# MAIMON RESEARCH LLC

# REPRESENTATIONAL MEASUREMENT FAILURE IN HEALTH TECHNOLOGY ASSESSMENT

# HEALTH TECHNOLOGY ASSESSMENT REBUILT:

# EVIDENCE AND VALUE

## A MEASUREMENT BASED INTERROGATION USING AI AND LARGE LANGUAGE MODELS

# PROGRAM INTRODUCTION AND OVERVIEW

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

## INTRODUCTION: A NECESSARY PROGRAM

For those who recognize the imperative of meeting the axioms and theorems of representational measurement, health technology assessment is in crisis. This is not a matter of technical disagreement or methodological preference. It is a failure at the level of scientific foundations. After more than forty years of development, HTA presents itself as a quantitative, evidence-based discipline, yet the constructs on which it relies have not been shown to satisfy the requirements for measurement.

What has changed is that this failure can now be demonstrated. With the use of AI and large language models, it is possible to interrogate and deconstruct the HTA knowledge base in a systematic and reproducible way. National frameworks, research groups, academic journals, and core components of HTA—multiattribute instruments, utility systems, the QALY, and the reference case—can all be evaluated against a common standard. The results, expressed as categorical probabilities and transformed into normalized logits, provide a consistent and unambiguous picture.

Across more than one hundred standardized interrogations spanning over thirty countries and multiple institutional settings, the findings converge. The understanding of measurement theory of the requirement that measurement must precede arithmetic is largely absent. The 24-item canonical diagnostic, covering both true propositions grounded in measurement theory and false propositions embedded in standard HTA practice, reveals a systematic pattern. True statements are weakly endorsed or ignored, while false statements are strongly supported. This is not random error. It is a structured inversion of the principles required for measurement.

HTA therefore operates within what can only be described as a state of measurement inversion. Arithmetic is applied without first establishing that the quantities involved possess the properties required to support it. Utilities are treated as if they were interval or ratio measures, multiplied by time, and aggregated into QALYs. Composite cost estimates are combined with these constructs to produce cost-effectiveness ratios. These outputs are then interpreted as evidence and used to inform decisions that affect patient care.

The conclusion is unavoidable. HTA is not failing at the margins; it is failing at its core. Yet this also presents an opportunity. With the tools now available, the knowledge base can be interrogated, its limitations made explicit, and a path established toward a framework grounded in measurement and the evolution of objective knowledge.

## STANDARDIZED LOGIT PROFILES

**This website presents over 100 logit profiles, each with categorical probabilities, logits, and a detailed review, providing a comprehensive picture of measurement failure in HTA. Profiles are organized by country, allowing readers to examine their own setting and the consistent pattern of measurement inversion across knowledge bases.**

## WHY THIS COURSE IS IMPORTANT

Health technology assessment has, over the past four decades, established itself as the principal mechanism for informing decisions on pricing, access, and the allocation of healthcare resources. It presents itself as a scientific enterprise, grounded in evidence, quantitative analysis, and transparent reasoning. The language of HTA is the language of measurement: outcomes are quantified, benefits are compared, and value is expressed in numerical terms. This appearance of rigor has supported its widespread adoption across national health systems and its authority in shaping patient access to therapies.

Yet this foundation has remained largely unexamined. The central assumption, that the quantities employed in HTA are measures, has not been subjected to systematic scrutiny. Instead, constructs such as preference-based utilities, the QALY, and cost-effectiveness ratios have been accepted as if their measurement properties were self-evident. Over time, these constructs have become embedded within guidelines, academic literature, and institutional practice, forming what appears to be a coherent and stable framework.

The difficulty is that coherence does not imply validity. When examined against the axioms of representational measurement, the core components of HTA fail to meet the conditions required for measurement. Utilities derived from ordinal preference data lack the properties of interval or ratio scales. Their multiplication by time to produce QALYs presupposes scale characteristics that are not present. Composite cost estimates combine heterogeneous resource units into aggregates that lack dimensional homogeneity. The resulting cost-effectiveness ratios are therefore constructed from quantities that cannot support the arithmetic operations applied to them.

