MAIMON WORKING PAPER No 30 DECEMBER 2025

RASCH MEASUREMENT AS THE NECESSARY REGULATORY GATEKEEPER FOR PATIENT-REPORTED OUTCOMES

Paul C Langley, Ph.D Adjunct Professor, College of Pharmacy, University of Minnesota, Minneapolis MN

ABSTRACT

Patient-reported outcomes (PROs) have become central to regulatory submissions, health technology assessment (HTA), and pricing decisions. Their widespread acceptance is routinely justified by reference to guidance issued by the U.S. Food and Drug Administration (FDA) and the European Medicines Agency (EMA). This paper argues that such reliance reflects a fundamental category error. Regulatory acceptance is not measurement, and it cannot confer the quantitative legitimacy required for arithmetic operations on PRO scores.

The objective of a PRO, if it is to function as a measure, is to quantify possession of a latent trait. Representational measurement theory establishes that this objective can be met only through conjoint simultaneous measurement yielding an invariant scale. For latent variables, this requirement is operationalized exclusively by the Rasch model, which produces a logit ratio measure with separable person and item parameters. Without Rasch, summed ordinal scores, preference indices, and composite scales do not measure anything at all; they merely score responses.

A critical review of FDA and EMA PRO guidance shows that regulators emphasize content validity, reliability, responsiveness, and interpretability, while remaining silent on scale type, invariance, and lawful arithmetic. This silence has been widely misinterpreted by HTA as implicit measurement approval. As a result, PRO scores lacking measurement properties have been embedded in cost-utility analyses, reference-case models, and pricing decisions, where arithmetic operations are performed without justification.

The paper argues that this situation is no longer defensible. Either PROs are measured, in which case Rasch must be mandatory as a regulatory gatekeeper, or they are not, in which case quantitative claims based on them must cease. The current ambiguity has enabled non-measurement to masquerade as evidence for more than two decades. Restoring scientific credibility requires explicit enforcement of measurement conditions. Without Rasch, there is no measurement, and without measurement, arithmetic has no meaning.

INTRODUCTION

The discussion of FDA and EMA guidance on patient-reported outcomes has too often obscured the central issue rather than clarified it [1] [2] [3], by focusing on what regulators do or do not permit, HTA has mistaken procedural acceptance for epistemic legitimacy. This paper begins from a more fundamental position: if patient-reported outcomes are to be used as *measures* and not merely as descriptive or narrative endpoints then Rasch measurement is not optional [4] It is the entry portal. Without it, there is no measurement, only scoring.

This point reframes the role of FDA and EMA entirely. Regulators do not, and should not, define the conditions for measurement. Their remit is to evaluate whether evidence is fit for regulatory purposes: labeling claims, benefit–risk communication, and consistency of reporting. That remit is legitimate and limited. The problem arises when HTA treats regulatory acceptance as if it answered a prior and logically necessary question: whether a numerical outcome is a measure capable of supporting arithmetic. That question belongs not to regulators, but to measurement theory; specifically, the seminal contribution of Stevens on scales of measurement in 1946 and the formalization by Krantz et al of the axioms of representational measurement in 1971 [5] [6] .

If a PRO is intended to inform arithmetic claims then the rules of representational measurement apply. Latent traits cannot be measured by summing category responses, correlating scores, or demonstrating responsiveness. They can be measured only if a conjoint simultaneous measurement structure exists that yields invariant units and separates person location from item difficulty. This is precisely what the Rasch model provides. No alternative framework satisfies these requirements. Without Rasch, there is no demonstration of unidimensionality, no invariant scale, no meaningful unit, and no basis for claiming that numerical differences correspond to differences in a latent attribute.

From this perspective, FDA–EMA PRO guidance is not the foundation of quantitative legitimacy; it is orthogonal to it. Content validity, reliability, and interpretability are necessary for regulatory communication, but they do not establish measurement. A PRO can be entirely acceptable for labeling and still be mathematically incapable of supporting addition, multiplication, or comparison as a quantity. Treating such scores as measures is not an extension of regulatory guidance; it is a category error.

