verily

Performance Requirements for Novel Digital Measurements

March 24, 2020
Digital Health Technologies in Drug Development
Session III - Digital Tools for Pivotal Trials

Ritu Kapur, Ph.D.
Head of Digital Biomarkers
Verily Life Sciences
ritukapur@verily.com


OUR MISSION
Our mission is to make the world's health information useful so that people enjoy healthier lives.
verily

Signal Verification

Is the sensor working?

Requires bench testing

Analytical Validation

Am I measuring what I think I am measuring?

Measured against corroborative devices, human raters

Clinical Validation

Does what I am measuring matter?

Requires datasets in which the clinical outcome of interest varies

Step Count

Signal Verification	Bench testing of accelerometer
Analytical Validation	Human observer, Corroborative device
Clinical Validation	Parkinson's vs. non-PD By severity of PD

Step Count

Signal Verification	Bench testing of accelerometer
Analytical Validation	Human observer, Corroborative device
Clinical Validation	Parkinson's vs. non-PD By severity of PD

Challenge

Collecting **naturalistic** data in sufficient quantities

Bradykinesia

Signal Verification	Bench testing of accelerometer
Analytical Validation	Human observer, Corroborative device
Clinical Validation	By severity of PD

Bradykinesia

Signal Verification	Bench testing of accelerometer
Analytical Validation	Human observer, Corroborative device
Clinical Validation	By severity of PD

Challenge

Human observer ratings are subjective, effectively establishing a ceiling for accuracy

To date, accuracy of corroborative devices is pinned to subjective human observation

Bradykinesia

Signal Verification	Bench testing of accelerometer
Analytical Validation	Human observer, Corroborative device
Clinical Validation	By severity of PD

Challenge

Lack of common methods for evaluating digital metrics against clinical ratings

- kappa, ICC
- consensus scoring

Geo-Social Quality of Life

Signal Verification	?
Analytical Validation	?
Clinical Validation	By severity of PD, QOL & mood instruments

Geo-Social Quality of Life

Signal Verification	?
Analytical Validation	?
Clinical Validation	By severity of PD, QOL & mood instruments

Challenge

Engagement is higher for "bring your own smartphone", which means 100's of combinations of hardware x versions of OS

Signal verification becomes prohibitively difficult, and who sets standards for performance?

Geo-Social Quality of Life

Signal Verification	?
Analytical Validation	?
Clinical Validation	By severity of PD, QOL & mood instruments

Challenge

Same as for bradykinesia - lack of common methods for evaluating digital metrics against clinical ratings

Questions to Consider

- **To prove you're as good as a (subjective) clinical rating:**

  How should we go about creating a common methodology for evaluating digital metrics against clinical ratings?

- **To move from a subjective to an objective measurement:**

  How do we create a common approach to establishing performance criteria for quantitative measurements that are not limited by subjective ratings?

- **To build completely new measurements:**

  Is there a world in which algo performance on clinical datasets suffices if

  - Training and testing datasets are separated appropriately
  - Analyses on multiple **independent** datasets and show similar results