Opportunities and Pitfalls When Delphi Expert Consensus Substitutes for Missing Evidence in Rare Diseases- A Targeted Literature Review

AUTHORS: Deepthi Chinchapattanam, M.Pharm[1], **Manasa Vishnubhotla, MS[1],** Mahendra Kumar Rai, PhD[2].

1 Trinity Life Sciences, Mumbai, India; 2 Trinity Life Sciences, Singapore, Singapore;

RWD20

Poster presented at ISPOR | Asia Pacific Summit 2026

TRINITY

1 Background

- Rare diseases pose a structural evidence-generation problem that conventional clinical research methods struggle to solve.
- Small, geographically dispersed populations, heterogeneity, and long natural-history timelines constrain adequately powered trials and slow real-world evidence generation.
- Gaps affect diagnostic criteria, treatment sequencing, HRQoL and utility, resource use, and economic model inputs.

EVIDENCE GAP

SMALL	DISPERSED	HETEROGENEOUS
Few patients	Across geographies	Variable courses

Delphi is a structured, iterative way to turn dispersed expert experience into a documented statement, threshold, pathway, or model input. Its appeal in rare diseases is intuitive: when trial-level or registry data cannot exist at meaningful scale, aggregating the tacit knowledge of clinicians and other stakeholders who manage these patients is often the only feasible alternative.

Four-step Delphi cycle

1 | Recruit experts → 2 | Anonymous rating → 3 | Controlled feedback → 4 | Assess consensus

2 Objective

To evaluate, through a targeted literature review of the rare disease literature, the applications, methodological characteristics, strengths, and limitations of Delphi and modified Delphi methods used to address missing clinical, epidemiological, humanistic, and economic evidence, and the opportunities and pitfalls when consensus is used as a surrogate for empirical data in HEOR and HTA contexts.

3 Methods

- Design: Targeted literature review for landscape mapping across heterogeneous rare disease literature.
- Data source: Structured PubMed search with predefined eligibility criteria.
- Synthesis: Descriptive counts and proportions plus thematic synthesis.
- Classification: Evidence-gap type, disease area, methodological design, and reported limitation type.

Five-step eligibility flow

A English language → B Study design → C Rare or orphan disease → D Delphi or structured elicitation → E Outcome of interest

Study identification and screening flow

287 records screened

- 256 included 89%
- 31 excluded 11%

EXCLUSION REASONS

Population not of interest, n=14
Study design not of interest, n=2
Intervention not of interest, n=4
Non-English, n=11

Extraction and categorization

- Included records captured bibliographic identifiers, abstracts, decisions, exclusion reason where relevant, and co-occurring study-type tags: Clinical, HRQoL, Economic, HCRU, and Economic Modelling or Decision-analytic.
- Studies were categorized by evidence-gap use, disease area, Delphi design, panel features where reported, and recurring limitation type.

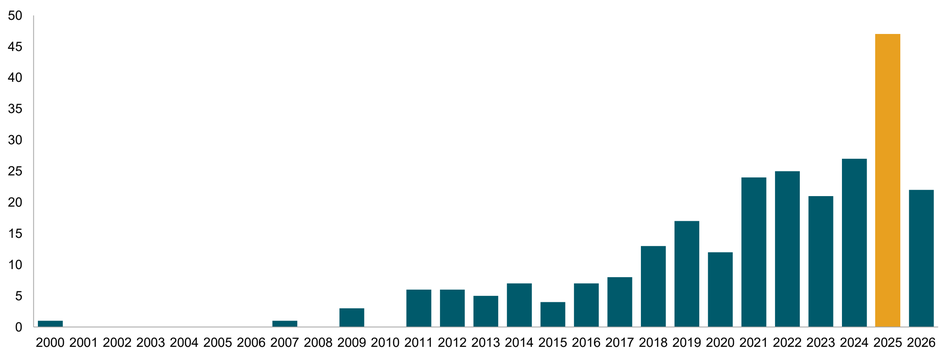
The review maps where Delphi is used, how it is implemented, and which strengths and weaknesses recur across rare-disease evidence generation

