Author's response to reviews

Title: A cross-national survey on patient safety culture: Japan, Taiwan and the U.S.

Authors:

  Shigeru Fujita (sfujita@med.toho-u.ac.jp)
  Kanako Seto (setokana@med.toho-u.ac.jp)
  Shinya Ito (s1007m09@med.toho-u.ac.jp)
  Yinghui Wu (cutesun_712@hotmail.com)
  Huang Chiu-Chin (cchuang@must.edu.tw)
  Tomonori Hasegawa (tommie@med.toho-u.ac.jp)

Version: 4 Date: 13 November 2012

Author's response to reviews: see over
Dear Mr. Danrolf de Jesus,

We are grateful to reviewers for the critical comments and useful suggestions that have helped us to improve our paper. We have taken all these comments and suggestions into account in the revised version of our paper as follows.

Comments from editor

1. The title is not consistent with the content. “A cross-national survey” sounds as though the authors collected data from more than one nation, which is not strictly the case, as they used existing AHRQ data (which is in the public domain) but currently present it as though it were original data. This needs to be made much clearer.

→ The title was changed as follows:
Pre: “A cross-national survey on patient safety culture: Japan, Taiwan and the U.S.”
Post: “The characteristics of patient safety culture in Japan, Taiwan and the U.S.”

2. Before the Japanese data were collected (January 2010), the Taiwanese results were already presented and compared with US (2009). A better explanation is therefore needed for what was already in the literature prior to the present study, what was actually done "de novo" for this present paper, and how this adds new understanding to the area. Better explanation is also needed as to why these three countries were chosen for comparison.

→ Chen et al (2010) compared Taiwanese results with US results. But their subjects were limited in physicians, nurses and administrative workers at teaching hospitals, and the number of respondents was only 788. We added following sentence. (Background, P.4, L.19-23)
“Chen et al (2010) compared Taiwanese PSC with US PSC using statistical analysis, showed some characteristics of Taiwanese healthcare workers.[7] But the subjects were from only teaching hospitals, and the sample size seems to be not enough to understand the characteristics of Taiwanese healthcare workers.”

As for the reason to choose the 3 countries, we added following sentence. (Background, P.5, L.2-6)
“We hypothesized that the characteristics of PSC among East Asian countries, whose culture was closely related with one another, were similar, although they were different from those of western countries which had heterogeneous customs such as the interpersonal relationships.”
3. The study says that the Japanese version of the PSC was developed and tested by Ito et al. (2011). How was it possible to use that test in 2009?

→The data set of Ito et al (2011) is the same as our data set although we added more data. Ito’s paper was published in 2011. But the data was collected in 2009.

4. There are very different response rates among the three countries, a point which is not adequately addressed / discussed at present.

We mentioned about this issue in the limitation of our study. (p.12, L.5-10) But the reasons of low response rate in the U.S. were unknown because the situation in each US hospital was not open to the public.

5. The analysis is unclear. How was Cronbach's alpha calculated? was this for each country, or globally?

→We changed the sentence as follows: (Data Analysis, p.7, L.5-6)
Pre: “Cronbach’s α was used to estimate internal consistency.”
Post: “For estimating the internal consistency, Cronbach’s α was calculated by sub-dimensions in each country.”

6. The standard of English is rather poor which makes the paper hard to read, and there were some sections that I could not follow at all. I would recommend revision by a native English speaker.

→Native English speaker checked our paper.

Other more detailed points are as follows:

METHODS

1. the response rates and details of participating hospitals would fit better in the results section rather than methods. This material may also be better presented in a table to allow comparison between the three countries, rather than in text

We summarized them in Table 1.
2. how were the 14 Japanese hospitals chosen?

Japanese hospitals and US hospitals were not randomly selected, and they voluntarily participated. It was also written in Table 1.

3. the methods section is rather clunkily to read as large sections of text are repeated between the sections for each country. Can this be restructred to avoid this repetition?

We summarized them in Table 1.

4. how were the sample sizes chosen, and why are they so different across the three participating countries?

The sample sizes in Japan and Taiwan were calculated without making allowance for international comparison because those data were originally collected to verify the internal consistency or availability in each country. The sample size in the U.S. was big because it was a part of the national database from 2006 to 2009.

5. it is stated that chi square test was used for categorical variables - what about other variables? For example a p value is presented for a comparison of the patient safety grade among the countries, but these are ordinal and not categorical data.

→Patient Safety Grade is an ordinal data. But we compared the proportion of the respondents who answered “Very Good” or “Excellent”. Therefore, we changed the sentence as follows: (Results, p.8, L.3-6)

Pre: “As regards to Patient Safety Grade, US respondents were more likely to answer “Excellent” or “Very good” (70.8%) than that of Japan (44.6%, P<0.01) and Taiwan (37.7%, P<0.01).”

Post: “As regards to Patient Safety Grade, the proportion of US respondents who answered “Excellent” or “Very good” (70.8%) was significantly higher than that of Japan (44.6%, P<0.01) and Taiwan (37.7%, P<0.01).”

RESULTS

1. Table 3 and table 5 · is mean appropriate for all of the data? I suspect many of these measures are ordinal and therefore would be better represented by the median.

Each question was ordinal data. But we calculated Percent Positive Scores and Total Scores by aggregating questions in each sub-dimension. Percent Positive Scores and Total Scores were quantitative data. The detail is written in “Measures”. (p.6, L13-22)
DISCUSSION

1. It is stated that the Deming cycle should be used - but is this the only suitable approach? Better justification for this is needed.

