ISCTM Working Group session, 16th Annual Scientific Meeting, 21 February 2020, Washington DC

While previous working groups focused on defining what an RAAD was and on how to best characterize response (in terms of the most appropriate scales); this working group session focused on the following three questions:

Question 1:
None of the commonly used depression rating scales were designed to measure rapid changes in symptoms. Measurement of symptoms that change in timeframes for which traditional scales were not developed for is a challenge for this class of compounds. While some researchers have developed protocols that use only certain items of scales (e.g., MADRS), there is no formal guideline around selection of items from existing scales. In this working group, the primary goal will be to determine which items from which existing scales will be important to retain and include in a guideline document. Next steps will include asking key stakeholders to contribute data for an analysis that will test the items derived by expert consensus.

Question 2:
The limitations in the measurement of the effects of RAADs are well-established by this group and elsewhere. The next generation of assessments will likely supplement currently used measures with passive data collection. To date, we are unaware of any researchers that are applying passive electronic measurement techniques to assess symptoms. For this working group, the primary goal will be to determine which categories can be accurately measured using these techniques and best methods to leverage these techniques to assess over time. This group will not be concerned with identifying any specific product but rather general categories for future evaluation.

Question 3:
Interest in the use of the compounds associated with RAADs has increased over time as researchers and practitioners have assessed the benefits in the treatment of depression. There has also been interest in other transdiagnostic indications including: PTSD, anxiety disorders, suicidality and even psychosis. This working group will be concerned with identifying those scales that would most appropriately measure symptoms across time in association with these conditions and producing a guideline document.

The session format included assignment of group individuals to questions that they might have specific expertise through scientific knowledge, industry and clinical trial experience or from the

regulatory perspective. There were six groups with two independent groups assigned to each of the three questions so that consensus could be determined around some aspects while providing space for each group to contribute unique suggestions or guidance.

Question 1 Summary:

There were fairly concrete suggestions about the existing measures, specifically that any assessment after the 24-hour point is probably fine to use, but that anything within 24 hours may be primarily descriptive and indicative of distress. To this end, the idea of measurement within the 24-hour period might include Patient Reported Outcomes PRO (or electronic PROs that could be used by the patient outside of the clinic) or Ecological Momentary Assessment (EMA) that could likewise provide more descriptive data within this time period. There were general suggestions about using unidimensional scales, or distress-related scales like the HAMD6, MADRS-6 or the VAS.

According to Dr. Holly Lisanby (NIMH) the FDA has accepted letters of intent from the Critical Path Institute regarding "Patient- Reported Outcome (PRO) Consortium, Depression Working Group to qualify two new measures for rapidly acting antidepressant action" that include the following:

1) Momentary Assessment - Symptoms of Major Depressive Disorder Momentary Assessment (SMDDMA).
2) 24hr assessment - Symptoms of Major Depressive Disorder Diary (SMDDD)

Using ecological momentary assessment alone or in conjunction with patient diaries was agreed to be an effective way of capturing mood changes with the caveat that whatever technology enabled this must be unobtrusive and easy to use.

Additionally, the Immediate Mood Scaler (IMS) may be an option for this population to assess rapid acting anti-depressant activity. The scale was designed to assess more discrete (momentary) mood states related to anxiety and depression. The scale was developed by a DARPA subnet group interested in measuring limbic activity. A validation study was carried out by Nahum et al., 2017 wherein the IMS was administered alongside the PHQ-9, GAD-7 and Ruminative Response Scale. It was noted that: "Although IMS-12 shows good correlation with standardized scales, it further captures mood fluctuations better and significantly adds to the prediction of the scales". This would appear to make the IMS a good candidate for measuring rapid changes in mood state.

Other topics included the real-time measurement of possible relevant biomarkers such as hemoglobin A1C that may augment or support putative changes in clinical measures like the IMS. Proxy determinations of change were also discussed in terms of measuring neurocognitive domains such as reward sensitivity function for which there are tests available.

Question 1 Action:

The intention of the sub-working group will be to develop a consensus statement around the best practices and additional data gathering modalities that can be used to assess rapid change in mood symptoms.

Question 2 Summary:
In response to question two there was some consensus around the use of passive data gathering technologies or tasks that could assess "feature spaces", i.e., those elements that cannot be feasibly assessed by a clinician's questioning. Suggested possible methods included:

- keystroke analysis
- passive analysis
- eye gaze of the environment

While these methods have been used to study depressive symptoms, there are no current studies using these technologies to assess rapid change in depressive symptoms. Keystroke analysis is fairly sophisticated and has reasonable diagnostic validity (e.g, Mastoras et al., 2019 and Zuluteta et al., 2018) the technique does not appear to have been trialed with respect to depression severity or change over time.

