From One Regulatory Trial to Many HTA Questions: The Economics of PICO Multiplicity Under the EU Health Technology Assessment Regulation

Ethan Miller [1*], Olivia Carter [2], Jacob Anderson [3]
Aram V. Chobanian & Edward Avedisian School of Medicine, Boston University, Boston, Massachusetts, USA

* Corresponding Author: **Ethan Miller**

Article Info

ISSN (online): 3049-0421
Volume: 02
Issue: 02
March-April 2025
Received: 11-02-2025
Accepted: 12-03-2025
Page No: 60-69

Abstract
The EU Health Technology Assessment Regulation, which has applied to new cancer medicines and advanced therapy medicinal products since January 2025, was designed so that a health technology developer submits clinical evidence once at Union level and Member States stop repeating the same clinical assessment. The Regulation also requires the assessment scope of a joint clinical assessment to be inclusive and to reflect Member States' needs, and the scopes agreed so far have contained several distinct population, intervention, comparator and outcome (PICO) questions. Debate about this PICO multiplicity has concentrated on procedural feasibility, on how many questions simulation studies predict, and on whether indirect comparison methods can answer them. Much less attention has been paid to the economic question that follows, because every additional question consumes scarce assessor and developer capacity and that capacity has an opportunity cost. This article reframes PICO generation as a resource allocation problem. We define four quantities, namely question generation cost, analytical cost, feasibility loss and expected decision value, and use them to develop three ideas. These are a PICO feasibility frontier beyond which the submitted evidence cannot support further questions, analytical deadweight loss, and a cost incidence account of who requests a question, who answers it and who gains from the answer. We propose four tests for prioritising PICOs and argue that the efficiency of joint assessment should be judged by decision value generated per unit of assessment resource rather than by the number of questions addressed.

DOI: https://doi.org/10.54660/IJPGRR.2025.2.2.60-69

Keywords: Health Technology Assessment (HTA), PICO Framework, Resource Allocation, Decision Value

1. Introduction

Regulation (EU) 2021/2282 on health technology assessment (the HTA Regulation) has applied since 12 January 2025 to medicinal products containing a new active substance indicated for the treatment of cancer and to advanced therapy medicinal products, with orphan medicinal products following in 2028 and all remaining centrally authorised products in 2030 [1]. Its stated purpose is to establish a mechanism under which the evidence required for the joint clinical assessment (JCA) of a health technology is submitted by the developer only once at Union level, together with common rules and methodologies for carrying out that assessment [1]. The efficiency argument appears in the Regulation's recitals, which identify parallel assessment by multiple Member States and divergent national evidence requests as a source of duplication [1], and it has a long lineage in European HTA cooperation [3].

The architecture creates a structural tension. One regulatory evidence package and one joint assessment must serve twenty-seven national decision contexts with different treatment pathways [2], different reimbursed comparators and different evidentiary conventions [4, 5, 6]. The Regulation resolves that tension in favour of inclusiveness. Article 8(6) provides that the designated subgroup initiates a scoping process and that the assessment scope "shall be inclusive and reflect Member States' needs in terms of parameters and of the information, data, analysis and other evidence to be submitted by the health technology developer" [1].

A joint assessment therefore does not narrow the assessment question. It aggregates national questions, and the aggregation is expressed as a set of questions defined by population, intervention, comparator and outcomes (PICO).

The consequence is now observable rather than hypothetical. The first JCA report published under the Regulation, on tovorafenib for paediatric low-grade glioma, was published on 9 June 2026 after endorsement by the Member State Coordination Group on Health Technology Assessment (the Coordination Group). Its scope contained eight PICO questions across three populations, and comparative effectiveness evidence was available for only one of them, through an unanchored matching-adjusted indirect comparison [7]. Two further reports followed in July 2026, one with seven PICOs, all addressed by direct or indirect comparisons, and one with a single PICO informed by a randomised controlled trial [8, 9].

A substantial literature describes this phenomenon. Simulation and mock scoping exercises estimate how many consolidated PICOs a JCA is likely to generate [10, 11], survey work documents the comparator heterogeneity that drives the count [4], methodological reviews assess whether indirect comparison techniques can bear the resulting analytical load [12, 13], implementation studies record stakeholder concern about capacity and timelines [14, 15, 16], and procedural responses have been proposed [17, 18]. What this literature does not do is treat PICO generation as an economic decision. It counts questions, tests their answerability and proposes procedures for managing them, but it does not ask whether the marginal question is worth asking. Two conference abstracts propose informal triage heuristics for developers [19, 20], and recent work examines the value of developing an HTA process and the design of appraisal processes as population health optimisation problems [21, 22], yet the question of when an additional PICO earns its analytical cost remains open.

This article addresses four questions. First, how should the economic cost of PICO multiplicity in JCAs be conceptualised? Second, when does an additional PICO generate enough decision value to justify its analytical and evidentiary cost? Third, how does PICO multiplicity distribute costs across Member States, assessors and developers? Fourth, under what conditions does joint clinical assessment reduce total European assessment burden, and when can coordination costs offset the savings from avoided duplication? In answering them we reframe PICO generation as a marginal resource allocation problem, define a PICO feasibility frontier and the analytical deadweight loss arising beyond it, set out a cost incidence account, and propose four tests for prioritisation together with an efficiency measure for the joint system. We use evidence available to 8 August 2026 and avoid building the argument on current assessment counts, which will change quickly.

2. Why PICO Multiplicity Arises

PICO multiplicity is a designed feature of the system, not a malfunction. Under the Coordination Group's scoping guidance, assessors draft an assessment scope proposal, Member States identify their national needs through a PICO survey, and the responses are consolidated. The guidance is explicit that "the objective of the consolidation is to ensure that MS needs are translated in the lowest possible number of PICOs" [23]. Consolidation therefore already embodies a restraint principle, but the principle is expressed as a rule about counting rather than about value, and it operates only after national requests have been made.

Three mechanisms generate the requests. The first is comparator heterogeneity. A mapping of publicly available HTA sources across twenty-nine European countries found that population and outcome requirements are largely aligned between countries and that comparator specification is the dominant driver of divergence [4]. Because standard of care differs by reimbursement history, guideline adoption and local pricing, the same intervention is compared against different alternatives in different systems, and each distinct comparator generates a distinct analysis.