Evidence-to-decision pathway

Evidence gap → Structured elicitation → Consensus output → Decision use → Future validation

4 Results

A rapidly growing Delphi and other consensus approaches evidence base

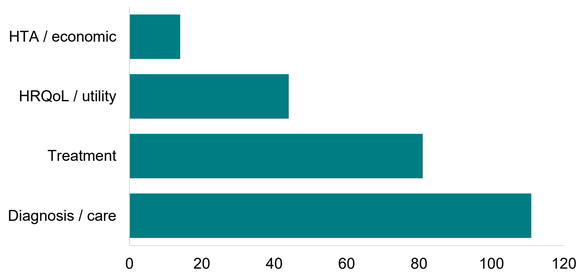


47
records in 2025
Evidence use is accelerating.

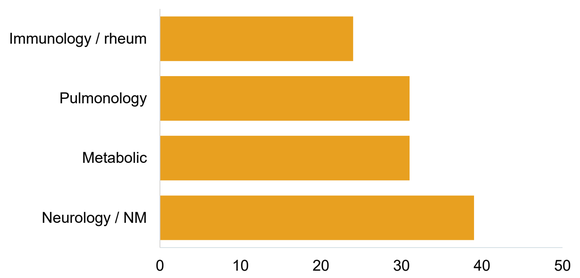
Delphi and other consensus approaches are becoming more common in rare-disease research. 2000-mid-2026; 256 records. <10/year before 2017; 24-27 annually in 2021-2024; 47 in 2025; 22 in H1 2026.

Predominant evidence-gap use cases and leading disease areas

Decision contexts



Frequently represented disease areas



Consensus is used most often to define diagnosis and care; 196 of 256 records (77%) were rare disease populations. Smaller use in HTA and economic modelling carries higher decision risk. Modified Delphi was the most common design (n=42); 40 panels international, 31 online, but only 29 pre-specified a consensus threshold.

4 Results

Opportunities and Pitfalls

OPPORTUNITY	PITFALL
Visible Makes tacit expert knowledge documented, transparent, and citable.	**Selection** Panel recruitment may not be representative.
Feasible Enables criteria and pathways where no gold standard or adequate trial exists.	**Voice** Specialist dominance can crowd out patient and caregiver perspectives.
Iterative Anonymity, repeated rounds, and controlled feedback can reduce dominance and social pressure.	**Validation** Consensus may never be checked against external or emerging data.
Timely Supports decisions that cannot wait for mature trials or registries.	**Rules** Thresholds and stopping criteria may be undefined or post hoc.
Opportunity is greatest when empirical evidence is genuinely unobtainable and the process is explicit.	**Precision** A precise-looking number can hide profound uncertainty.

Delphi Expert Consensus

Structured expert judgment

Recurring Weaknesses Reported

24	20	14	10	6
Transparency	General methods	Validation	Representation	Thresholds

Delphi creates value when expert judgment is transparent and testable. It creates risk when consensus appears empirically precise without transparent selection, thresholds, validation, or uncertainty.

HEOR and HTA implications

Missing empirical parameter → Structured expert elicitation → Utility / resource use / transition probability → Economic model → HTA / reimbursement decision

	Clinical Guideline Use	HEOR and HTA Use
Primary value	Codifies tacit clinical knowledge and care pathways.	Supplies parameters when empirical estimates are unavailable.
Main risk	Overgeneralization beyond the panel population	An unvalidated or threshold-ambiguous estimate may materially influence an ICER or reimbursement decision.
Safeguard	Transparent selection, multidisciplinary representation, and reporting of disagreement.	Explicit uncertainty, sensitivity analysis, and clear labeling of expert-derived inputs.

1. Use Delphi to prioritize future real-world evidence, not permanently replace it.
2. Flag every Delphi-derived model input so reviewers can identify expert-derived evidence.
3. Propagate elicited uncertainty through deterministic, probabilistic, and scenario analyses.