Passive analysis using smartphone data (e.g., GPS locational data and accelerometry) similarly has found some utility in the detection of depression, there does not seem to be evidence around use in change of depressive symptoms over time.

Eye gaze analysis that can include standardized tasks that have the advantage of norm-referenced populations and repeatability do appear to have some utility in assessing depressive symptomology. That said, Li et al., 2016 found that abnormalities in fixation, saccadic length and other measurable features were strongly correlated with depression. It remains unclear however if, given the mechanism of action of many rapid acting compounds whether this would confound such results.

Natural language processing (NLP) has also found utility in the diagnosis and characterization of depression though it is unclear if this, as a method, could be used in the assessment of rapid change. There was a general question of whether data gathered in these manners would be too variable to characterize a response, and whether it would be possible to have a challenge protocol in order to measure the response to a particular stimulus. Alghowinem et al., 2018 suggest a multimodal approach, which for the purpose of detection of change may prove to be the most salient methodology to proceed with.

Question 2 Action: Develop a challenge protocol that would use a multimodal approach to passive data acquisition to determine whether this method would be appropriate.

Question 3 Summary:
Question 3 focused in transdiagnostic indications and how best to measure symptom change as a result of RAADs. The groups discussing this question also indicated a strong tendency towards the use of passive measurement over more traditional rater administered clinical assessments. There was agreement that across indications there should be a focus on domains of disorders and what methods are currently validated to assess these. Although there were some concerns over timeframe and frequency, there is some evidence that these considerations have been taken into account for at least some passive measurement modalities (e.g., Cohen et al., 2020 in press; Ying et al., 2020 in press).

While AI and Machine Learning were recognized as having a future role in the analysis, interpretation and categorization of disorders, there was some concern about the nature of AI and, if allowed to "learn" over the course of a clinical trial if the results could be meaningfully interpreted.

Question 3 Action:
The action for this sub-working group will be to develop guidelines around which domains are appropriate targets for passive data collection depending on the indication.

References:

https://www.fda.gov/drugs/drug-development-tool-ddt-qualification-programs/ddt-coa-000109-symptoms-major-depressive-disorder-momentary-assessment-smddma

https://www.fda.gov/drugs/drug-development-tool-ddt-qualification-programs/ddt-coa-000108-symptoms-major-depressive-disorder-diary-smddd

https://www.fda.gov/drugs/drug-development-tool-ddt-qualification-programs/clinical-outcome-assessment-coa-qualification-program Symptoms of Major Depressive Disorder Scale (SMDDS)

Alghowinem S, Goecke R, Wagner M, Epps J, Hyett M, Parker G, Breakspearet M. Multimodal Depression Detection: Fusion Analysis of Paralinguistic, Head Pose and Eye Gaze Behavior. IEEE Transactions on Affective Computing. 2018; (9)4: 478-490. doi: 10.1109/TAFFC.2016.2634527.

Cohen A, Cowan T, Le T, Schwartz E, Kirkpatrick B, Raugh I, Chapman H, Strauss G. (in press). Ambulatory digital phenotyping of blunted affect and alogia using objective facial and vocal analysis: Proof of concept. Schizophrenia Research

Narziev N, Goh H, Toshnazarov K, Lee SA, Chung KM, Noh Y. STDD: Short-Term Depression Detection with Passive Sensing. Sensors (Basel). 2020;20(5):1396. Published 2020 Mar 4. doi:10.3390/s20051396

Sarda A, Munuswamy S, Sarda S, Subramanian V. Using Passive Smartphone Sensing for Improved Risk Stratification of Patients With Depression and Diabetes: Cross-Sectional Observational Study. JMIR Mhealth Uhealth. 2019;7(1):e11041. Published 2019 Jan 29. doi:10.2196/11041

Ying-min Zou Y, Ni K, Wang Y, Yu E, Lu, S, Zhou F, Yang H, Cohen A, Strauss , Cheung E, Chan R. (in press). Effort-Cost Computation in a Transdiagnostic Psychiatric Sample: Differences Among Patients with Schizophrenia, Bipolar Disorder, and Major Depressive Disorder. PsyCh Journal.

Zulueta J, Piscitello A, Rasic M, Easter R, Babu P, Langenecker SA, McInnis M, Ajilore O, Nelson PC, Ryan K, Leow A. Predicting mood disturbance severity with mobile phone keystroke metadata: a BiAffect digital phenotyping study. J Med Internet Res. 2018; 20(7):e241. doi: 10.2196/jmir.9775.