The second is divergence in evidentiary conventions. Comparative work across German, Spanish and French assessment bodies, the European Medicines Agency and the EU HTA framework found that the evidence each will accept differs materially, with Union-level guidance and the regulator more permissive than national assessment bodies [5], and an environmental scan of national methods guidance found broad agreement on systematic evidence identification but substantive divergence on indirect treatment comparison [6]. Multi-stakeholder survey work across seventeen countries had already identified heterogeneity of national treatment standards and evidence requirements as a central obstacle to a joint assessment delivering added value over national practice [24].

The third is the mismatch between regulatory and assessment evidence. Pivotal trials are designed to demonstrate benefit and risk to a regulator, not to answer national comparative questions. Among pivotal studies for rare disease medicines approaching launch, thirty-six per cent were single-arm and eighty-two per cent used a surrogate primary endpoint [25]. Applying JCA criteria retrospectively to products supported by single-arm trials, expert assessment concluded that ninety-four per cent would not qualify for joint assessment on evidence grounds, and that the remainder would support only a limited number of PICOs at low certainty [26]. A retrospective simulation across seven national agencies in onco-haematology generated at least four PICOs per medicine and found the evidence available at marketing authorisation insufficient to address all of them [11].

Two features of the legal framework reinforce these mechanisms. Article 13(1)(d) prevents a Member State from requesting at national level information the developer has already submitted at Union level, which raises the cost of failing to register a requirement during scoping, while Article 13(1)(a) requires only that Member States give due consideration to the published report and preserves their competence to reach their own conclusions [1]. A Member State therefore faces a strong incentive to state every plausible requirement at scoping and a weak obligation to use the resulting output. Legal analysis has noted that disparities between the joint scope and national needs can themselves justify a Member State in setting the report aside, with additional workload for developers and national bodies as the consequence [27]. Indication wording adds a further layer, since the wording approved at authorisation can differ from the wording available when scoping begins, usually editorially but sometimes by adding or omitting a sub-population [28].

3. What the Emerging Evidence Shows

The empirical picture has three components.

The first is volume. A peer-reviewed analysis applying the

consolidation rules developed by the European network for Health Technology Assessment (EUnetHTA) 21 project to two hypothetical medicines produced ten consolidated PICOs for first-line non-small cell lung cancer across six Member States and sixteen for third-line multiple myeloma across nine, rising to fourteen and eighteen when a national scope was used as a proxy for the remaining countries. Because each PICO requires analysis of multiple outcomes, the authors estimated a minimum of 280 and 720 discrete analyses respectively [10]. A targeted review of fourteen publications reporting thirty-five PICO exercises across twenty indications found an average of eight consolidated PICOs per exercise, and the authors' own simulation reflecting twenty-five countries in first-line metastatic non-small cell lung cancer produced sixty-seven [29]. Conference analyses of nineteen and then twenty-four case studies report medians of eight and nine PICOs with upper ranges of eighteen and twenty-one, and around two-thirds requiring an indirect treatment comparison [30, 31]. These conference reports are not peer reviewed and their assumptions differ, but their central tendency matches the peer-reviewed estimates.

The second is answerability, which matters more than volume. A hand search of twenty-three EUnetHTA relative effectiveness assessments identified sixty-four required comparisons, with a median of four comparators per assessment and a range from one to eighteen. Twenty-five comparisons were informed by indirect evidence, and the suitability of the indirect comparison was judged unclear in twenty-four of those twenty-five [12]. Critical review of the Coordination Group's methodological guideline on direct and indirect comparisons argues that its restrictive treatment of unanchored and population-adjusted methods increases rather than absorbs the analytical load created by additional questions [13, 32], and broader analysis of the guidance identified thirty-two topics requiring concretisation, concentrated in evidence synthesis and eligibility criteria [33]. The record from the German orphan drug system shows what happens when a demanding scope meets a thin evidence base [34].

The third is process cost. A proof-of-concept study of automated PICO extraction reported that two indications required the curation of forty-nine national reports totalling 5,463 pages from sixteen countries [18]. Assessor-side gap analyses identify capacity and resources as explicit implementation risks [15], Italian stakeholder interviews concluded that the promised gains in speed depend on how far Member States require complementary analyses [16], and Belgian interview work named the compilation of a harmonised PICO as a primary barrier [35]. Developers, meanwhile, now invest in anticipatory PICO simulation before a scope exists, which is itself a cost created by the architecture [30, 31].

The scale is set out in the Coordination Group's work programme, which anticipates around thirty-five joint assessments of new cancer medicines, fifteen of advanced therapy medicinal products and three on variations during 2026, alongside the first assessments of medical devices [36]. The three reports published to 8 August 2026 contained eight, seven and one PICO [7, 8, 9]. We treat these figures as illustrations of a mechanism rather than estimates of long-run performance. The relevant point is not the current average but the width of the distribution and the fact that answerability varies independently of volume.

Table 1 summarises how characteristics of the regulatory evidence package interact with sources of national divergence to generate analytical work of differing expected value.

Table 1: From regulatory evidence to national assessment questions

Characteristic of the regulatory evidence package	Source of national divergence	Analytical consequence in the joint scope	Dominant cost	Likely decision benefit
Single active comparator chosen for regulatory purposes	Reimbursed standard of care differs by country [4]	Additional indirect comparisons against nationally relevant comparators	Analytical cost, assessor and developer time	High where the comparator is widely used, low where it is used in one or two systems
Broad claimed indication	National restriction to sub-populations by biomarker, line of therapy or age [29]	Sub-population PICOs requiring subgroup analysis	Analytical cost and feasibility loss	Moderate, contingent on subgroup pre-specification and power
Surrogate or composite primary endpoint	National preference for final patient-relevant outcomes [5, 6]	Additional outcome specifications across every PICO	Analytical cost, multiplied across the scope	Variable, high where the surrogate is contested
Single-arm pivotal design	National tolerance of unanchored comparison differs [26, 34]	Unanchored matching-adjusted or naive comparison, or no comparison	Feasibility loss	Low, often no usable answer
Indication wording changed between scoping and authorisation	Timing of national processes [28]	Rescoping, or PICOs that no longer match the licence	Question generation cost, rework	None, this is pure process cost
Disconnected evidence network	Different national comparator sets [12]	Network meta-analysis with unverifiable similarity assumptions	Analytical cost and feasibility loss	Low, with high uncertainty attached

