Wednesday, February 19, 2020

10:00-11:00 AM

Scholar Seminar

"Put your suit on, we’re getting in the pool! Approaches to pooling datasets"

Briana Mezuk, PhD – Analysis Core Co-Leader

The suggested readings, PowerPoint slides as well as the video recording of this seminar will be sent out to everyone. (Do not have an email address on record.)

Attending (see screenshot of list on last page):

1. (participant with Chinese name/writing)
2. 248-879-2731
3. 561-654-8256
4. Angie Zell
5. Carolyn Still
6. Danyelle Dawson
7. Dayna Johnson
8. DeAnnah Byrd
9. Guest 3
10. Jason Winston
11. Julie Ober Allen
12. Katrina Ellis
13. Kaydian (S. Reid)
14. M. Dottingham (Dot) Fullwood
15. Mamadou Seck
16. Ola Rostant
17. Patty Keys
18. Preethy Samuel
19. Riana Anderson
20. Rie Suzuki
21. Rodlescia Sneed
22. Shekinah Fashaw
23. Sheria Robinson
24. Teri Rosales
25. Verna Keith
26. **Wassim Tarraf, MCUAAAR Analysis Core Co-Leader**

*Additionally*, recording and materials sent upon request to…

1. Hawi Teizazu
2. Maria Roche-Dean
3. Janelle Goodwill
4. Patsy Smith
5. …

In response to DeAnnah Byrd’s question - Difference between “pooling” and “harmonization”? (47:26):

Briana – In this context, harmonizing looks at “what are the comparable measures in terms of the actual items and the responses to those items, i.e. Likert vs. Yes/No, and the approach to analyzing those data”.

Wassim - “Harmonization, relative to pooling, is finding similar measures across different data sets even if the response categories are different. It is a first step to help to understanding correlation structures within and across datasets in order to develop prediction models or some ‘matching scheme’ to impute patterns that don’t exist.” NIH is very interested in this because they have invested a lot of money in these datasets – harmonizing and pooling across the datasets to make sure that these resources are conducive to better, more generalizable and reproducible results across datasets.

Ola Rosant – Commenting on “harmonization: [Scott Hofer, PhD (University of Victoria)](https://www.uvic.ca/socialsciences/psychology/people/faculty-directory/hoferscott.php) – links data on one master scale by measurement. Prediction model=integrative data analysis – Once the measures are linked a mega dataset is created and used as a prediction model.

Recommended Reading:

“Use of a Pooled Cohort to Impute Cardiovascular Disease Risk Factors Across the Adult Life Course”, Adina Zeki Al Hazzour et.al. PMCID: PMC6659365

Additional readings on “harmonization” will also be provided.

