tested already, we would hesitate to recommend testing other types to cover every alternative. Given the rapid technological development, a lot of new types of sources and signals will appear, but this should not be the reason to do conduct further experimental studies.

5. Case definition

Thank you. I find this section more convincing now, obviously, while feeling it deserve some rewriting for concision and clarity.

Response: Fine. As suggested, we have made slight modifications to this section, including rephrasing and shortening of sentences.

Please find below some additional suggestions for wording or presentation.

Response: Thank you very much for your suggestions. We feel that most of them are more precise than the ones we used and have included your suggestions in the manuscript. Otherwise we have commented on the use of different wording.

p.4, l.57: "and that these individuals could be masked in heterogeneous study groups that include individuals with e.g. somatic diseases or mental disorders" / who could be masked in heterogeneous study groups that include subjects misattributing somatic diseases or mental disorders to EMF exposure.

Response: Ok, but changed “subjects” to “individuals”, “who” replaced with “their responses”, and added “symptoms provoked by e.g.” before “somatic diseases or …”

p.5, l.42: "the experimental research designs differ between studies investigating symptom development and those investigating physiological and cognitive functioning" precluding their joint analysis in this review.

Response: Ok, “precluding their joint analysis” has been added.

p.10, l.9: start a new paragraph here

Response: Ok.

p.10, l.17-24: "we assessed whether the study potentially included individuals with diseases/disorders that may equally explain the health complaints but are unrelated to EMF exposure and whether the study design was inappropriate for the included participants" / we assessed whether the study included individuals with somatic diseases or mental disorders that may explain their symptoms, and whether the study design was appropriate for the included participants"
Response: Ok. Further added “their EMF-attributed symptoms”

p.10, l.31: "i.e., which may be due to e.g. habituation effects" / e.g., which may be due to habituation effects

Response: We noted that we have to be more precise here to differentiate between sequence and period effects. So this sentence now reads: “Under performance bias, we considered biases related to knowledge of which exposure condition was used and biases related to the sequence (which may be due to carry-over effects) and to the period of the exposure conditions (which may be due to habituation or variable stress levels as a function of time).”

p.11, l.14: "When the power to detect an effect of exposure is low because of a low number of participants or repetitions of trials" / When the power to detect an effect of exposure for lack of enough participants or trials

Response: Ok, but obviously you meant to say “When the power to detect an effect of exposure is low due to …” We further removed “a lack of enough participants or trials” by “too few participants or trials.”

p.14, l.43-5: "with diseases or disorders that may equally explain the symptoms but are unrelated to EMF exposure" / with somatic diseases or mental disorders that may explain their symptoms

Response: Ok, Further added “their EMF-attributed symptoms”

p.15, l.28-33: "because they did e.g. not provide a statistical power estimate (n=21, 75%) and/or because they did not adjust for multiple comparisons" / because they did not e.g. provide a statistical power estimate (n=21, 75%) and/or adjust for multiple comparisons

Response: Ok.

p.17, l.60: "These studies did not consider a careful assessment of potential participants" / These studies did not carefully assess potential participants (if that is what you mean)

Response: Ok.

p.18, l.30: "one can argue that the inclusion of healthy individuals in the group of IEI-EMF individuals" / one can argue that the involuntary inclusion of healthy individuals in IEI-EMF groups

Response: “unintentional” seems more appropriate here

p.19, l.40: "many different tools (e.g., validated or customized self-report questionnaires) were used to record the symptoms, but standardized questionnaires developed for studies with IEI-EMF individuals [6] have been rarely used [21,61]." / … were used to record the symptoms. Standardized questionnaires…”
Response: Ok. Further added “However, standardized…”

p.25, l.21: "area, resulting in less public concern with this topic which would support the notion of" / resulting in less public concern for this topic. This would support the notion of

Response: Ok.

p.28, l.1: "Because participates will" / Because participants

Response: Changed

p.28, l.7-11: "However, fear or stress levels caused by the experimental situation can be reduced by habituation sessions (as already included in the assessment tool), but also by other approaches such as at-home testing [81]." / Further studies should try to minimize the stress caused by the experimental situation, through habituation sessions or other approaches such as at-home testing [81].

Response: Ok, added “stress level” and changed “further studies” into “future studies”.

p.29, l.33: "a high study quality must be ensured to" / a high quality study must ensure to

Response: We have removed this term from the manuscript and use different wording, e.g. “We encourage researchers to achieve a high credibility of the results …”

Reviewer #3: Dear Authors,

I have re-read the manuscript and supplemental material and I am satisfied with the adjustments made and I think the ms reads well; thank you for your gracious consideration of my recommendations. All the best with your research!

Response: We are happy to hear that you were satisfied with our revised version of the manuscript. Thank you again for all your valuable comments and suggestions for being more concise with the analyses and descriptions. They were very much appreciated!