This reframing resolves a long-standing confusion. The question is not whether FDA or EMA should “require Rasch”; the question is whether HTA wishes to claim that PROs are measured. If the answer is yes, and HTA practice unequivocally assumes yes, then Rasch rules must be enforced regardless of regulatory silence. Guidance that permits arithmetic on PRO scores without requiring Rasch does not merely fall short; it abandons measurement altogether while continuing to speak its language.

The argument of this paper is simple and uncompromising. Patient-reported outcomes may play many roles in clinical research and regulation. But if they are to function as measures, Rasch is the necessary and sufficient gateway. Anything else produces numbers without quantity, arithmetic without meaning, and decisions justified by scores that cannot bear the weight placed

upon them. Measurement is not conferred by acceptance, consensus, or convenience. It begins and can only begin with Rasch.

WHY RASCH IS ESSENTIAL: NO RASCH, NO MEASUREMENT FOR PROs

The objective of a patient-reported outcome, if it is to function as a *measure*, is unambiguous: to quantify the possession of a latent trait. Pain, fatigue, mobility, emotional distress or need fulfillment have to be inferred from patterns of responses to items. The moment an investigator claims that a PRO score represents "more" or "less" of such an attribute, that claim invokes measurement. At that point, the rules of representational measurement theory apply, without exception.

Those rules have been settled for decades. To measure a latent trait, it is necessary to demonstrate unidimensionality, invariance, and a lawful mapping from empirical observations to numbers such that differences in numbers correspond to differences in the attribute. The result must be a scale with a meaningful unit and permissible arithmetic. For latent variables, there is only one framework that satisfies these axioms: conjoint simultaneous measurement, operationalized through the Rasch model. This is not a matter of methodological preference or disciplinary fashion. It is a logical consequence of what measurement requires where the Rasch rules for instrument development are consistent with the axioms of representational measurement [7].

Rasch measurement provides a logit ratio scale in which person location (trait possession) and item difficulty are estimated independently on the same metric. This separability is the defining feature of measurement. Without it, item scores and person scores are confounded, units are unstable, and comparisons are sample-dependent. Summed scores, averages, factor scores, and index values fail precisely because they lack this separability. They are scoring rules, not measurement models. They do not yield invariant units and therefore cannot support claims about magnitude or change.

The insistence on Rasch is therefore not ideological. It follows directly from the axioms of representational measurement. If an attribute is latent, and if numerical claims about that attribute are to be made, then the measurement model must demonstrate that the numerical structure mirrors the empirical structure. Rasch does this by enforcing specific requirements: items must work together to define a single latent variable; response probabilities must follow a lawful monotonic function; and the resulting scale must be invariant across relevant subpopulations. Without these properties, there is no basis for claiming that a numerical difference corresponds to a difference in the attribute.

This is why statements such as "PROs can be measured without Rasch" are not debates; they are errors. There is no alternative model that yields a ratio-scale logit measure of a latent trait while satisfying the axioms of measurement. Classical test theory does not do this. Item response theory models that relax Rasch constraints abandon invariance. Factor analysis describes covariance, not measurement. Responsiveness, reliability, and validity coefficients do not create units. They are ancillary properties of scores, not foundations of measurement.

Once this is understood, the implications for PRO guidance are unavoidable. If FDA and EMA documents permit PRO scores to be treated as quantitative outcomes without requiring Rasch measurement, then those documents are not guidance on measurement. They are guidance on data collection and interpretation. That distinction matters. It means that regulatory acceptance does not, and cannot, confer measurement status. If regulators are unaware of this distinction, the appropriate response is not accommodation but correction. A refresher in measurement theory is not an insult; it is a necessity. After 60 years it is difficult to understand why it is still necessary to make this point.