These concerns are not new. What is new is that they can now be demonstrated. Through the interrogation of HTA knowledge bases using large language models, it is possible to evaluate systematically whether the field recognizes and applies the principles of measurement. The results are consistent across agencies, journals, and national frameworks. The logit profiles derived from these interrogations reveal a clear pattern: propositions that reflect the axioms of measurement are weakly endorsed or absent, while propositions embedded in standard HTA practice are strongly supported. This pattern of measurement inversion is not confined to particular settings. It is a defining feature of the HTA knowledge base.

The implications are profound. HTA decisions determine which interventions are funded and which are denied. They influence clinical practice and patient outcomes. A framework that does not measure cannot produce evidence in the scientific sense. Without measurable attributes, claims cannot be interpreted, compared, or evaluated. Simulation models may generate projections, but these projections are contingent on assumptions and cannot be subjected to empirical test within meaningful timeframes. The result is a system that produces numerical outputs without a basis for validation.

This is not a matter of improving data quality or refining models. It is not a question of adjusting parameters or adopting alternative assumptions. The problem is structural. The current HTA framework is built on constructs that do not meet the requirements of measurement and therefore

cannot support its claims to evidence-based decision-making. Incremental change cannot address this. A different foundation is required.

That foundation is provided by a measurement-based framework. It begins with the recognition that measurement precedes arithmetic. Value claims must be grounded in attributes that can be measured. For observable phenomena, this requires linear ratio scales. For latent constructs, it requires transformation through Rasch measurement to achieve invariant scales. These are not methodological preferences; they are the conditions under which quantitative claims can be made and evaluated.

The conclusion follows directly. If HTA is to retain credibility as a scientific discipline, it must move to a framework grounded in measurement. This is not an option among alternatives. It is an imperative.

## INTERROGATING A KNOWLEDGE BASE

A knowledge base, in the strict sense required for scientific inquiry, is not simply a repository of information but a structured body of claims, assumptions, rules, and inferential commitments that together define what is accepted as valid knowledge within a domain. It is both an archive and a logic system. Its defining feature is not volume, but coherence and constraint: what can be said, what must be assumed, and what is excluded.

At a minimum, a knowledge base has three core parameters. The first is ontological scope: what entities or phenomena are recognized as legitimate objects of inquiry. In health technology assessment, for example, these might include clinical outcomes, patient-reported experiences, resource use, and preference scores. The second is epistemological rules: the criteria by which claims are judged to be knowledge. These include standards for evidence, replication, measurement, and inference. The third is formal structure: the mathematical, statistical, or logical frameworks used to express and manipulate claims. This includes scale types, permissible operations, and model structures. Together, these parameters determine not only what is known, but what can be known.

The content of a knowledge base is correspondingly layered. At the most visible level are explicit statements: published findings, guidelines, models, and accepted metrics. Beneath these lie implicit assumptions: rarely stated but widely shared beliefs about what constitutes a valid measure, what kinds of evidence are acceptable, and what transformations of data are permissible. Deeper still are axiomatic commitments, often inherited from earlier intellectual traditions, which define the permissible relationships between objects, numbers, and operations. These axioms are rarely interrogated, yet they are decisive. They determine whether a system produces meaningful measures or merely numerical artifacts.

A functioning knowledge base also exhibits closure properties. It is not an open-ended collection of ideas but a system in which new claims must be expressed in terms that are already recognized. This creates path dependence. Once a particular representation, say a composite index or a simulation model is accepted, it becomes increasingly difficult to introduce alternatives that violate

its assumptions. The knowledge base thus becomes self-reinforcing, not through explicit defense, but through structural inertia.

To interrogate a knowledge base is to move beyond its surface content and examine its internal consistency, measurement validity, and susceptibility to falsification. This is not a matter of literature review; it is a process of structured challenge. The first step is to identify canonical statements that express the core commitments of the system. These statements may be explicit ("utilities are interval measures") or implicit ("it is acceptable to aggregate across dimensions without demonstrating unidimensionality"). Each statement is then evaluated against established principles—in your case, the axioms of representational measurement.

