Author's response to reviews

Title: Internet-based Self-Assessment as Valuable Monitoring for Public Mental Health

Authors:

Stefan Vetter (stefan.vetter@access.uzh.ch)
Astrid Rossegger (astrid.rosegger@me.com)
Thomas Elbert (thomas.elbert@uni-konstanz.de)
Juliane Gerth (juliane.gerth@gmail.com)
Frank Urbaniok (frank.urbaniok@ji.zh.ch)
Mario Mueller (mario.mueller@dgsp.uzh.ch)
Wulf Rossler (roessler@dgsp.uzh.ch)
Jérôme Endrass (jerome.endrass@me.com)

Version: 3 Date: 2 February 2010

Author's response to reviews: see over
**Reviewer:** Hans van Oers

**Reviewer's report:**
All the comments given in the first review report have been addressed well by the authors. Methods and statistical analyses are described in a more clear way, better understandable for the reader, and the discussion section has improved very much. One minor remark: titles of the two figures are missing.

Thank you.
We added the titles of the two figures.

**Reviewer:** Peter van der Velden

**TITLE**
The title is incorrect. The authors do not present any data that supports the title, especially that it valuable. The authors present the design of the study, characteristics of those who participated and their experiences with this site. Since there are many limitations, the title should express something like “Internet-based Self-Assessment after the Tsunami among Swiss: lessons learned”.

We followed the suggestion of Mr. van der Velden to adapt the title of the manuscript. The new title is “Internet-based Self-Assessment after the Tsunami: lessons learned”

**ABSTRACT**
The abstract should be re-written to provide a better summary of the background, methods, results and conclusions. The last sentence of the conclusions is not supported by the study, since no data is available about the use of mental health services.

We have rewritten the Abstract.

**INTRODUCTION**
The introduction remains puzzling and problematic. For instance, the authors state “In the latter (i.e human-made/technological disasters,) there is usually an obvious group of direct victims, whereas in the former (i.e. natural disasters), it is more difficult to identify who was affected and to what degree and thus to assess the impact on public mental health” and later on page 3 “However, the problem remains, that in natural disasters it is unclear, who was affected by the disaster and therefore who should be screened”. First, I do not know any study that supports that statement (I do not believe that in principle this identification problem has anything to do with “human-made/technological disasters”
The introduction was shortened and clarified.

Second, the authors still do not explain why all victims should be screened. As said before, referring to the NICE guidelines is insufficient (cf. Wessely, S.C. (2003). The role of screening in the prevention of psychological disorders arising after major trauma: Pros and cons. In: R.J. Ursano, C.S. Fullerton, & A.E. Norwood (Eds.), Terrorism and disaster. Cambridge: Cambridge University Press).

We do make a reference to the NICE guidelines but it is not the purpose of paper to state that all victims need to be screened nor do we find it necessary to discuss this matter into great detail. We decided to follow the NICE guidelines. The reader can make up his mind whether he or she finds this reasonable.

Furthermore the authors state “.. the result of the screening relies fully on self-assessment and Internet-based questionnaires are likely to inflate scores [9] leading partially to false positive or over-diagnosed individuals. On the other hand, there is evidence that psychometric properties of Internet-based questionnaires are not biased and that questionnaire format and presentation order do not affect rates of psychological symptoms reported by participants [10]. There is evidence, that there are no significant differences between assessment techniques [11], suggesting that Internet-based methods are a suitable alternative to more traditional methods”.

Remarkably, on page 8 the authors wrote “For the PTTS-10 the cut-off was raised to 14 (compared to 12 defined in the PTTS-10 manual), since it has been shown that self-evaluations via Internet might lead to up to 10-20 % higher scores [9]. Within the 10-item depression-scale we set the cut-off at 9. This line of reasoning is very difficult to follow, if not contradictory.

We do not find this difficult to follow. What we address here is the problem of calibration. There is empirical evidence supporting the hypothesis, that the psychometric properties of online versions of questionnaires are not systematically biased when compared to paper-pencil versions. We cited however literature that suggests, that web-based instruments need to be recalibrated, since there is 10-20% of over-reporting. We followed this suggestion by adapting the threshold accordingly.

The authors refer to the Harvey et al. post-disaster study, stating that “One of the first applications of an Internet-based instrument was the employment of a screening instrument after the hurricane Katrina. The web-based monitoring system was used to assess the traumatic experiences of 102 hurricane disaster evacuees before and after the hurricane, and to measure self-reported levels of psychological distress [12]. This strategy was found to support individuals who could not get direct treatment from mental health professionals.

However, the authors of this Harvey et al. study tells us something totally different: “While the web-based system has the facility to deliver evidence-based therapy (e.g.,
cognitive behavioral therapy), this feature was not used for this project as researchers wanted to measure only the impact of the Hurricane Choir on mental health”. Their conclusion about support was not about the web-site, but in particular: “It should be recognized that the choir provided a strategy to assist survivors of the hurricanes who might not otherwise have been able to get direct psychological support from mental health professionals”. Such misleading remarks should be omitted.

We deleted this section.