→ We added a following sentence. (p.9, L8-10)
“For the continuous improvement of safety and quality, the organizations are needed to construct Deming Cycle or to analyze the course of adverse events for preventing the recurrence.[17]”

2. Where data are presented on the number of nurses per bed, it is not clear whether this is based on the present study or from the literature - please clarify.

→ We added a following sentence. (p.10, L23-25)
“According to “OECD Health data 2010” and national statistics of Taiwan, the number of nurses per bed in the U.S. (3.4) was 5 times higher than that of Japan (0.7), and 4 times higher than that of Taiwan (0.8).[23,24]”

3. Better justification is needed for the statement that by varying the number of part time workers, US hospitals might be better able to adapt the number of workers to changes in demand.

→ We added a following sentence. (p.11, L4-7)
“In the US hospitals, temporary nurses who are called “Agency nurse” or “Travel nurse” are positively hired.[25] On the other hand, most of hired nurses in the Japanese and Taiwanese hospitals are permanent staff.”

4. In relation to staff experience with event reporting, this may also be due to differences in the underlying number of events, or different perceptions as to what constitutes a reportable event.

→ It is difficult to say that the number of reports reflects the real number of events. Therefore, we added a following sentence: (p.11, L16-17)
“In addition, the definitions or perceptions of events that has to be reported might be different across the countries.”
Reviewer: Said Bodur

1. The title is not consistent with the content. “A cross-national survey” sounds like the authors collected data from more than one nation, which is not the case:

→The title was changed as follows:
Pre: “A cross-national survey on patient safety culture: Japan, Taiwan and the U.S.”
Post: “The characteristics of patient safety culture in Japan, Taiwan and the U.S.”

a. They used AHRQ data, which can be used in discussion but is wrong to use like original data.

→We wrote it clearly in the Methods. (p.5, L9-15)
“For comparing the PSC, a questionnaire-based, anonymous, and self-administered cross-sectional survey was conducted towards healthcare workers in Japan and Taiwan. The questionnaire of HSOPS was used to measure the PSC in each country. The data of Japan and Taiwan was also compared with “2010 HSOPS Comparative Database” in the U.S. which was provided by AHRQ.”

b. Before Japanese data were collected (January 2010), Taiwanese results were already presented and compared with US (2009). If SF designed the study, how can this design effect the past?

→Dr. Chiu-Chin Huang, who is one of our co-authors, compared Taiwanese data with US data, and reported them in the conference (2009). But the results have not reported as a paper yet, and this is the first paper regarding her survey. On the other hand, Chen et al (2010) also reported the results of comparison between Taiwanese data and US data. But their subjects were limited in physicians, nurses and administrative workers at teaching hospitals, and the number of respondents was only 788. We added following sentences: (Background, P.4, L.19-23)
“Chen et al (2010) compared Taiwanese PSC with US PSC using statistical analysis, showed some characteristics of Taiwanese healthcare workers.[7] But the subjects were from only teaching hospitals, and the sample size seems to be not enough to understand the characteristics of Taiwanese healthcare workers.”

2. The study says Japanese version of PSC was developed and tested by Ito et al. (2011). How is it possible to use that test in 2009?
The data set of Ito et al (2011) is the same as our data set although we added more data. Ito's paper was published in 2011. But the data were collected in 2009.

3. The study includes only acute care hospitals.

The questionnaire of HSOPS was developed for acute care hospitals. Therefore, long-term care hospitals or facilities should be compared by using another questionnaire (Nursing Home Survey on Patient Safety by AHRQ).
Reviewer: Peter Spurgeon

Major Compulsory Revisions

The authors adopt a quasi scientific style of simply reporting scores and unfortunately there are so many across the three countries with a host of “a bit higher here, lower here, higher there, those two higher than the third” etc that overall one is lost as to what is the message about patient safety culture.

→ We arranged the sentences as follows: (p.8, L12-17)
“In Japan, “Organizational Learning – Continuous Improvement” was the lowest rating among 3 countries, and “Hospital Management Support for Patient Safety” was lower than the U.S. In Taiwan, “Frequency of Events Reported” was the lowest rating among 3 countries, and “Communication Openness” was lower than the U.S. In the U.S., “Staffing” was the highest rating among 3 countries.”

The topic needs a higher order interpretation. For example there are many more or less nurses represented in the samples across the countries - does this matter, do nurses have an influence on the observed pattern? There is no comment when it may be key to understanding what the numbers mean.

→ We added a following sentence. (p.11, L13-16)
“The small proportion of nurses in the US respondents might have an effect on the little experience in event reporting because it is widely known that most of event reports come from nurses.[26]"

A large number of hospitals exist in each country sample and clustering of high PSC scoring groups in each country might be very revealing. As the authors conclude “The results suggested that the PSC could vary among different countries as well as hospitals”. So both country and institution could be an influence but I am not clear how, particularly in respect of the latter.

→ In this study, we didn’t verify the effects of the differences of hospitals. Therefore the following words were deleted: (Conclusions, p.13, L5-6)
“The results suggested that the PSC could vary among different countries as well as hospitals.”

Minor Essential Revisions
The English style is stilted and this sometimes makes it difficult to understand exactly. Once the key messages of the data are discerned the paper needs to be revised probably with a native English speaker involved.

Native English speaker checked our paper.

Finally, we thank the reviewers for his/her constructive comments that have helped to improve our manuscript. Looking forward to hearing from you.

Yours truly,