There is no room for compromise on this point. Either a PRO instrument yields a Rasch logit ratio measure of trait possession, or it does not measure the trait at all. In the latter case, it may still be useful descriptively, clinically, or narratively—but arithmetic must stop. Treating non-Rasch scores as measures is not approximation; it is misrepresentation.

The consequence for HTA and regulatory science is stark. If patient-reported outcomes are to underpin quantitative claims, pricing decisions, or comparative assessments, Rasch must be the entry portal. Without Rasch, there is no PRO measurement. Without measurement, there can be no legitimate arithmetic. And without legitimate arithmetic, claims of quantified patient benefit collapse into numerical storytelling. This is not a matter for debate. It is the settled logic of measurement itself.

RASCH PROs: THE NECESSARY STEPS

If patient-reported outcomes are to function as measures rather than scored descriptions, the pathway is neither discretionary nor flexible. Rasch measurement imposes a sequence of necessary yet sufficient steps that follow directly from the objective of quantifying possession of a latent trait. Skipping any step does not weaken measurement; it eliminates it. What follows is therefore not a recommended workflow but a set of conditions that must be satisfied before numerical claims can be made.

The first step is explicit construct definition. A Rasch PRO begins with a single, clearly articulated latent trait. This is not a thematic domain or a conceptual umbrella but a precise attribute whose possession varies monotonically among persons. Ambiguity at this stage is fatal. If the construct cannot be expressed as “more versus less of the same thing,” measurement is impossible. Multi-attribute notions must be decomposed or abandoned; Rasch does not rescue conceptual vagueness.

Second, instrument item generation must be theory-driven and exhaustive, not opportunistic. Items are not indicators to be correlated; they are locations on a latent continuum. Each item must represent a distinct level of the trait, differing in difficulty or intensity, not in kind. Face validity and stakeholder preference are insufficient. Items are hypotheses about the structure of the latent variable and must be constructed to test that structure.

Third, conjoint simultaneous measurement must be demonstrated. Rasch analysis estimates person location and item difficulty on the same metric and tests whether responses conform to the model’s probabilistic expectations. This step establishes unidimensionality empirically, not

by assumption. Items that misfit do not merely reduce reliability; they invalidate the measurement structure and must be revised or removed. Fit statistics are not diagnostics to be tolerated; they are criteria to be enforced. The fit to the Rasch model is not for debate.

Fourth, invariance must be shown across relevant groups and contexts. Differential item functioning is not a nuisance parameter but a threat to measurement. If items behave differently across populations, languages, disease stages, or time, then the scale lacks a stable unit. Without invariance, comparisons are sample-dependent and arithmetic is meaningless. Demonstrating invariance is therefore a prerequisite for any claim of generalizability.

Fifth, the scale must be reported in logit units, not raw scores. The logit is the ratio-scale unit produced by Rasch measurement, expressing the logarithm of the odds of endorsing an item given person location. Raw scores are ordinal summaries; logits are measures. Reporting in logits makes explicit that the outcome is a measurement result, not a scoring convention, and it defines the permissible arithmetic.

Sixth, interpretation must respect the measurement structure. Change scores, group differences, and longitudinal analyses must be conducted on the Rasch scale, not on raw totals or transformed indices. Claims about improvement or decline refer to movement along a measured continuum, not to changes in summed categories. Where the scale does not support a claim, the claim must not be made.

Finally, governance must enforce these rules. Rasch measurement cannot be optional, advisory, or retrospective. It must be embedded at the point where PROs are introduced as quantitative endpoints. Guidance that allows arithmetic first and measurement later is incoherent. If a PRO is intended to support quantitative inference, Rasch must be present from inception through reporting. Measurement must always precede arithmetic.