In the revised manuscript, again clear research questions are absent.

*The objective of the study was made clear in the manuscript: “The objective of this article was to report the experiences that were made using an Internet-based, self-screening instrument after a large-scale disaster.” The design of the study was clearly naturalistic and thus formulating a clear research question would have had a post-hoc character.*

**METHOD**

**Instruments**
The used instruments need a more detailed and more structured description, such as the scoring/range of the scales, number of questions related to depression, anxiety, obsessive-compulsive disorder and suicidal ideation. Since no validated questionnaires with respect to the aforementioned disorders/symptoms was used, a precise description of the used questions should be presented. Perhaps it is better to present them as symptoms of .... I assume that the authors focused of tsunami-related PTSD (that respondents were asked to fill in the scale with the tsunami in mind). If this is correct, this should be added.

PTSD measure, in accordance with *dsm III, dsm III-r, dsm IV or dsm IV-tr*?

Referral to doctor: were the respondents recommended to visit their doctor when they scores high on one or more scales? Why was suicidal ideation not included in this algorithm?

*The number of items is listed in the paper. The items are dichotomous. Otherwise is would not have mad sense to allocate one point for a “yes” answer. However, we have made this clear by explicitly stating that the items were dichotomous. DSM III-R was used, respondents were recommended to see a physician if at the score of at least one scale was high, and suicidal ideation was included in this algorithm.*

We have rewritten the mentioned section and clarified the concerns raised by the reviewer.

**Study population**
As said, all publicity efforts were aimed at Swiss affected citizens. Therefore, there is no reason to include respondents from other countries.

As said before, the absence of any information about mental health services utilization after the disaster (especially).
Since ONSET was also used by citizens from other countries we do not see why non-Swiss citizens should be excluded. The percentage of Swiss respondents was made clear.

Analyses
It is still unclear on which arguments and how the five distinct groups were formed and to what extent members of the five groups experienced other possible stressful events. This information is crucial because the analyses are based on these five groups. Is witnessing the tsunami less stressful that the death of a family member? Why did the authors use t-transformed scores. In either way, raw mean-scores and SD should be provided. For suicidal ideation and re-experiencing, logistic regression was used instead of anova. The authors need to explain this strategy (and used cut-off scores), to be able to interpret the findings.

We did not use “t-scores” but rather “T scores”. Using standardized scores corresponds to common practice and does normally not need to be explained since the benefits are so obvious. Providing raw mean-scores is not necessary since it does not contribute in any way to the clarity of the paper nor does it provide any information that a reader would need in order to understand the results. Contrary to the reviewer we believe that dumping raw scores and coefficients that are not discussed on the reader will make the paper more difficult to understand. Suicidal ideation as well as re-experiencing are dichotomous questions and thus using logistic regressions and reporting OR are the methods of choice and do not need to be justified.

RESULTS

Exposure
According to the manuscript, about 60% of the users were not directly affected by the Tsunami. It is unclear what is meant by that.

In table 2 the types of exposures are listed. Those respondents who were not exposed were referred to as “not being directly affected by”. We have clarified this in the manuscript. Since this interpretation was given just above the table containing all the relevant information, we believe that there is no further specification needed.

Psychiatric symptoms
What is meant by traumatic alarm reaction? The authors state that 45% of the users reported a degree of symptomatology relevant to PTSD, suggesting that the fulfilled or almost fulfilled the criteria of PTSD. Given the “fact” that about sixty% of the users were not directly affected by the disaster, this prevalence is rather high! For all five groups probable PTSD prevalences should be provided.

Selye identified the alarm response as a first stage in his model: When the threat or stressor is identified, the body’s stress response is a state of alarm. The periphery is
activated via a fast and a slow route: sympathetic activation will release adrenaline from the adrenal medulla in order to bring about the fight-or-flight response. The major defense system of the body, the HPA axis, begins to produce cortisol within minutes. Depending on the individual experience, the fast route may also be dominated by a parasympathetic response with vasovagal modulation of a dissociative shut-down.

For all five groups we added probable PTSD prevalence using a cut-off score of 14 points on the PTSS.

How do the authors explain PTSD scores among those who were not directly witnessed the Tsunami themselves, were not injured themselves, did not lose a family member, had no injured family member, did not lose friends / acquaintance, had no injured friends / acquaintance, and did not lose lost property? Is this an indication that the questions about exposure are insufficient?

We do not believe that the questions about exposure were insufficient. People who did not directly suffer from the Tsunami but from other traumatic events apparently felt that ONSET addressed questions that were relevant for them. The lesson learned would rather be to ask more questions regarding previous traumatic events.

Depression, anxiety and obsessive-compulsive: better to say “symptoms of.”

Table 5: please provide details. ANOVA (F-value, DF, N). Add note explaining G1-G5

The description of scores is not uniform, for example (M=52.1 (SD=?), Mgroup3=9.4, Mgroup4=10.4). (Group 5, M=49.1, SD=9.6).

We have provided the necessary information in the legend and were more careful referring to e.g. “symptoms of anxiety” instead of “anxiety”.

