

Minutes for Data Analysis Group Meeting

Date: February 3, 2016

Time: 3:00 PM CDT

Means: Conference Call

Attendees: Lealem Mulugeta
Marc Horner
Martin Steele

Not Available: Tina Morrison
Lu Tian

Agenda:

1. Review Martin's analysis
(a) https://simtk.org/websvn/wsvn/cpms/dat/Survey/MS_Analysis_04242015.xlsx
2. Determine primary analysis criteria
3. Determine timeline to complete our analysis

Action Items:

See Notes for details.

- **Lealem:**
 - Post meeting minutes
 - Present analysis findings a the next meeting. See notes for agenda item #2.
- **Marc:**
 - Present analysis findings a the next meeting. See notes for agenda item #2.
- **Martin:**
 - Present analysis findings a the next meeting. See notes for agenda item #2.
 - Provide a better summary of the analysis approach for the rest of the committee to get oriented with the overall analysis methodology
- **Lu**
 - Run analyses to determine if the effect of the repeated questions on the results and if the outcome can be used to help us
 - Provide a summary of the analysis approach for the rest of the committee to get oriented with the overall analysis methodology
- **All**
 - Review Martin's analysis spreadsheet to get familiarized with the analysis Lealem, Marc and Martin will complete for February 24, 2016.

Notes:

1. Martin's analysis spreadsheet.

- Martin walked us through his analysis methodology.
 - Confirmed the bar charts represented the total number of survey participants who ranked each

question with extremely important and very important, and not a summation of the values assigned to each importance level (e.g. Extremely important = 5)

- It is important to note that each question was not answered by all participants. The influence of the missing responses were not taken into account for the overall assessment. This includes for cases where all responses to questions were blank either because the participants chose “Disagree” to take the survey, or chose “Agree” but did not fill out the survey. However, it there are recommendations to eliminate all data where the participant:
 - Chose to “**Disagree**” to take the survey,
 - Chose to “**Agree**” but didn't provide any responses, or
 - Indicated that they were, “**Not using nor interested**” in M&S.

refer to the “complete-mod” and “Anomalous Data” tabs in MS_Analysis_04242015.xlsx spreadsheet for more information.

- We will also need to figure out how to deal with the responses that did not fill out the survey in full. Especially for those who skipped many questions. For example, it is not clear if we should:
 - Completely drop their responses?
 - Include their response but normalize the analysis for each question so that the non-responses does not affect the overall outcome?
 - Use the repeat questions as a criteria to assess whether or not the participants in question were good quality participants?

Perhaps Lu's analysis can help us determine how to tackle this problem?

- It is also possible that some of the questions were blank because they did not understand the question, or other factor that made them skip the question (e.g. they may have missed it by accident)
- In the discussions, the question came up on if and how we should weight the responses from the different participants. For example, a response from a person who is “Not using or interested” in M&S may might not be given as great emphasis as someone who has 5+ years of experience in the field. Marc pointed out that the level of interest and expertise in M&S also has a bias factor from each respondent. So a more objective way to look at the data is to simply present the distribution of the respondents with respect to their level of experience and interest in M&S. There was a general consensus that Marc's approach is probably more appropriate since attempting to provide a weighting criteria based on level of experience also adds a layer of bias from the perspective of the Committee.
- It seems that the data used by Martin and Lu used may not be the same. We will need to reconcile the inclusion/exclusion criteria so that we use the same data set to analyze the data with the two different analysis methods.
- Once we reconcile the inclusion/exclusion criteria, and we have the overall analysis criteria solidified, we will be able to compare and contrast between the two analysis methodologies in order to leverage the best aspects of the two approaches in our final analyses.
- Given that there is a much larger community of M&S experts around the world, there was some concern expressed that:
 - Do we have enough data to represent this wider community?
 - If we could get more data, how would we go about getting more data without having some people repeat the survey?

After some discussion, based on feedback we received from Lu in the past about the statistical significance of the 180+ responses we revived, it was concluded that there is enough data to run analysis for the first publication of the Credible Practice Guideline. It is best to work with what we have now, and re-launch the survey again in the future based on the needs of the stakeholder community.

- Given the perspectives of the M&S community in healthcare are likely to evolve over time, we will

re-engage the global stakeholder community periodically to update the guideline to reflect the evolutionary state of the field.

- 2. Analysis criteria.** For the next meeting, the attendees agreed to perform and present the following analyses (see complete-mod tab of spreadsheet):

Look at the number of responses for Extremely important and Very important, as well as overall response distribution based on the following criteria.

- **Marc:**
 - **Columns BA and BB:** *How familiar are you with computational modeling and simulation / (M&S)?*
 - **Column AZ:** *What is the primary reason for your use/interest in leveraging / modeling and simulation (M&S) for healthcare research and / practice?*
- **Martin:**
 - **Columns AW and AX:** *What is your primary field of academic/professional training?*
 - **Column AY:** *What is your highest level of education?*
- **Lealem:**
 - **Columns AU:** *What is your geographical location?*
 - **Column AV:** *What is the primary setting you work in?*

- 3. Data analysis timeline.** An overarching timeline for the full analysis has not been set. However, the meeting attendees will provide their findings at the February 24, 2016 meeting as agreed above.