PICO population, intervention, comparator and outcomes

4. The Economics of Additional PICO Questions

4.1. Four Quantities

We characterise any additional PICO question by four quantities. Question generation cost (C_q) is the coordination, survey, consultation and consolidation effort required to formulate and agree the question. Analytical cost (C_a) is the cost of answering it, comprising evidence identification, direct or indirect comparison, subgroup and sensitivity analysis, assessor and co-assessor review, and developer response. Feasibility loss (C_f) is the expected shortfall arising because the submitted evidence cannot answer the question to a standard usable in a national decision. Expected decision

value (V_d) is the expected improvement in the quality, relevance or confidence of the national decisions the question informs.

Decision value is the least familiar of these in scoping practice and the most important. It is not a measure of national relevance but of the probability that the answer changes something. This is the logic of value of information analysis, which holds that evidence is worth acquiring when the expected cost of a decision made without it exceeds the cost of acquiring it, and which underpins the argument that inference for its own sake is irrelevant to a decision maker [37, 38, 39]. Good practice guidance is well established [40, 41], the computational barriers that once made such analysis impractical have largely fallen [42, 43], and the approach has been operationalised for near-real-time prioritisation inside a research organisation [44]. We do not claim that a formal expected value of sample information calculation should be performed for every candidate PICO. We claim that the underlying question, whether resolving this uncertainty would change a decision, is the right question to ask at scoping and is not currently asked.

4.2. Marginal Cost and Opportunity Cost

The costs in Table 2 are real but largely invisible, because no party currently reports them. Systematic review workload gives an order of magnitude, with registered protocols requiring a mean of five people and around sixty-seven weeks from registration to publication [45]. A PICO does not require a full de novo review, but it does require a defensible evidence identification step, a comparison and a certainty assessment, repeated across each outcome in the scope.

Table 2: Cost categories generated by PICO multiplicity

Cost category	Description	Who bears it	Candidate measurable proxy
Question generation	PICO survey, national consultation, consolidation, validation, rescoping after indication change	Member States, assessors, Coordination Group secretariat	Staff hours per scoping round, number of survey iterations
Developer analysis	Evidence identification, direct and indirect comparisons, subgroup and sensitivity analyses, dossier preparation, responses to requests	Health technology developer	Analyst hours and external spend per PICO
Assessor analysis	Review of submitted analyses, certainty assessment, drafting, co-assessor review, Coordination Group review	National HTA bodies and their staff	Assessor hours per PICO and per report
Anticipatory prediction	Simulation of likely scopes before a scope exists	Health technology developer, disproportionately smaller firms	Number and cost of pre-submission PICO simulations
Delay	Extension of assessment timelines attributable to scope breadth	Patients, national payers, developer	Days from procedure start to report publication, by number of PICOs
Feasibility loss	Effort expended establishing that a question cannot be answered	Assessors and developer, jointly	Proportion of PICOs closed without a usable comparison
Displaced work	Assessments, consultations and national appraisals not undertaken because capacity was committed elsewhere	The joint system as a whole	Queue length and time to initiation for subsequent procedures
Downstream duplication	National reanalysis where the joint report does not meet national needs	Individual Member States	Proportion of national appraisals commissioning additional clinical analysis

HTA health technology assessment, PICO population, intervention, comparator and outcomes

What matters economically is not the level of these costs but their opportunity cost. Assessment capacity within national agencies and the Coordination Group's subgroups is fixed in the short run, so analyst hours devoted to an unanswerable question are hours unavailable for another assessment, a joint scientific consultation or national appraisal work. Where additional analysis extends the timeline, part of the opportunity cost is borne in delayed access, and delay carries a measurable health cost as well as an option value [46]. Health opportunity cost is well developed in this literature, with empirical estimates of the health displaced at the margin in England and in Spain [47, 48, 49] and a clear account of what a threshold does and does not represent [50]. Those estimates are not transferable to assessment capacity, but they establish the principle that resources consumed by a process that does not improve decisions displace health elsewhere, and that the displacement is real even though the patients affected are unidentified [51].

4.3. The PICO Feasibility Frontier

Marginal analysis of this kind is not new to health care priority setting, where programme budgeting and marginal analysis has long been used to ask what is gained and given up at the margin of a fixed budget [52]. What is distinctive about PICO generation is that the marginal product falls for two reasons at once.

The first is ordinary diminishing returns. Early questions in a consolidated scope tend to address the populations and comparators that dominate treatment in most Member States, so they carry high expected decision value, whereas later questions address narrower sub-populations, less widely used comparators or outcome specifications requested by one or two countries.

The second reason is specific to a joint system built on a single regulatory evidence package. As the scope widens, an increasing proportion of questions falls outside what the trial programme can support, because there is no comparator arm, no measured outcome, no adequately powered subgroup or no connected network. Beyond this point, which we term the PICO feasibility frontier, additional analytical effort does not produce evidence. It produces a documented statement that evidence is unavailable. Figure 1 represents this relationship. The frontier is a property of the evidence base rather than of the assessors. It moves outward when the developer's trial

programme is designed with national requirements in mind, which is the economic justification for joint scientific consultation, and inward when the indication is broad, the pivotal design is single-arm or the primary endpoint is a surrogate [25, 26].