The second step is to assess endorsement within the corpus. This involves determining the extent to which the knowledge base supports, ignores, or contradicts each canonical statement. Traditionally, this might involve systematic review and expert judgment. More recently, it can be operationalized through large-scale text interrogation using language models, which estimate the probability that a given statement is consistent with the corpus. These probabilities can then be transformed into logits to provide a continuous measure of alignment or contradiction. The result is not a summary of opinions but a profile of structural coherence.

The third step is to examine logical closure and contradiction. A knowledge base that simultaneously endorses incompatible statements, for example, treating ordinal data as if they support ratio operations, reveals a fundamental incoherence. Such contradictions are not minor technical errors; they indicate that the system lacks a consistent set of axioms. In a scientific context, this is fatal. Without consistency, there can be no stable representation of reality.

The fourth step is to evaluate falsifiability and replicability. A credible knowledge base must generate claims that can be tested and potentially refuted. If its core outputs such as modeled projections or composite indices cannot be independently evaluated or reproduced, then it fails as a scientific enterprise. Interrogation therefore asks not only whether claims are internally consistent, but whether they are empirically accountable.

Finally, interrogation must consider evolutionary capacity. A viable knowledge base is not static; it must be able to incorporate new evidence and revise its assumptions. However, this requires openness to foundational challenge. Systems that treat their core constructs as given, immune to scrutiny, become closed belief systems rather than scientific frameworks. They persist not because they are valid, but because they are insulated.

In sum, a knowledge base is defined by its scope, rules, and structure; populated by explicit and implicit content; and sustained by closure properties that resist change. To interrogate it is to expose its underlying commitments, test them against formal principles, and assess whether the system can support credible, evaluable, and replicable claims. Where it cannot, the conclusion is not refinement but reconstruction.

A critical feature of any knowledge base interrogation particularly when supported by large language models is that the primary knowledge base is fixed at a defined temporal cutoff. This fixed corpus constitutes the reference knowledge base. It is a bounded set of training data,

incorporating published literature, methodological frameworks, and domain-specific content available up to a specific date. In the present case, for models such as those developed by OpenAI, this cutoff represents a hard epistemic boundary: no information beyond that point is part of the foundational knowledge structure unless explicitly introduced during interaction.

This fixed reference base is essential for two reasons. First, it provides reproducibility. Interrogations conducted against the same underlying corpus, using the same canonical statements, should yield consistent probability structures (subject to stochastic variation). Second, it establishes a common evaluative frame. The knowledge base being interrogated is not fluid or continuously updated during the analytical process; it is stable, allowing for meaningful comparison across jurisdictions, institutions, or timepoints.

However, this is only part of the system. As interrogation proceeds, a second layer emerges: a supplementary, interaction-derived knowledge base. This is constructed through the accumulation of user inputs, clarifications, constraints, and previously generated outputs within a session or, where enabled, across sessions via memory functions. This layer does not replace the fixed reference base but conditions how it is accessed and expressed.

In practical terms, the supplementary knowledge base serves several functions. It provides contextual framing, allowing the system to align responses with the user's specific analytical objectives; for example, emphasizing representational measurement theory, Rasch requirements, or the rejection of composite metrics. It also enables iterative refinement, where prior outputs are corrected, extended, or constrained in subsequent interactions. Over time, this produces a form of local coherence: responses become increasingly consistent with the user's methodological stance and preferred terminology.

Importantly, this supplementary layer does not introduce new validated knowledge in the scientific sense. It does not alter the underlying corpus or expand the evidentiary base beyond the cutoff. Rather, it acts as a filter and weighting mechanism, influencing which elements of the fixed knowledge base are prioritized, how conflicts are resolved, and how arguments are structured. In this sense, it is analogous to a working interpretive framework applied to a stable dataset.

The distinction between these two layers has direct implications for knowledge base interrogation. The fixed reference base ensures that the interrogation is anchored in a defined and reproducible corpus. The supplementary interaction base allows the interrogation to be directed**,** cumulative, and adaptive, reflecting the evolving focus of the analysis. Together, they create a system that is both stable and responsive.

