Considerations for the Generalizability of Clinical Outcome Assessment Validation Data Across Rare Diseases

Chris J Edgar[1], Pamela Ventola[2]
1 Cogstate, London, UK; 2 Cogstate; Yale Child Study Center, New Haven, US

Introduction

- Development, validation and selection of clinical outcome assessments (COAs) requires:
  - ✓ A clear understanding of concepts of interest (COIs) for measurement and intended context of use (COU)
  - ✓ Evidence for content validity
  - ✓ Evidence for psychometric validity and reliability
- The ability to reuse or adapt existing COAs for novel COUs is valuable in reducing cost and timelines
- In rare disease, availability of patients and issues of disease course and severity may make adaptation an imperative

Evaluation of Existing Clinical Outcome Assessments

Properties	Highly generalizable COAs	Specific COAs
Concept(s) of interest (COIs) for measurement:	COI(s) common to many diseases/conditions e.g. 'cognition' or 'ambulation'	COIs disease specific e.g. 'severity of mitochondrial disease'
Evidence for validity and reliability:	Established in several diseases/conditions	Established in a single disease/condition
Intended context(s) of use:	Intended for use across a broad range of ages, intellectual ability	Developed for a narrow age range, specific stage of disease, narrow range of intellectual ability

- Per updated FDA guidance, where a primary endpoint may be subject to bias or rater judgment, supportive secondary endpoints such as performance-based (PerfO) outcome assessments (e.g., cognitive and ambulation tests), may complement reports from caregivers and patients regarding relevant aspects of day-to-day functioning
- Such tests may have the advantage of being widely validated and suited to multiple contexts of use (COUs)

Key Elements of Validity and Reliability

Element	Description
Content validity	Evidence COA measures the relevant COI(s) (qualitative and quantitative)
Construct validity	Evidence relationships between items, domains, COAs align with a priori hypotheses
Reliability	Stability of scores over time when no change is expected; consistency between raters
Ability to detect change/sensitivity	Evidence COA can identify differences in scores over time in individuals/groups who have changed
Interpretation/meaning of score changes	Relationships between COA and other endpoints used to measure treatment benefit

- Where tests are highly standardized/fully automated, inter-rater reliability may not be relevant
- Reliability increasingly important with smaller sample sizes
- Test-retest (cf intra-rater) reliability critical, impacts sample size calculation, and easier to measure/confirm in typical trial designs than inter-rater
- Selection and application of widely used performance-based outcome (PerfO) assessments (e.g., cognitive tests) can complement reports of day-to-day functioning

Generation of Additional Validation Evidence

Application of adaptive seamless designs:

- It is common for exploratory evidence from earlier phases to inform later trial design, including endpoints
- FDA note that "adaptive seamless trial designs may allow early evidence to be used later in a study, especially helpful when there are limited numbers of patients"
- Exploration of multiple COAs early on in a trial, perhaps as part of a run-in period, can allow for specification of a reduced set of endpoints based on a priori criteria

Adaptation of existing COAs:

- Several COA paradigms may be more easily adapted to novel COAs
- Example: Clinical Global Impression Severity and Improvement (CGI) can be modified with disease specific anchor text
  - Tailored to symptom presentation and functioning level of population
  - Increases both specificity and sensitivity of endpoints, which is highly relevant to rare disease studies with smaller numbers of patients
- The use of an established COA can be supplemented by asking respondents to identify a 'most bothersome symptom' to increase personalization
- Consistent with FDA's guidelines related to importance of patient involvement

Conclusions

- When selecting a COA, a comprehensive understanding of the measured concepts, key properties, reliability, and validity is imperative
- In rare disease trials, by definition, there are a limited number of patients, so sensitivity and specificity of the endpoints are of critical importance
- Several approaches may provide needed efficiencies in the context of rare disease clinical trials
  - Investigators may need to employ adaptive designs and explore multiple endpoints early in the trial/program
  - Adapting existing COAs, such as the CGI, to specific indications using clear, objective anchors tailored to the symptom presentation and functioning level of the patient can increase specificity and sensitivity of the measure
  - Adaptive designs and adaptations/personalization of COAs are aligned with FDA's recommendation on involving patients and families in selecting endpoints and determining meaningful changes