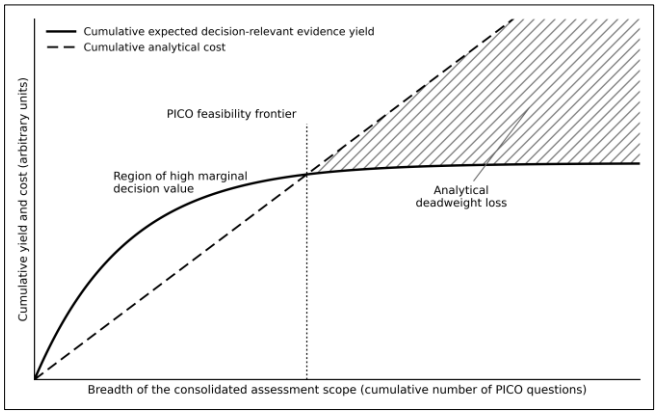

Fig 1

Fig. 1 The PICO feasibility frontier. Both axes are stylised and unitless. The horizontal axis represents the breadth of the consolidated assessment scope, expressed as the cumulative number of PICO questions. The solid curve shows cumulative expected decision-relevant evidence yield, which rises steeply for the first questions and flattens as later questions address narrower populations, less widely used comparators and outcomes the trial programme did not measure. The dashed line shows cumulative analytical cost, which rises approximately linearly with the number of questions and their associated outcomes. The vertical dotted line marks the PICO feasibility frontier, the point beyond which the submitted evidence base cannot support further questions and additional analytical effort yields documented evidence gaps rather than usable comparative evidence. For simplicity the frontier is drawn where the two curves intersect. The hatched area to the right represents analytical deadweight loss. The position of the frontier is a property of the evidence base rather than of the assessment process, and moves outward when trial programmes are designed with national assessment requirements in view. *PICO* population, intervention, comparator and outcomes

4.4. Analytical Deadweight Loss

We define analytical deadweight loss as the assessment resource consumed by analyses that cannot change, clarify or support a national decision. The tovorafenib assessment illustrates the concept without impugning the work that produced it. Eight questions were formulated because eight distinct national needs existed, and the report records that comparative evidence could be brought to bear on one of them [7]. The resource consumed in establishing that the other seven could not be answered was not wasted in a procedural sense, since documenting an evidence gap has value, but it was not decision value in the sense used here and much of it was foreseeable at scoping.

The argument must be kept symmetric. Under-scoping is also costly. Ignoring decision uncertainty imposes real losses, and a joint assessment that omits a comparator genuinely in use in a Member State forces that country either to disregard the report or to commission its own analysis, reproducing the duplication the Regulation was designed to remove [27, 53]. The claim is not that fewer PICOs are better. It is that PICOs should be selected on expected decision value per unit of assessment resource, which will sometimes mean adding a question that no single Member State requested.

4.5. Coordination Cost and the Efficiency Test

The efficiency case for joint assessment rests on avoided duplication [1, 2, 3], but centralisation substitutes one set of costs for another. The insight that organising an activity centrally is worthwhile only up to the point at which the cost of coordination equals the cost of the decentralised alternative is the foundational result of transaction cost economics [54, 55]. Applied here, a joint clinical assessment replaces partially duplicated national assessments with one assessment plus a scoping, consolidation, review and dissemination apparatus. Whether this is efficient depends on how much national activity it displaces and how large the apparatus becomes. A decade of voluntary cross-country collaboration in European pharmaceutical policy points the same way, since sustained collaboration has required dedicated national resources and political commitment rather than delivering savings automatically [56].

PICO multiplicity is the mechanism through which the second term grows. If the joint scope must reproduce the union of national questions, the joint assessment approaches the sum of the national assessments in analytical content while adding the coordination layer on top. This does not make the joint system inefficient, but it makes the efficiency claim contingent and testable rather than automatic. Formal treatment of appraisal process design as a population health optimisation problem, and recent work on the value of developing an HTA process, supply the machinery for such a test [21, 22]. The appropriate summary measure is decision value generated per unit of assessment resource, not the number of questions addressed or the number of Member States whose requests were accommodated.

5 Who Pays and Who Benefits

The costs and benefits of an additional PICO fall on different parties, and the misalignment is systematic.

A Member State that requests an additional question bears the cost of drafting it and almost none of the cost of answering it. The analytical cost falls on the assigned assessor and co-assessor, on the developer, and, through longer timelines, on every Member State awaiting the report. The benefit, if the question is answerable, accrues largely to the requesting

country. The assessment capacity of the joint system is non-excludable to participating Member States but congestible in use, which places it closer to a club good than a pure public good and makes over-appropriation the expected outcome absent governance rules [57, 58]. This is not a criticism of any Member State. Registering a national requirement is individually rational, particularly given the disincentive created by Article 13(1)(d) against relying on a later national request [1].

The distributional consequences are uneven. Larger Member States typically maintain the capacity to conduct their own assessment if the joint report does not meet their needs, and several have well-resourced traditions of demanding comparator specificity [34]. Smaller Member States have the most to gain from a usable joint output and the least capacity to absorb the consequences of an unusable one, yet fewer resources to contribute to assessment and consolidation work. A joint system in which scope is driven by the best-resourced agencies, and in which the resulting reports are then set aside by those same agencies, would transfer cost towards the countries the Regulation was intended to help. Comparative studies of national coverage recommendations attribute divergence to the evidence considered, its interpretation and deliberative judgement, not only to differences in the questions posed [59, 60]. Answering a question at Union level therefore does not resolve the sources of divergence operating downstream of the clinical assessment, and the expected decision value of a nationally specific PICO cannot be assumed from the fact that a Member State asked for it. Proposals to deepen Union-level coordination in pricing and procurement face a version of the same distributional question, namely how the gains and the administrative burden of collective action are shared between larger and smaller markets [61].

Developers face a distinct problem. Because scopes are not known when trials are designed, and often not known when a dossier is prepared, firms invest in predicting them [30, 31, 62]. This anticipatory expenditure is a direct product of scope uncertainty and appears in no assessment budget. It is also regressive, since the fixed cost of PICO prediction weighs more heavily on smaller developers and on products with narrow indications.

6. A Decision Rule for PICO Prioritisation

We propose four tests, set out in Table 3, to be applied during consolidation rather than as an additional procedural stage.