These steps are demanding, but they are not negotiable. They reflect what measurement of latent traits requires, not what institutions prefer. PROs that satisfy these conditions yield invariant, interpretable, and lawful measures of trait possession. PROs that do not may still be useful descriptively, but they are not measures. The distinction is decisive. Rasch is not an enhancement to PRO development; it is the gateway. Without it, there is no measurement to justify the numbers that follow.

WHERE THE FDA AND EMA FAILED

The failure of FDA and EMA guidance on patient-reported outcomes is not subtle, technical, or excusable by regulatory remit. It is foundational. For more than two decades, regulatory guidance has treated PROs as if their numerical outputs could function as quantitative evidence, while never once specifying the conditions under which such quantification is possible. The result has been the institutional legitimization of what can only be described as epistemic nonsense: summed ordinal scores presented, interpreted, and propagated as if they were measures.

The core failure is simple. Neither FDA nor EMA guidance embeds any requirement for measurement. There is no insistence on unidimensionality as a necessary condition, no requirement for invariant units, no specification of scale type, and no recognition that latent traits cannot be measured without a conjoint simultaneous measurement model. Rasch measurement, available since 1960 and explicitly designed to solve the problem of measuring latent variables, is entirely absent. This omission is not a matter of emphasis or prioritization; it is a categorical error. Guidance that speaks the language of quantitative outcomes while omitting the conditions that make quantification possible is not neutral. It is not just misleading; it is a catastrophic failure.

By focusing almost exclusively on content validity, reliability, responsiveness, and interpretability, FDA and EMA conflated *usefulness* with *measurement*. An instrument may be useful for eliciting patient narratives or for structuring clinical discussion without being capable of measurement. Reliability does not create units. Responsiveness does not establish equal intervals. Correlation does not confer quantity. Yet regulatory guidance repeatedly allowed these properties to stand in for measurement, thereby encouraging sponsors, journals, and HTA bodies to believe that PRO scores were numerically meaningful.

The consequence was predictable. PRO instruments based on Likert-type items were developed, validated, licensed, translated, and deployed at scale. Their scores were summed, averaged, compared, and analyzed as if differences had magnitude. None did. Ordinal category counts were treated as distances. Changes in score were interpreted as changes in an attribute. This is not measurement error; it is measurement absence. The numbers produced do not represent quantities at all.

What makes this failure inexcusable is that Rasch measurement has been available, mature, and widely applied for more than sixty years. Its purpose is precisely to address the problem regulators claimed to care about: how to transform subjective responses into invariant measures of latent traits. That regulators failed even to acknowledge this framework cannot be defended by appeal to remit. Once guidance permits arithmetic interpretation of PRO scores, it has crossed into measurement territory whether it acknowledges it or not. At that point, ignorance of measurement theory is not neutrality; it is negligence.

The failure is compounded by the absence of any requirement that PRO instruments demonstrate invariance across target patient populations. Regulatory guidance repeatedly emphasizes use in the “intended population,” yet provides no measurement framework to ensure that scores mean the same thing across subgroups, cultures, disease stages, or time. Without Rasch-based tests of differential item functioning, claims of generalizability are empty. A score that does not function equivalently across patients is not merely biased; it is uninterpretable as a measure.

For more than twenty years, this regulatory silence has had consequences. PROs have been incorporated into labeling claims, comparative studies, and HTA submissions as if they quantified patient benefit. They did not. They described responses. The difference is decisive. By failing to draw this line, FDA and EMA guidance enabled a global ecosystem in which summed ordinal scores were mistaken for evidence, and arithmetic was performed on non-existent

quantities. All the FDA and EMA achieved were to put numerical storytelling in the box seat for therapy response claims.

Calling this epistemic nonsense is not rhetorical excess. It is a precise description of what happens when numbers are endowed with meaning they do not possess. The fault lies not with the idea of patient-reported outcomes, but with regulators who allowed those outcomes to masquerade as measures without ever requiring the conditions that measurement demands. If PROs are to be used quantitatively, and regulators have allowed and encouraged this, then Rasch must be mandatory. Without it, PRO guidance does not merely fall short. It is deception. The corrective is obvious and overdue. Measurement must be explicit, enforced, and non-negotiable. Until that happens, FDA and EMA guidance on PROs will continue to authorize the appearance of quantification while delivering none of its substance.