From a methodological standpoint, it is essential to recognize that results derived from such interrogations are a function of both layers. The probabilities assigned to canonical statements reflect the interaction between the underlying corpus and the interpretive constraints imposed during the session. Consequently, transparency requires that both the cutoff-defined reference base and the presence of an accumulating interaction context be acknowledged.

## A NEW FRAMEWORK FOR HEALTH TECHNOLOGY ASSESSMENT

The limitations of the current HTA approach are not the result of isolated technical weaknesses. They are structural. They arise from the use of constructs that do not meet the requirements of measurement and from a decision process that substitutes numerical outputs for evaluable evidence. These limitations cannot be addressed through incremental refinement. They require a fundamental shift in how value claims are defined, measured, and evaluated.

The proposed framework begins with a simple but essential principle: measurement precedes arithmetic. Value claims must be grounded in attributes that can be measured. For manifest attributes, such as hospitalizations, drug possession, or mortality, this requires linear ratio scales with a true zero and equal intervals. These attributes are directly observable and support meaningful comparison. For latent constructs—such as quality of life or symptom burden—measurement requires transformation through models that satisfy the conditions of invariance, most notably the Rasch model. These two approaches constitute the only defensible basis for measurement in HTA.

On this foundation, the framework is organized around single, protocol-driven value claims. Each claim specifies the attribute of interest, the target population, the measurement standard, and the timeframe for evaluation. This replaces composite constructs and long-term simulation with clearly defined hypotheses that can be tested within meaningful timeframes. The emphasis shifts from projecting outcomes to observing and evaluating them.

This reorientation restores the connection between evidence and decision-making. When attributes are measurable, differences between interventions can be interpreted and compared. Claims can be replicated across settings and refined over time. The role of modelling changes accordingly. Rather than generating non-falsifiable projections, models support study design, hypothesis formulation, and the identification of measurable endpoints.

A central feature of this framework is accountability. Measurable claims and explicit protocols create the conditions for evaluation. Outcomes can be assessed against expectations, and decisions can be revisited in light of evidence. This is the mechanism through which objective knowledge evolves. Without measurement, this process is not possible.

The transition to a measurement-based framework is not optional. A system that does not measure cannot produce evidence, and without evidence there can be no defensible basis for decisions that affect patient care. This framework sets out the principles and practical steps required for implementation. It is not an alternative to the current HTA standard of measurement inversion, but a replacement. unless the axioms of representational measurement are to be set aside in favor of simulation models that generate imaginary claims.

## WHAT WILL THE HTA PROFESSION GAIN—AND WHERE DO WE GO AFTER FORTY YEARS?

For the professional working in HTA, the question is no longer whether the current framework can be refined. After forty years of development, it is clear that the issue is not one of incremental improvement but of foundations. The interrogation of HTA knowledge bases, now possible through AI and large language models, has made this explicit. Across agencies, journals, and national systems, the same pattern emerges: a failure to meet the axioms of representational measurement. The implication is unavoidable. The profession is faced with a choice: continue within a framework that cannot produce evidence, or move to one that can.

What this program offers is not simply a critique, but a professional advantage grounded in clarity. It provides the ability to distinguish between numbers that measure and numbers that merely persuade. This is not a semantic distinction. It changes how value claims are interpreted, how submissions are evaluated, and how decisions are justified. The professional who completes this program will be able to identify when a cost-effectiveness claim lacks a measurable foundation, when a utility-based model is applying arithmetic to ordinal data, and when a simulation output cannot be evaluated as evidence. These are not abstract skills. They go directly to the credibility of HTA as a decision-making discipline.

More importantly, the program provides a way forward. It sets out how to construct value claims that are measurable, testable, and replicable. It replaces composite constructs with clearly defined attributes. It distinguishes between manifest attributes, which require linear ratio scales, and latent constructs, which require transformation through Rasch measurement. It introduces protocol-driven claims that can be evaluated within meaningful timeframes. In doing so, it repositions HTA as a discipline capable of contributing to the evolution of objective knowledge.