Table 3: Four tests for prioritising PICO questions during consolidation

Test	Question to be answered	Practical indicator	Consequence if the test is not met
Decision relevance	Would a credible answer plausibly change or materially clarify a national appraisal or reimbursement decision?	The requesting body can state the decision the answer would inform	Retain only if another test is strongly met, otherwise consolidate
Evidence feasibility	Can the submitted evidence answer the question without extrapolation or comparison too fragile to survive national scrutiny?	Comparator arm, outcome and connected network exist, or a defensible adjustment is possible	Record explicitly as unanswerable, with reasons, and refer to evidence generation
Marginal information value	Does the question resolve an uncertainty not already resolved elsewhere in the scope?	No adjacent PICO shares the same evidence base and clinically similar comparator	Consolidate with the adjacent question
Marginal resource cost	Is the expected benefit proportionate to the additional analytical burden and its effect on timelines?	Estimated additional analyses and their effect on the assessment schedule	Sequence rather than exclude, or address in a subsequent update

PICO population, intervention, comparator and outcomes

The first is decision relevance. Would a credible answer plausibly change or materially clarify a national appraisal or reimbursement decision? A question a national body would treat as informative but not decision-relevant is a weaker candidate than one that determines a comparator choice or a restriction.

The second is evidence feasibility. Can the question be answered from the submitted evidence without extrapolation, unanchored comparison or subgroup analysis so fragile that the answer would not survive national scrutiny? This test can be applied at scoping, because the trial programme is known before the scope is fixed.

The third is marginal information value. Does the question resolve an uncertainty not already resolved elsewhere in the scope? Where two comparators are clinically close and the same body of evidence informs both, the second question adds cost without adding information.

The fourth is marginal resource cost. Is the expected benefit proportionate to the additional analytical burden on assessors and developer, including the effect on timelines?

The rule that follows is graduated rather than binary. Questions scoring highly on relevance and feasibility are prioritised. Questions with low incremental information value are candidates for consolidation with an adjacent question. Questions that are relevant but fall beyond the feasibility frontier are not silently dropped. They are recorded explicitly as unanswerable from the available evidence, with the reason stated, which converts a sunk analytical cost into a signal for future evidence generation.

This is not a mechanism for excluding Member States from the scope, and it does not sit in tension with the inclusiveness required by Article 8(6) [1]. Inclusiveness governs whose needs may be expressed. The tests govern how a fixed analytical budget is allocated across the questions that result, a distinct decision the current guidance leaves implicit. Nor is it a new procedural framework. It is a criterion for choices assessors and the Coordination Group already make.

7. Implications for the Design of Joint Clinical Assessment

Four design implications follow.

The first concerns scoping discipline. If the four tests are to bind, Member States submitting a PICO requirement should be asked to state briefly how the answer would be used in their national process. This is a low-cost reason-giving requirement that makes expected decision value visible at the point of request, and it is consistent with proposals to structure scoping through explicit consensus methods [17].

The second concerns evidence alignment. The feasibility frontier moves outward only when trial programmes anticipate national requirements. Joint scientific consultation

is the instrument for this, and its value should be judged by whether it reduces the proportion of PICOs that are unanswerable at assessment, not by uptake alone.

The third concerns transparency. Reports should distinguish clearly between questions answered with direct evidence, questions answered indirectly with stated uncertainty, and questions that could not be answered. The published reports already do this in substance [7, 8, 9]. Recording it in a consistent, extractable form would let the answerability of joint scopes be tracked over time and give national bodies, developers and regulators a shared picture of where the evidence base falls short.

The fourth concerns the limits of automation. Automated scoping tools reduce the cost of identifying and consolidating national requirements [18], which is a genuine saving. They do not reduce the cost of answering a question, and they cannot create a comparator arm, an outcome or a subgroup the trial did not include. A system that lowers the cost of generating questions without lowering the cost of answering them will, other things equal, generate more questions beyond the feasibility frontier. That strengthens rather than weakens the case for a value-based selection rule.

8. A Research Agenda

The framework is conceptual, and each of its quantities is measurable.

Resource use studies should record assessor, co-assessor and developer hours and external analytical spend per PICO and per JCA, giving the first empirical estimate of C_a. Answerability studies should classify PICOs in published scopes as directly answerable, indirectly answerable or unanswerable from the submitted evidence and track the proportions over time, which would locate the feasibility frontier empirically. Decision impact studies should follow specific PICO analyses into national appraisals and test whether they altered conclusions, restrictions or evidence requests, giving an observable proxy for V_d. Comparative efficiency studies should compare total joint plus national resource use against a counterfactual of separate national assessments, using the pre-Regulation record as a baseline [3]. Distributional studies should examine whether Member States differ systematically in requests made, benefits obtained and resources contributed. Where model inputs allow, formal value of information analysis applied to selected PICO uncertainties would test whether the informal prioritisation we propose approximates the formal answer [40, 41, 43].

9. Conclusion

The EU HTA Regulation asks a single joint assessment to serve twenty-seven decision contexts, and resolves the resulting tension by requiring the assessment scope to be inclusive. That is a defensible political settlement, and PICO multiplicity is its predictable consequence. The question this article raises is not whether multiple PICOs are legitimate. They are. It is how a joint system with finite analytical capacity and a finite evidence base should choose among them.

Framed economically, the answer is that an additional PICO should be judged by the decision value it is expected to generate relative to the resources it consumes and the probability that it can be answered at all. A joint HTA system that maximises the number of nationally relevant questions is not the same as one that maximises the value of the decisions those questions inform, and the two objectives diverge precisely where the evidence base is thinnest. Judging the system by decision value per unit of assessment resource, rather than by questions addressed, would give the Coordination Group, national bodies and developers a common and testable standard for whether joint assessment is delivering the efficiency the Regulation promised.

Statements and Declarations

Funding. Details of any funding are provided on the separate title page, which has been withheld to preserve anonymity during peer review.

Competing Interests

Declarations of competing interests for all authors are provided on the separate title page, which has been withheld to preserve anonymity during peer review.

Ethics Approval

Not applicable. This article is a conceptual analysis of published policy documents and peer-reviewed literature and involved no human participants, human data or animals.

Consent to participate. Not applicable.

Consent for publication. Not applicable.

Availability of Data and Material

All sources analysed in this article are publicly available and are cited in the reference list. The policy evidence reviewed has a cut-off date of 8 August 2026.

Code availability. Not applicable.

Author Contributions

Author contribution statements are provided on the separate title page, which has been withheld to preserve anonymity during peer review.