IMPLICATIONS FOR FDA AND EMA

The implications of this analysis for FDA and EMA are uncomfortable but unavoidable. While regulators did not set out to define measurement theory, they nevertheless created an evidentiary environment in which numerical PRO scores were treated as if they possessed quantitative meaning. Once guidance documents framed PROs as acceptable endpoints for regulatory claims without explicitly distinguishing description from measurement, they opened the door to systematic deception and misinterpretation downstream. That misinterpretation was not marginal. It became global practice.

The core implication is that FDA and EMA cannot plausibly maintain methodological neutrality once PRO scores are allowed to function as numerical evidence. The moment guidance permits numerical comparison, change estimation, or effect interpretation, it implicitly invites arithmetic reasoning. At that point, silence on measurement conditions is no longer benign. It becomes enabling. Regulators may not have intended to authorize quantification, but their guidance has been repeatedly read as doing exactly that. The persistence of this reading over two decades indicates that the guidance failed to draw a critical boundary clearly enough.

This failure is magnified by the historical context. Rasch measurement has been available since 1960 and has been explicitly developed to address the precise problem regulators face with PROs: how to transform subjective responses into invariant measures of latent traits. That this framework is entirely absent from FDA and EMA guidance is not a trivial omission. It means that regulators endorsed the *use* of PRO scores without ever specifying what must be true for those scores to be treated as quantities. In effect, guidance documents spoke as if measurement were optional, when in fact it is logically prior. This is unacceptable.

The result is that regulators in their ignorance legitimized summed ordinal scores as evidence. Sponsors, journals, and HTA bodies unreasonably inferred that if regulators accepted PRO endpoints, then numerical operations on those endpoints must be permissible; endorsing a further level of legitimized ignorance. Guidance that emphasizes reliability, responsiveness, and interpretability, while remaining silent on scale type and invariance, encourages the belief that psychometric adequacy is sufficient for quantification. It is not.

There is also a population-level implication that regulators have not confronted. FDA and EMA routinely emphasize that PRO instruments must be appropriate for the target patient population, yet they provide no framework for ensuring that scores function equivalently across subgroups. Without Rasch-based tests of differential item functioning, claims of comparability across age, sex, culture, disease severity, or language are unfounded. Regulatory acceptance without invariance testing therefore risks endorsing instruments whose scores change meaning across patients; an outcome directly at odds with regulatory intent.

If FDA and EMA wish to continue endorsing PROs as numerical evidence, their guidance must change fundamentally. Measurement cannot be implied; it must be required. Guidance would need to specify that any PRO intended to support quantitative claims must demonstrate unidimensionality, invariance, and lawful scaling through Rasch measurement. Item selection, scale construction, and reporting would have to be governed by measurement rules, not by convention or convenience. Without these requirements, regulators are culpable in endorsing the appearance of quantification while withholding its substance.

Alternatively, regulators could choose a different path: explicitly restrict PROs to descriptive and communicative roles, making clear that regulatory acceptance does not imply quantitative measurement. This would align guidance with its actual content but would require regulators to resist downstream misuse aggressively. Silence is no longer defensible, because its consequences are now evident. The implication for FDA and EMA is therefore stark. Either embed Rasch measurement as a non-negotiable requirement for quantitative PRO use, or explicitly disclaim any measurement status for PRO scores. What cannot continue is the present ambiguity, which has allowed non-measurement to masquerade as evidence for more than twenty years. Regulatory authority carries epistemic responsibility. Clarifying this boundary is no longer optional.

