

# Minutes for Data Analysis Group Meeting

**Date:** April 27, 2016

**Time:** 3:00 PM CDT

**Means:** Conference Call

**Attendees:** Lealem Mulugeta  
Marc Horner  
Martin Steele  
Tina Morrison  
Lu Tian

**Not Available:** All in attendance

## Agenda:

1. Come to consensus on the data reduction method
  - (a) Lu presented a data reduction approach, and Martin had some thoughts as well
  - (b) EMBC abstract results section in relation to the preliminary data we will present
2. Task allocations for our analysis moving forward

## Action Items:

See Notes for details.

- **Lealem:**
  - Post meeting minutes
  - Present analysis findings at the next meeting. See notes for agenda item #2.
- **Marc, Martin and Tina:**
  - Present analysis findings at the next meeting. See notes for agenda item #2.
- **Lu**
  - Reduce new data based on new criteria and send to the team (Done)

## Notes:

### 1. Consensus on the data reduction method

- Following our meeting on February 27<sup>th</sup>, Lu proposed the following criteria for reducing the survey data:
  1. "No more than 2 missing responses for the 35 questions"
  2. "Total discrepancy over three pairs of replicated questions  $\leq 6$ "
  3. "Answer for "5"  $\leq 20$ "

This resulted in a sample size reducing to 162 from 185.

- Given the election of these criteria do have a subjective factor to them, the data analysis group examined these criteria extensively to ensure that they were reasonable. Throughout the discussion, we concluded that:
  1. Criterion #1 was reasonable since a participant who is fully engaged and seeks to provide thoughtful

response will make all efforts to not miss very few questions, if any at all.

2. Criterion #2 was refined to include a secondary rule:
  - The sum of the differences between the scores of the repeat questions should be no more than 6. Although this is somewhat arbitrary, it is reasonable to expect people to respond slightly different to the same question if asked in a slightly different way. Therefore, it was reasonable to expect:
    - (a) Response should generally not exceed  $\pm 2$  points difference between the repeated questions; and
    - (b) If any of the repeated questions exceeds the  $\pm 2$  threshold, only one of the three questions can exceed this limit and must be  $\leq 3$ . (e.g. 2 1 3 = 6 and consistent response; while 0 1 4 = 6 but is considered inconsistent since one of the differences is  $> 3$ )
3. Criterion #3 - Given that the survey guidelines had requested that participants should not respond "5" for more than 7 questions, we did not feel it was appropriate to stray too far from this threshold. However, we recognized the limit of 7 responses was a guideline and not a rule. Additionally, most respondents were not likely to keep a very close track of how many questions they scored as "5". Therefore, we felt it was reasonable to set the data reduction threshold for Criterion #3 to about double (15) instead of 3X of the original survey guideline.

- This resulted in the following revised data reduction criteria:
  1. No more than 2 missing responses for the 35 questions
  2. Total discrepancy over three pairs of replicated questions  $\leq 6$ , but
    - (a) only one of the three response differences for repeated questions can exceed  $\pm 2$  threshold limit and must be  $\leq 3$ . (e.g. 2 1 3 = 6 and consistent response; while 0 1 4 = 6 but is considered inconsistent since one of the differences is  $> 3$ )
  3. Gave a response of "5" for  $\leq 15$  of the 32 rules
- Lu will re-run his analysis based on the new criteria and forward outcomes to the data analysis team
- Results reported in the EMBC abstract will be based on the new criteria.

2. **Data analysis tasks:** Once the new data is available, the following data analysis tasks were assigned for the next meeting. However, everyone is not restricted to the assigned tasks. Everyone is encouraged to collaborate in performing their respective tasks and perform other analyses that may provide new insight.

- **Marc :**
  - summarize the demographic breakdown of the respondents based for Continent Region and Work Environment
- **Tina:**
  - Summarize the demographic breakdown of the respondents based for Discipline and Education Level
- **Martin:**
  - Summarize the demographic breakdown of the respondents based for Familiarity with M&S and Motivation to use M&S
- **Lealem:** Perform general analysis of the question responses using Martin's method (counting 4 and 5 responses)

New data is available here: <https://simtk.org/scm/viewvc.php/dat/Survey/?root=cpms&pathrev=308>