Use of Large Language Models

No generative artificial intelligence tool was used to produce the substantive content, analysis or conclusions of this article. Any use of such tools was limited to assistance with language and formatting, for which the authors take full responsibility.

Key Points for Decision Makers

- PICO multiplicity improves the national relevance of a joint clinical assessment, but each additional question also consumes assessor and developer capacity that could be used elsewhere, and some questions cannot be answered from the evidence that exists.
- PICO requests should be judged by their expected contribution to national decisions relative to their answerability and their marginal resource cost, rather than by national relevance alone.
- Evaluating the European joint assessment system requires measuring both the duplicated national assessment it avoids and the coordination and analytical burden it creates, and reporting the result as decision value per unit of assessment resource.

References

1. European Parliament and Council of the European Union. Regulation (EU) 2021/2282 of the European Parliament and of the Council of 15 December 2021 on health technology assessment and amending Directive 2011/24/EU. OJ L 458, 22.12.2021, p. 1. https://eur-lex.europa.eu/eli/reg/2021/2282/oj/eng. Accessed 8 Aug 2026.
2. Bikowski L. Startup ecosystems as catalysts of regional innovation and competitive advantage. International Journal of Innovation Studies. Advance online publication. https://doi.org/10.1016/j.ijis.2026.100209.
3. Hawlik K, Rummel P, Wild C. Analysis of duplication and timing of health technology assessments on medical devices in Europe. Int J Technol Assess Health Care. 2018;34(1):18-26. https://doi.org/10.1017/S0266462317001064.
4. Eberle K, Hagemann LM, Schweitzer MK, Justl M, Maurer J, Carls A, *et al*. The PICO puzzle: can public data predict EU HTA expectations for all EU countries? J Mark Access Health Policy. 2025;13(3):32. https://doi.org/10.3390/jmahp13030032.
5. Tiemens A, Sauboin C, Wolters S, Bond SJ, Bentvelzen JOM, Monteiro S, *et al*. Differences in evidentiary requirements for oncology drug assessments across health technology assessment bodies in Germany, Spain, and France, the European Medicines Agency, and the EU Health Technology Assessment Regulation. Expert Rev Pharmacoecon Outcomes Res. 2026;26(6):783-93. https://doi.org/10.1080/14737167.2026.2691190.
6. Sarri G, Vinals L, Leisle L, Chau IC, Smalbrugge D, Lucassen K, *et al*. Mapping methods gaps between EU joint clinical assessments and local health technology assessment decision-making: an environmental scan of guidance in select EU markets and harmonization challenges. J Comp Eff Res. 2025;14(5):e240240. https://doi.org/10.57264/cer-2024-0240.
7. Member State Coordination Group on Health Technology Assessment. Joint clinical assessment report on tovorafenib (Ojemda). Version 1.0. Brussels: European Commission, Directorate-General for Health and Food Safety; 2026. https://health.ec.europa.eu/publications/joint-clinical-assessment-report-tovorafenib-ojemda_en. Accessed 8 Aug 2026.
8. Member State Coordination Group on Health Technology Assessment. Joint clinical assessment report on tarlatamab (Imdylltra). Version 1.0. Brussels: European Commission, Directorate-General for Health and Food Safety; 2026. https://health.ec.europa.eu/publications/joint-clinical-assessment-report-tarlatamab-imdylltra_en. Accessed 8 Aug 2026.
9. Member State Coordination Group on Health Technology Assessment. Joint clinical assessment report on lurbinectedin (Zepzelca). Version 1.0. Brussels: European Commission, Directorate-General for Health and Food Safety; 2026. https://health.ec.europa.eu/publications/joint-clinical-assessment-report-lurbinectedin-zepzelca_en. Accessed 8 Aug 2026.
10. van Engen A, Krüger R, Parnaby A, Rotaru M, Ryan J, Samaha D, *et al*. The impact of additive population(s), intervention, comparator(s), and outcomes in a European joint clinical health technology assessment. Value Health. 2024;27(12):1722-31. https://doi.org/10.1016/j.jval.2024.07.024.
11. Starinieri B, Allocati E, Gerardi C, Joppi R, Banzi R. Comparing national health technology assessment approaches and the new European joint clinical assessment framework: the case of onco-hematological medicines. Eur J Clin Pharmacol. 2026;82(7):191. https://doi.org/10.1007/s00228-026-04123-6.
12. van Beekhuizen S, Che M, Monfort L, Hashim M, Azough A, Kubitz N, *et al*. Indirect treatment comparisons in EUnetHTA relative effectiveness assessments: learnings and recommendations for the implementation of EU joint clinical assessments. PharmacoEcon Open. 2025;9(4):597-609. https://doi.org/10.1007/s41669-025-00575-1.
13. Aballéa S, Toumi M, Wojciechowski P, Clay E, Falissard B, Simoens S, *et al*. Between rigor and relevance: why the EU HTA guidelines on indirect comparisons miss the mark. J Mark Access Health Policy. 2026;14(2):30. https://doi.org/10.3390/jmahp14020030.
14. Wang T, McAuslane N. Ensuring the efficiency and effectiveness of joint clinical assessment in national HTA decision-making: insights from the 2024 CIRS multi-stakeholder workshop. J Mark Access Health Policy. 2025;13(1):9. https://doi.org/10.3390/jmahp13010009.
15. Willemsen A, Rutten-van Mölken M, Al Dulaimi R, Schelleman H, Goettsch W, Timmers L. Preparing for the EU HTA Regulation: insights from the Dutch perspective. J Mark Access Health Policy. 2025;13(3):35. https://doi.org/10.3390/jmahp13030035.
16. Meregaglia M, Costa F, Cavallaro L, Armeni P, Otto MH. Implementing the EU HTA regulation and joint clinical assessment: a multi-stakeholder perspective from Italy. Int J Technol Assess Health Care. 2026;42(1):e39. https://doi.org/10.1017/S026646232610364X.
17. Arcà E, Barlassina A, Eze A, Strammiello V. Enhancing patient engagement in HTA: using consensus research to overcome PICO scoping challenges under the EU HTAR. J Mark Access Health Policy. 2025;13(2):27. https://doi.org/10.3390/jmahp13020027.
18. Boverhof BJ, Takatzoglou N, Redekop K, Uyl-de Groot CA, Visser JJ, Rutten-van Mölken MPMH. Development and validation of an artificial intelligence tool for automated population, intervention, comparator, outcome scoping for the European joint clinical assessment: a proof of concept. Value Health. 2026. https://doi.org/10.1016/j.jval.2026.05.009.
19. van Schoubroeck HM, Kroep S, Arcà E, Fenwick E. PD93 uncertainty metric to guide PICOS assessment and evidence synthesis planning for joint clinical assessment submission. Int J Technol Assess Health Care. 2024;40(S1):S131. https://doi.org/10.1017/S0266462324003404.
20. Dimiou S, Lisica A, Tanner S, Falla E. HTA221 life-cycle stage-specific prioritization of joint clinical assessment PICOs: balancing access potential,