IMPLICATIONS FOR HTA

The regulatory failure to specify measurement conditions would have been contained had HTA treated FDA–EMA acceptance of PRO instruments for what it was: authorization for use in defined regulatory contexts, not certification of measurement. Instead, HTA made a decisive and consequential inference. Regulatory acceptance was taken as implicit measurement approval. From that point onward, PRO scores were treated as quantitative effects, eligible for arithmetic, aggregation, and incorporation into cost-effectiveness models. This inference was never justified, yet it became foundational. Instrument directories were developed which were also oblivious to the Rasch rules.

HTA's reliance on regulatory approval as a proxy for measurement competence reveals a deeper ignorance: the absence of Rasch measurement from HTA's conceptual apparatus. If HTA had possessed even a rudimentary understanding of representational measurement theory, the inference would have been impossible. Regulatory silence on scale properties would have been recognized as a warning, not a license. Instead, HTA proceeded as if measurement were guaranteed upstream, thereby absolving itself of responsibility to establish it downstream.

This abdication explains why HTA methods documents, reference cases, and journal standards never require demonstration of unidimensionality, invariance, or lawful transformation for PRO-based endpoints. Utilities derived from ordinal responses were accepted as cardinal. Summed Likert scores were treated as effects. Changes in composite indices were interpreted as magnitudes. None of these practices can be defended without Rasch. Yet Rasch is nowhere required, rarely mentioned, and often misunderstood when it appears. The field behaved as if measurement were optional, even as it insisted on arithmetic outputs.

The consequences for HTA are profound. Cost-utility analysis presupposes ratio-scale measurement of outcomes. ICERs presuppose commensurable units in numerator and denominator. Lifetime modeling presupposes additive and multiplicative properties that only ratio measures possess. When PRO-derived scores lacking these properties are admitted as effects, the entire edifice becomes incoherent. This is not a matter of bias or uncertainty; it is a category error. Arithmetic performed on non-measures is undefined, regardless of how carefully it is executed.

HTA's defense has typically been procedural: adherence to reference cases, journal acceptance, and regulatory alignment. None of these defenses addresses measurement. They substitute consensus for validity. The fact that FDA and EMA accepted PRO instruments for labeling claims does not rescue HTA's arithmetic; it indicts HTA's assumption. Regulators never authorized what HTA assumed, and HTA never demonstrated what arithmetic requires. The gap was filled by habit.

This also explains the uniformity of HTA practice across jurisdictions. The same endorsement patterns recur because the same assumption recurs: if regulators allow PROs, they must be measurable. Once that assumption is embedded, Rasch becomes unnecessary by definition. To require Rasch would be to admit that decades of HTA outputs lack quantitative meaning. The field therefore treats Rasch as optional, marginal, or overly technical; anything but necessary.

That position is untenable. If HTA wishes to claim that PROs measure patient benefit, Rasch is non-negotiable. If HTA cannot or will not require Rasch, then PROs must be treated as descriptive inputs, not quantitative endpoints. What cannot be justified is the current middle ground: non-measurement presented as measurement, and arithmetic performed as if legitimacy were inherited rather than earned.

The implication is not incremental reform but categorical choice. HTA must decide whether it is a quantitative science or narrative storytelling. HTA's mistake was not trusting regulators; it was mistaking regulatory acceptance for measurement approval. That mistake cannot be repeated now that the distinction is explicit. Continuing to justify non-measurement by appeal to precedent is no longer error; it is willful disregard. Measurement is not optional, and HTA cannot claim exemption from its requirements.