This is where the profession must now go. The past forty years have been characterized by the institutionalization of a framework that has not been subjected to measurement standards. The result is a body of work that is internally coherent but externally untestable. Continuing along this path is not neutral. It implies acceptance of a system that produces claims without a measurable basis and decisions without accountability.

The transition is therefore not optional. It is a requirement if HTA is to retain any claim to scientific credibility. This program equips professionals to make that transition. It provides the conceptual foundation, the methodological tools, and the practical framework needed to move from numerical storytelling to measurement-based evaluation.

The question is not whether change is needed. The question is whether the profession is prepared to act on what is now clear. A new HTA framework is imperative

## PROGRAM OVERVIEW

This program was developed in response to feedback from colleagues reviewing the logit profiles, initially for the United States and the United Kingdom. Their conclusion was clear: there is now compelling evidence of measurement inversion or more bluntly, measurement failure in health technology assessment. In the absence of an established textbook or a body of peer-reviewed work addressing this issue, there was a need for a structured course at the level of a first-year graduate program.

The objective was to create a program that is accessible, professionally relevant, and priced to reflect the cost of production and distribution. Given the global relevance of the findings, rapid and wide dissemination was considered essential, supported by an established network of over 8,000 professional contacts.

The program is organized into three units, each comprising four modules. Each module includes detailed notes and supporting Questions and Answers, providing both conceptual grounding and practical application. The full program comprises approximately 85,000 words of material.

The program is priced at US$250.00. Student discounts can be arranged for class use. All payments will go through a secure PayPal portal. Logit country profiles are free to download.

### PROGRAM ENROLLMENT

**IF YOU WISH TO ENROLL IN THIS PROGRAM PLEASE CLICK HERE AND YOU WILL BE TAKEN TO THE PAYPAL PORTAL. THIS WILL GIVE YOU IMMEDIATE ACCESS TO ALL 12 MODULES FOR 6 MONTHS. ALL MODULES CAN BE DOWNLOADED TO CREATE A MASTER FILE FOR PERSONAL USE.**

To support continued engagement, subscribers will have access to additional specialized modules, with two already available, as well as small-group Zoom workshops. These are designed to extend and apply the core concepts introduced in the program.

For further information or comments, please contact: langleylapaloma@gmail.com

**PROGRAM CONTENT**

**MAIMON FOUNDATION COURSE**

**HEALTH TECHNOLOGY ASSESSMENT REBUILT: EVIDENCE AND VALUE**

**TABLE OF CONTENTS**

## UNIT 1: FOUNDATIONS OF MEASUREMENT IN SCIENCE

### Module 1: ATTRIBUTES AND THE LOGIC OF MEASUREMENT

- **Attributes and measurement in scientific inquiry**
- **Distinguishing objects and the attributes they possess**
- **Examples of attributes in clinical and health services research**
- **Variation and ordering along an attribute dimension**

**Questions & Answers**

1. **Why are attributes the starting point for therapy impact assessment?**
2. **Why must we distinguish between objects and the attributes they might possess?**
3. **Why is ordering the first structural step required for measurement?**
4. **What do we mean by "mapping between empirical structure and numerical representation?**

### Module 2: SCALES OF MEASUREMENT AND DIMENSIONAL HOMOGENEITY

- **Stevens' typology of measurement scales**
- **Stevens' typology, latent attributes and the Rasch transformation**
- **Stevens and dimensional homogeneity**
- **Linear ratio and logit ratio measures cannot be combined**

**Questions & Answers**

1. **Why does HTA dismiss Rasch in favor of ordinal scales for latent constructs?**
2. **Arithmetic depends entirely on the measurement properties of the scale being used, Explain**
3. **Stevens' typology was critical but is only the starting point. Explain**
4. **Where manifest and latent claims are presented to support product acceptance, they cannot be combined to create a composite claim. Explain**

### Module 3: THE AXIOMS AND THEOREMS OF REPRESENTATIONAL MEASUREMENT

- **Understanding axioms**
- **The key axioms of fundamental measurement**
- **From axioms to theorems in representational measurement**
- **The absence of measurement axioms in health technology assessment**
- **Dimensionless numbers and dimensional homogeneity**
- **The origins of dimensional homogeneity**
- **Discounting and the failure of measurement**
- **Why the QALY fails dimensional homogeneity**