feasibility and likelihood. Value Health. 2025;28(12):S448. https://doi.org/10.1016/j.jval.2025.09.2005.
21. Murphy P, Walker S, Putri S, Sculpher M, Griffin S. What is the value of developing an HTA process? PharmacoEconomics. 2026;44(7):791-8. https://doi.org/10.1007/s40273-026-01613-4.
22. Johannesen KM, Claxton K, Sculpher MJ, Wailoo AJ. How to design the cost-effectiveness appraisal process of new healthcare technologies to maximise population health: a conceptual framework. Health Econ. 2018;27(2):e41-e54. https://doi.org/10.1002/hec.3561.
23. Member State Coordination Group on Health Technology Assessment. Guidance on the scoping process. Version 1.0. Brussels: European Commission; 2024. https://health.ec.europa.eu/publications/guidance-scoping-process_en. Accessed 8 Aug 2026.
24. Julian E, Gianfrate F, Solà-Morales O, Mol P, Bergmann JF, Salmonson T, *et al*. How can a joint European health technology assessment provide an 'additional benefit' over the current standard of national assessments? Insights generated from a multi-stakeholder survey in hematology/oncology. Health Econ Rev. 2022;12(1):30. https://doi.org/10.1186/s13561-022-00379-7.
25. Jommi C, Bonfanti M, Guardigni M, Aiello A, Marcellusi A, Canonico PL, *et al*. Pivotal studies for drugs about to be launched for rare diseases: will they better support health technology assessment and market access than in the past? J Mark Access Health Policy. 2025;13(3):37. https://doi.org/10.3390/jmahp13030037.
26. Kloc K, Toumi M, Łukomska E, Kowalska M, Tyrała-Chowaniec I, Simoens S, *et al*. Joint clinical assessment in the EU HTA Regulation: would drugs supported by single-arm trials fit under evaluation? J Mark Access Health Policy. 2026;14(2):36. https://doi.org/10.3390/jmahp14020036.
27. Xander NSH, Rutten-van Mölken MPMH, Bertin L, Miljković N, Uyl-de Groot CA. Legal implications of joint clinical assessments under the EU HTA regulation. Health Econ Policy Law. 2026;1-13. https://doi.org/10.1017/S1744133126100413.
28. Heikkinen I, Goodall M, Steck N, Poulakou M, Piso K. Impact of indication changes on scoping for European Union joint clinical assessment: scale of the problem and how to address it. Int J Technol Assess Health Care. 2024;40(1):e62. https://doi.org/10.1017/S0266462324004641.
29. D'Oca K, Darquennes E, Garrigues C, Draganigos A, Steck N. European joint clinical assessment PICO scoping process: analysis of current approaches and recommendations. J Mark Access Health Policy. 2026;14(1):3. https://doi.org/10.3390/jmahp14010003.
30. van Engen A, Amoura L, Curtis O, Falla E, Ibok A, Kotseva F, *et al*. HTA6 predicting PICOs to prepare for joint clinical assessment: learnings from 19 case studies. Value Health. 2024;27(12):S353. https://doi.org/10.1016/j.jval.2024.10.1831.
31. Dimiou S, Falla E, Tanner S, Amoura L, Curtis O, Faivre P, *et al*. HTA300 simulating PICOs to predict evidence requirement implications of the EU joint clinical assessment: learnings from 24 case studies. Value Health. 2025;28(12):S464. https://doi.org/10.1016/j.jval.2025.09.2083.
32. Member State Coordination Group on Health Technology Assessment. Methodological guideline for quantitative evidence synthesis: direct and indirect comparisons. Brussels: European Commission; 2024. https://health.ec.europa.eu/publications/methodological-guideline-quantitative-evidence-synthesis-direct-and-indirect-comparisons_en. Accessed 8 Aug 2026.
33. Goetz G, Schandelmaier S, Busse R, Wild C, Panteli D. Implementation of the EU's Health Technology Assessment regulation: where does existing methods guidance require concretization and what are the relevant methodological options? Int J Technol Assess Health Care. 2025;41(1):e9. https://doi.org/10.1017/S0266462324004793.
34. Kranz P, McGauran N, Ünal C, Kaiser T. Results of health technology assessments of orphan drugs in Germany: lack of added benefit, evidence gaps, and persisting unmet medical needs. Int J Technol Assess Health Care. 2024;40(1):e68. https://doi.org/10.1017/S026646232400062X.
35. Desmet T, Brijs M, Vanderdonck F, Tops S, Simoens S, Huys I. Implementing the EU HTA regulation: insights from semi-structured interviews on patient expectations, Belgian and European institutional perspectives, and industry outlooks. Front Pharmacol. 2024;15:1369508. https://doi.org/10.3389/fphar.2024.1369508.
36. Member State Coordination Group on Health Technology Assessment. Annual work programme 2026. Brussels: European Commission; 2025. https://health.ec.europa.eu/publications/2026-work-programme-member-state-coordination-group-hta-htacg_en. Accessed 8 Aug 2026.
37. Claxton K. The irrelevance of inference: a decision-making approach to the stochastic evaluation of health care technologies. J Health Econ. 1999;18(3):341-64. https://doi.org/10.1016/S0167-6296(98)00039-3.
38. Sculpher M, Claxton K. Establishing the cost-effectiveness of new pharmaceuticals under conditions of uncertainty: when is there sufficient evidence? Value Health. 2005;8(4):433-46. https://doi.org/10.1111/j.1524-4733.2005.00033.x.
39. Claxton KP, Sculpher MJ. Using value of information analysis to prioritise health research: some lessons from recent UK experience. PharmacoEconomics. 2006;24(11):1055-68. https://doi.org/10.2165/00019053-200624110-00003.
40. Fenwick E, Steuten L, Knies S, Ghabri S, Basu A, Murray JF, *et al*. Value of information analysis for research decisions: an introduction. Report 1 of the ISPOR Value of Information Analysis Emerging Good Practices Task Force. Value Health. 2020;23(2):139-50. https://doi.org/10.1016/j.jval.2020.01.001.
41. Rothery C, Strong M, Koffijberg H, Basu A, Ghabri S, Knies S, *et al*. Value of information analytical methods. Report 2 of the ISPOR Value of Information Analysis Emerging Good Practices Task Force. Value Health. 2020;23(3):277-86. https://doi.org/10.1016/j.jval.2020.01.004.
42. Kunst N, Wilson ECF, Glynn D, Alarid-Escudero F, Baio G, Brennan A, *et al*. Computing the expected value of sample information efficiently: practical