CONCLUSION

The argument developed here admits of no ambiguity. Patient-reported outcomes cannot function as measures unless they satisfy the axioms of representational measurement. For latent traits,

those axioms are operationalized only through Rasch measurement. Without Rasch, there is no invariant unit, no lawful scale, and no defensible arithmetic. Scores may exist, correlations may be calculated, and narratives may be constructed, but measurement does not occur. Any claim to the contrary is simply false. This invalidates hundreds of PRO instruments in HTA

FDA and EMA guidance did not resolve this problem because it was never designed to. Regulators focused on procedural acceptability, patient relevance, and evidentiary support for labeling claims. They neither claimed nor attempted to establish measurement validity. The concept was entirely foreign. The scientific revolution need never have occurred. That limitation was appropriate to their remit. The failure arose when HTA treated regulatory acceptance as if it answered a prior question that regulators never addressed: whether PRO scores are quantities. They are not, unless Rasch rules are satisfied.

HTA compounded this error by embedding PRO scores into cost-utility analyses, reference-case models, and pricing decisions without ever requiring measurement demonstration. Ordinal responses were summed, utilities were multiplied by time, and composite indices were treated as if they possessed ratio properties. None did. This was not approximation or uncertainty; it was non-measurement masquerading as quantification. The persistence of this practice reflects not disagreement but culpable ignorance of Rasch and the axioms it enforces.

The implications are stark. HTA must choose between enforcing measurement or abandoning arithmetic claims based on latent traits. There is no defensible middle ground. PROs that are not Rasch-measured may still inform clinical understanding and regulatory communication, but they cannot support quantitative comparison, aggregation, or economic modeling. Continuing to pretend otherwise is not pragmatism; it is epistemic failure.

The corrective is clear and unavoidable. If HTA insists on quantified patient benefit, Rasch must be mandatory. If Rasch is absent, arithmetic must stop. Measurement is not conferred by acceptance, repetition, or authority. It is earned by satisfying necessary conditions. Until HTA accepts this, it will continue to generate numbers that look scientific while lacking the properties required to mean anything at all.

ACKNOWLEDGEMENT

This work was prepared with the assistance of ChatGPT, an artificial intelligence language model developed by OpenAI. ChatGPT was used to support drafting, editing, and refining the content. The author takes full responsibility for the final content and any interpretations presented.

A NEW START IN MEASUREMENT FOR HEALTH TECHNOLOGY ASSESSMENT

For readers who are looking for an introduction to measurement that meets the required standards, Maimon Research has just released two distance education programs. These are:

- Program 1: Numerical Storytelling – Systematic Measurement Failure in HTA.
- Program 2: A New Start in Measurement for HTA, with recommendations for protocol-supported claims for specific objective measures as well as latent constructs and manifested traits.

Each program consists of five modules (approx. 5,500 words each), with extensive questions and answers. Each program is priced at US$65.00. Invitations to participate in these programs will be distributed in the first instance to 8,700 HTA professionals in 40 countries.

More detail on program content and access, including registration and on-line payment, is provided with this link: https://maimonresearch.com/distance-education-programs/

REFERENCES

[1] Guidance for Industry: Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims. U.S. Department of Health and Human Services, Food and Drug Administration. December 9, 2009. Guidance for Industry. Silver Spring, MD: FDA.

[2] European Medicines Agency. *Appendix 2 to the Guideline on the Evaluation of Anticancer Medicinal Products in Man: The Use of Patient-Reported Outcome (PRO) Measures in Oncology Studies.* EMA/CHMP/292464/2014, April 22, 2016

[3] European Medicines Agency. *Reflection Paper: Regulatory Guidance on the Use of Health-Related Quality of Life (HRQL) Measures in the Evaluation of Medicinal Products.* EMA

[4] Rasch G, Probabilistic Models for some Intelligence and Attainment Tests. Chicago: University of Chicago Press, 1980 [An edited version of the original 1960 publication]

[5] Stevens S. On the Theory of Scales of Measurement. *Science*. 1946;103(2684):677-80

[6] Krantz D, Luce R, Suppes P, Tversky A. Foundations pf Measurement Vol 1: Additive and Polynomial Representations. New York: Academic Press, 1971

[7] Wright B. Solving measurement problems with the Rasch Model. *J Educational Measurement.* 1977;14(2):97-116