- **Measurement inversion in health technology assessment**

**Questions & Answers**

1. **What do we mean when we say that axioms are constraints on permissible operations?**
2. **The QALY fails for one simple reason: dimensional homogeneity. Explain**
3. **What do we mean by measurement inversion as recognized in the logit analysis?**
4. **What is the role of theorems in implementing the axioms of representational measurement?**

### Module 4: REPRESENTATIONAL MEASUREMENT AND LARGE LANGUAGE MODELS

- **Understanding the large language model**
- **Textual response and interpretation**
- **A fixed information system**
- **Interrogating the HTA belief system**
- **The 100 plus interrogations**
- **The large language model as a paradigm enforcement tool**
- **Application of the large language model as enforcer**
- **The large language model and dynamics of change**
- **The memeplex of false measurement**

**Questions & Answers**

1. **What is a large language model?**
2. **What does large language model interrogation mean?**
3. **Why is the large language model described as a fixed reference point?**
4. **How can a large language model be seen as an enforcer within a paradigm?**

## UNIT 2: MEASUREMENT FAILURE IN CONTEMPORARY HEALTH TECHNOLOGY ASSESSMENT

### Module 5: CHEERS 2022 - THE FLAGSHIP OF MEASUREMENT FAILURE IN HEALTH TECHNOLOGY ASSESSMENT

- **The origins of false measurement in HTA**
- **The irrelevance of representational measurement**
- **The triumph of the ordinal scale**
- **The global HTA memeplex**
- **Relativism and CHEERS 2022**
- **Deconstructing CHEERS 2022**

- **Interrogation results" logit distribution**

**Questions & Answers**

1. **In the CHEERS 2022 deconstruction what does 'measurement inversion' mean?**
2. **What initial decision guaranteed HTA measurement failure?**
3. **What do we mean by the triumph of the ordinal scale in HTA?**
4. **What do we mean by 'interrogating the knowledge base'?**

## Module 6: THE QLQ-C30 AS THE EPITOME OF ORDINAL SUMMED SCORES AND MEASUREMENT FAILURE

- **Why ordinal scores fail representational measurement?**
- **Defining the QLQ-C30 Knowledge Base**
- **Deconstructing the QLQ-C30**
- **Interrogation results: logit distribution**
- **Can the QLQ-C30 be rescued?**
- **The consequence of measurement failure**
- **Duty of care and objective knowledge**
- **Is the cat out of the bag**

**Questions & Answers**

1. **Why doesn't the QLQ-C30 produce quantitative estimates of latent attributes?**
2. **Why is the QLQ-C30 knowledge base incomplete?**
3. **Why is the rejection of Rasch measurement in the QLQ-C30 logit profile so significant?**
4. **Could a case be made that the measurement status of the QLQ-C30 means a failure in duty of care?**

## Module 7: THE GENESIS OF MEASUREMENT FAILURE WITH THE EQ-5D-3L AND EQ-5D-5L

- **A red flag for a fatal error**
- **Utilities are not measures**
- **The knowledge base of the EQ-5D instruments**
- **The EQ-5D interrogation**
- **Interrogation results: logit distribution**
- **The EQ-5D instruments: design for failure**
- **Implications for future use**

**Questions & Answers**

1. **Was the application of multiattribute instruments to create utilities doomed from the outset?**
2. **Why does the time trade-off technique fail the standards of representational measurement?**
3. **Utilities are only a number: why is there no equivalence in measurement theory?**
4. **Does the ultimate triumph of multiattribute utilities guarantee QALY failure through the lack of dimensional homogeneity?**

### Module 8: THE FAILURE OF THE REFERENCE CASE KNOWLEDGE BASE FOR SIMULATION MODELLING

- **Thirty years of numerical storytelling with the NICE reference case**
- **Deconstructing the reference case**
- **The reference case knowledge base**
- **Interrogating the knowledge base**
- **Interrogation results: logit distribution**
- **Quality of manufacturer submissions**
- **Measurement: decisions without understanding**
- **Meningitis: decisions without measurement**
- **Global measurement illiteracy**
- **CHEERS 2022: promoting the reference case illiteracy**
- **The reference case as a measurement dead end**
- **Duty of care and meaningless cost-effectiveness claims**
- **Measurement and objective knowledge**
- **From numerical storytelling to a true HTA framework**