Recommendations

1. State why Delphi is appropriate and what evidence gap it addresses.
2. Define the evidence gap before panel recruitment.
3. Document expert selection, panel size, and specialty mix.
4. Include multidisciplinary and, where appropriate, patient or caregiver perspectives.
5. Pre-specify consensus thresholds and stopping rules.
6. Report whether thresholds were prospective or post hoc.
7. Report disagreement and dropout, not only consensus.
8. Conduct sensitivity analyses around Delphi-derived model inputs.
9. Avoid unsupported causal or comparative claims based solely on consensus.
10. Follow established reporting guidance such as CREDES.

5 Limitations

- The review used PubMed and was targeted, not a multi-database systematic review. Relevant studies indexed elsewhere may not be captured; the 89% inclusion rate reflects broad population criteria. Panel size, rounds, format, and consensus threshold were available only at aggregate level and are reported where available.
- Definitions of Delphi, modified Delphi, and expert consensus were heterogeneous.
- Rare disease settings vary substantially, limiting uniform generalizability.
- Publication and reporting bias may understate methodological weaknesses.
- Title and abstract screening required judgment for borderline records.

6 Conclusion

- Across 256 included studies, Delphi methods are increasingly used to address evidence gaps in rare diseases, particularly for diagnostic and treatment decision-making.
- Delphi studies also support HTA, HRQoL, and economic evaluations when trial-level evidence is limited or unavailable.
- Key methodological challenges include inadequate transparency, limited validation, panel homogeneity, and inconsistent consensus definitions.
- The credibility of consensus-derived evidence depends on transparent panel selection, pre-specified consensus thresholds, validation where feasible, and explicit treatment of uncertainty.
- Standardized Delphi methodology and reporting practices are essential to strengthen the use of expert consensus in rare disease decision-making and HTA.

FUNDING & DISCLOSURE

- None

ABBREVIATIONS

CREDES, Guidance on Conducting and REporting DElphi Studies; HCRU, healthcare resource use; HEOR, health economics and outcomes research; HRQoL, health-related quality of life; HTA, health technology assessment; ICER, Incremental cost-effectiveness ratio; Neuro/NM, neurology/neuromuscular; rheum, rheumatology. RWE, real-world evidence; TLR, targeted literature review.

REFERENCES

1. Iolascon A, Bianchi P, Andolfo I, et al. Recommendations for diagnosis and treatment of methemoglobinemia. *Am J Hematol.* 2021;96(12):1666-1678. doi: 10.1002/ajh.26340.
2. Schoenmakers DH, Beerepoot S, van den Berg S, et al. Modified Delphi procedure-based expert consensus on endpoints for an international disease registry for metachromatic leukodystrophy: the European Metachromatic Leukodystrophy initiative. *Orphanet J Rare Dis.* 2022;17(1):48. doi: 10.1186/s13023-022-02189-w.
3. Facey K, Granados A, Guyatt G, et al. Generating health technology assessment evidence for rare diseases. *Int J Technol Assess Health Care.* 2014;30(4):416-422. doi: 10.1017/S0266462314000464.
4. Choon SE, van de Kerkhof P, Gudjonsson JE, et al. International consensus definition and diagnostic criteria for generalized pustular psoriasis from the International Psoriasis Council. *JAMA Dermatol.* 2024;160(7):758-768. doi: 10.1001/jamadermatol.2024.0915.
5. Demarest S, Pestana-Knight EM, Olson HE, et al. Severity assessment in CDKL5 deficiency disorder. *Pediatr Neurol.* 2019;97:38-42. doi: 10.1016/j.pediatrneurol.2019.03.017.

CONTACT

- Manasa Vishnubhotla, MS, Trinity Life Sciences, Mumbai, India | mvishnubhotla@trinitypartners.com
- Mahendra Kumar Rai, PhD, Trinity Life Sciences, Singapore | marai@trinitylifesciences.com