guidance and recommendations for four model-based methods. Value Health. 2020;23(6):734-42. https://doi.org/10.1016/j.jval.2020.02.010.
43. Jackson CH, Baio G, Heath A, Strong M, Welton NJ, Wilson ECF. Value of information analysis in models to inform health policy. Annu Rev Stat Appl. 2022;9(1):95-118. https://doi.org/10.1146/annurev-statistics-040120-010730.
44. Bennette CS, Veenstra DL, Basu A, Baker LH, Ramsey SD, Carlson JJ. Development and evaluation of an approach to using value of information analyses for real-time prioritization decisions within SWOG, a large cancer clinical trials cooperative group. Med Decis Making. 2016;36(5):641-51. https://doi.org/10.1177/0272989X16636847.
45. Borah R, Brown AW, Capers PL, Kaiser KA. Analysis of the time and workers needed to conduct systematic reviews of medical interventions using data from the PROSPERO registry. BMJ Open. 2017;7(2):e012545. https://doi.org/10.1136/bmjopen-2016-012545.
46. Eckermann S, Willan AR. The option value of delay in health technology assessment. Med Decis Making. 2008;28(3):300-5. https://doi.org/10.1177/0272989X07312477.
47. Claxton K, Martin S, Soares M, Rice N, Spackman E, Hinde S, *et al*. Methods for the estimation of the National Institute for Health and Care Excellence cost-effectiveness threshold. Health Technol Assess. 2015;19(14):1-504. https://doi.org/10.3310/hta19140.
48. Lomas J, Martin S, Claxton K. Estimating the marginal productivity of the English National Health Service from 2003 to 2012. Value Health. 2019;22(9):995-1002. https://doi.org/10.1016/j.jval.2019.04.1926.
49. Vallejo-Torres L, García-Lorenzo B, Serrano-Aguilar P. Estimating a cost-effectiveness threshold for the Spanish NHS. Health Econ. 2018;27(4):746-61. https://doi.org/10.1002/hec.3633.
50. Culyer AJ. Cost-effectiveness thresholds in health care: a bookshelf guide to their meaning and use. Health Econ Policy Law. 2016;11(4):415-32. https://doi.org/10.1017/S1744133116000049.
51. Claxton K, Sculpher M, Palmer S, Culyer AJ. Causes for concern: is NICE failing to uphold its responsibilities to all NHS patients? Health Econ. 2015;24(1):1-7. https://doi.org/10.1002/hec.3130.
52. Mitton C, Donaldson C. Twenty-five years of programme budgeting and marginal analysis in the health sector, 1974-1999. J Health Serv Res Policy. 2001;6(4):239-48. https://doi.org/10.1258/1355819011927558.
53. Griffin SC, Claxton KP, Palmer SJ, Sculpher MJ. Dangerous omissions: the consequences of ignoring decision uncertainty. Health Econ. 2011;20(2):212-24. https://doi.org/10.1002/hec.1586.
54. Coase RH. The nature of the firm. Economica. 1937;4(16):386-405. https://doi.org/10.1111/j.1468-0335.1937.tb00002.x.
55. Williamson OE. The economics of organization: the transaction cost approach. Am J Sociol. 1981;87(3):548-77. https://doi.org/10.1086/227496.
56. Vogler S, Knoll V, Sonemann A, Quanten A, Adams R, de Jager T, *et al*. Collaborating for affordable patient access to medicines: review of the first decade of the Beneluxa Initiative in pharmaceutical policy. Appl Health Econ Health Policy. 2026. https://doi.org/10.1007/s40258-026-01057-z.
57. Ostrom E. Governing the commons: the evolution of institutions for collective action. Cambridge: Cambridge University Press; 1990. https://doi.org/10.1017/CBO9780511807763.
58. Buchanan JM. An economic theory of clubs. Economica. 1965;32(125):1-14. https://doi.org/10.2307/2552442.
59. Allen N, Liberti L, Walker SR, Salek S. A comparison of reimbursement recommendations by European HTA agencies: is there opportunity for further alignment? Front Pharmacol. 2017;8:384. https://doi.org/10.3389/fphar.2017.00384.
60. Nicod E, Maynou L, Visintin E, Cairns J. Why do health technology assessment drug reimbursement recommendations differ between countries? A parallel convergent mixed methods study. Health Econ Policy Law. 2020;15(3):386-402. https://doi.org/10.1017/S1744133119000239.
61. Frischhut M, Gamba S, Magazzini L, Pertile P. The price of national pharmaceutical pricing in the EU: a proposal for 'common procurement'. Appl Health Econ Health Policy. 2026. https://doi.org/10.1007/s40258-026-01054-2.
62. Johns A, Andrade A, Johal S, Ryan J. EU joint clinical assessment: a framework for optimising use with cost-effectiveness decision-making. J Mark Access Health Policy. 2025;13(4):52. https://doi.org/10.3390/jmahp13040052.