**Questions & Answers**

1. **Could we characterize the reference case as a measurement wasteland?**
2. **The impossible QALY effectively demolishes the reference case? Explain**
3. **Can we claim that the reference case establishes whether a product is cost-effective?**
4. **Why has CHEERS 2022 no choice but to endorse the reference case?**

## UNIT 3: A NEW FRAMEWORK FOR HEALTH TECHNOLOGY ASSESSMENT

### Module 9: MEASUREMENT AND CLAIMS FOR MANIFEST ATTRIBUTES

- **Demonstrating measurement: the axioms governing manifest attributes**
- **Constructing evaluable claims for manifest attributes**
- **Clinical claims as manifest attributes**
- **The importance of compliance in the assessment of manifest attribute claims**

- **Agreement on attribute claims in formulary submissions**
- **Protocols to support the evaluation of manifest claims**

**QUESTIONS & ANSWERS**

1. **All therapeutic claims must evaluate attributes. Explain**
2. **How do we know that an attribute satisfies the axioms required for measurement?**
3. **Why should individual attribute claims be supported by a protocol?**
4. **Why is compliance important for claims assessment?**

## Module 10: RASCH MEASUREMENT AND INSTRUMENT DEVELOPMENT FOR LATENT ATTRIBUTES

- **The challenge of the Rasch model**
- **The logit framework**
- **Why logits were the only way forward**
- **Item selection in Rasch instrument development**
- **Creating a common logit scale for persons and items**
- **Estimating fit to the Rasch model**
- **Rasch software**

**QUESTIONS & ANSWERS**

1. **Latent attributes require transformation through Rasch modelling before they can support quantitative claims. Explain**
2. **What is the Rasch logit ratio scale of measurement?**
3. **Applying the Rasch model involves detailing separate attribute claims measured in logits. Why?**
4. **Why do we describe the Rasch model as the necessary and sufficient condition for measuring therapy impacts on latent attributes?**

## Module 11: RASCH MEASUREMENT AND ATTRIBUTE POSSESSION

- **Possession and the distribution of items**
- **Information and item distribution**
- **Assigning item difficulty on the logit scale**
- **Item difficulty is not psychological difficulty**
- **Estimating attribute possession uses the entire pattern of responses**
- **Instrument item response to respondent possession logit**
- **Statistical inference with Rasch attribute logit measures**
- **Attribute measurement and item endorsement probabilities**
- **Reporting Rasch therapy impact results**
- **Logit linear transformation**
- **Logits and probabilities**
- **From the Rasch logit interval scale to the Rasch logit ratio scale**

**QUESTIONS & ANSWERS**

1. **What do we mean by the possession of a latent attribute?**
2. **The information contributed by an instrument item depends on the relationship between item difficulty and respondent ability. Explain**
3. **The Rasch estimate of possession of a latent attribute is based on the entire pattern of item response. Why?**
4. **Because logits represent the logarithm of odds, differences in logits correspond to multiplicative changes in the odds of endorsing items. Explain**

## Module 12: FROM MEASUREMENT TO DECISIONS IN IMPLEMENTING A NEW HTA FRAMEWORK

- **The end of the reference case**
- **The lessons of logit profiles**
- **The two pillars of measurement**
- **The structure of a value claim**
- **From individual measurement to population measurement**
- **Replication, reproduction and the evolution of knowledge**
- **Decision making without models**
- **Contracting and accountability**
- **The new role of health technology assessment**

**QUESTIONS & ANSWERS**

**1.Why have the logit profiles been so critical in deconstructing the current HTA belief system?**

**2. Why are there only two pillars of measurement in HTA?**

**3. To what extent are value claims protocols the first step participating in the evolution of objective knowledge?**

**4. Could we claim that accountability is the hallmark of the proposed HTA framework?**