The authors further focus on nightmares and frequent reexperience of the disaster. They wrote “One out of five victims (19.8%, N=32) who witnessed the Tsunami waves reported frequent nightmares and flashbacks of the incident. In the group of those subjects who were not directly affected by the Tsunami, only 3.0% (N=51) reported a frequent reexperience of the disaster”. Why were nightmares compared to reexperience of the disaster in general? (or are we talking about the same thing). Moreover, these findings raise serious questions with respect to PTSD. If 19% of those who witnessed the tsunami reported intense levels of nightmares and flashbacks of the incident (reexperiencing), it is almost impossible that 45% of all users reported a degree of symptomatology relevant to PTSD.

The PTSS-10 includes two questions concerning nightmares (question 2) and flashbacks (question 4). DSM-IV-TR subsumes both symptoms as reexperience respectively as B criterion. Therefore we report about users scoring positive on one or on both of these two questions. – Even without scoring in questions 2 & 4 (sum of 0 points) individuals may obtain screening results/scores that surpass the cutoff.
for a potential PTSD pathology, explaining the observed difference between the prevalence of reexperiences and positive screening results.

According the authors the prevalence of suicidal ideation varies across the five strata between 21.5% and 25.7%!! This is almost incredible high compared to for example Katrina victims.

_Beyond any doubt, these are relatively high prevalence rates of suicidal ideation. Nevertheless, we do not consider them as representative for the population of all tsunami victims. The robustness of this result has to be investigated in subsequent investigations._

How many people were give advice to visit their GP/FP? In total and how many in the three periods?

_About 4 out of 10 users had been given the advice to contact their GP/FP._

**DISCUSSION**

The absence of clear research questions is very prominent in the discussion paragraph. The authors inform us about their experiences with several aspects of the site (how it started, copyright problems, only a “validated” PTSS questionnaire, etc), but not in a very structured way. Below I will give a few examples.

Again, this paragraph has some puzzling remarks such as “For ONSET – besides the PTSS – due to copyright issues and the difficulty of the multi-lingual design of the instrument, no validated instruments could be used and thus it was not possible to compare the ONSET users to a norm population. Since the developers of ONSET did not have the time to commission a legal expert opinion regarding data protection of a web-based mental health instrument, it was decided to guarantee very strict anonymity to users. This decision led to a severe limitation in interpreting the data, since no pre-existing traumatic events were assessed. Theses serious difficulties lead to a first, not very surprising conclusion, namely to develop an online instrument before a catastrophe”.

I what way does anonymity prohibit assessing pre-existing traumatic events? The discussion opens with “Besides the financial aspect, in order to reach out to the victims, the authorities would have to know who was affected”, suggesting that the Swiss authorities had no addresses. I do not understand this remark. The study of Kraemer et al (2009) from the same University among “the same” Swiss tourists who were affected by the Tsunami had the addresses of the Swiss tourists. According to this study of Kraemer (2009): “For the period of 1 month, 26 December 200425 January 2005, a tsunami helpline was operating day and night at the Swiss Federal Department of Foreign Affairs (EDA), Bern, Switzerland. A total of 3855 addresses of missing persons, couples and families were reported to the EDA by relatives, friends or anyone worried about a person possibly located in the region affected by the tsunami”. Please explain why the authors come to a lower estimate (“It is estimated that 2000-3000 Swiss tourists were affected”).
The number of 3855 addresses of missing persons, couples and families were reported to the EDA by relatives, friends or anyone worried about a person possibly located in the region affected by the tsunami. We on the other hand report an array of possibly really affected tourists. By far not all initially reported persons were actually in the region or in a devastated area (other holiday region respectively e.g. in North Thailand). Therefore the figures do not match.

In addition the authors wrote “The prevalence of 45% of the users reporting PTSD relevant symptoms differs substantially from epidemiological studies, which had reported a PTSD prevalence of 4.5% after a large-scale disaster [8] and is at the upper bound of the 5-60% interval presented by Galea et al. [1]. This finding is, as said before, remarkable and serves as an indication that the PTSD instrument used needs a better calibration. However, the authors come to the conclusion: However, results gathered from ONSET users after the Tsunami-wave are not conclusive in the sense that they do not give a robust estimate of the PTSD prevalence in the Swiss citizen affected by the Tsunami since we don’t have reliable information of the ONSET outreach. Besides the general difficulties and uncertainties that accompany the implementation of a new screening strategy, ONSET was developed as a cross-sectional instrument. Carrying out a screening just at one appointed date might fail to reach all affected individuals”. I do not understand why high prevalences are due to the cross-sectional nature of this study.

We did not say, that high prevalence was due to the cross-sectional nature of the study – this is an interpretation of the reviewer. We were cautious regarding the interpretation of the high prevalence of PTSD relevant symptoms in our sample. The reason for the caution was, that we do not know how representative our sample was (“outreach”) and that we were not able to follow-up our sample (“screening at just one appointed date”).

Thus, I my view the discussion needs a major revision, starting with clear and testable hypothesis /research questions in the introduction. After discussing the outcomes, I believe it is very interesting to end with a box summarizing suggestion for future internet use.