Annals of Pharmacy Practice and Pharmacotherapy
ISSN: 3062-4436
2024, Volume 4, Issue 2, Page No: 114-124

Available online at: www.galaxypub.co/page/journals

Galaxy Publication

Therapeutic Success Is Misclassified When Biomarkers Improve but the Medication Work Required from Patients Becomes Unsustainable

Nina Larsen[1]*, Ahmed Mansour[2], Julia Fischer[3]

[1]Department of Clinical Pharmacy and Patient-Reported Outcomes, Faculty of Health Sciences, Copenhagen Business School, Copenhagen, Denmark.
[2]Department of Pharmacy Practice and Treatment Burden, Faculty of Pharmacy, Alexandria University, Alexandria, Egypt.
[3]Department of Pharmaceutical Care and Health Outcomes, Faculty of Pharmacy, University of Zurich, Zurich, Switzerland.

***E-mail** ✉ nina.larsen@cbs.dk

Received: 12 August 2024; **Revised:** 03 November 2024; **Accepted:** 06 November 2024

ABSTRACT

To develop a methodological approach for classifying treatment outcomes when biological or clinical markers improve while the work required to sustain treatment becomes difficult, disruptive, or unacceptable to patients. A construct-explicit methodological synthesis was undertaken across literature addressing treatment burden, patient work, patient-reported outcomes, medication adherence and persistence, surrogate endpoints, functional outcomes, patient preferences, and value assessment. Biological response, symptom or functional response, treatment workload, experienced treatment burden, persistence feasibility, and patient-valued benefit were treated as analytically distinct constructs. Rules were specified for recognizing discordance among these domains and for preventing adherence, biomarker change, or treatment-burden scores from being used as interchangeable proxies. No patient-level dataset was created, no empirical threshold was estimated, and no composite score was derived. Biological improvement can coexist with high treatment workload, reduced capacity to sustain treatment, limited functional improvement, or unfavorable patient evaluation. Conversely, demanding treatment can remain sustainable when its benefits are substantial, the workload is bounded or supported, and the patient considers the trade-off worthwhile. Observed adherence or persistence does not establish low treatment burden, and nonpersistence does not establish that burden caused discontinuation. Biomarker improvement represents biological response unless direct patient-important evidence or an adequately justified surrogate relationship supports broader inference. Sustainable treatment success is therefore better treated as a multidomain interpretation than as a synonym for favorable biological response. Health-outcome assessment can misclassify treatment success when biological improvement is recorded without examining the work required to produce and maintain it. Biological response, treatment-work sustainability, patient-important outcomes, and persistence should be measured separately before preference-sensitive weighting or value assessment. The proposed approach is a testable methodological specification rather than a validated scoring instrument.

Keywords: Treatment burden, Patient work, Biomarkers, Patient-reported outcomes, Medication persistence, Value assessment

How to Cite This Article: Larsen N, Mansour A, Fischer J. Therapeutic Success Is Misclassified When Biomarkers Improve but the Medication Work Required from Patients Becomes Unsustainable. Ann Pharm Pract Pharmacother. 2024;4(2):114-24. https://doi.org/10.51847/xB0VMtRha0

Introduction

Modern treatment can improve disease markers while simultaneously increasing the amount of work required from patients. Medication preparation, administration, dose timing, monitoring, laboratory testing, appointments, refill management, dietary coordination, adverse-effect surveillance, device use, and adaptation of ordinary routines may all be necessary to produce a favorable biological response. Treatment burden therefore extends

beyond medication count or regimen complexity. An integrative review of multimorbidity research characterizes burden as a multidimensional experience involving practical, emotional, social, and financial consequences of managing healthcare [1]. Its magnitude also depends on whether the patient possesses sufficient capacity to carry out the required work.

This interaction between workload and capacity is central to minimally disruptive medicine. Patients with similar prescriptions may experience very different levels of burden because health status, cognition, finances, transportation, social support, competing responsibilities, and familiarity with treatment differ. The workload–capacity formulation therefore treats burden as relational rather than as a fixed property of a regimen [2]. A treatment that is feasible during a highly supported trial or acute treatment phase may become difficult to sustain in ordinary life, while an intensive regimen may remain acceptable when its expected benefit is large and the required period of effort is limited.

Clinical evaluation nevertheless often privileges disease control, laboratory results, imaging, physiological measurements, or other markers of biological response. These measures are frequently indispensable, but they answer a different question from whether treatment produces outcomes that matter directly to patients or can be sustained over time. Methodological work on surrogate endpoints emphasizes that a surrogate represents a target outcome only when the relationship is adequately justified in the relevant context [3]. Biological response therefore cannot automatically be promoted into evidence of complete therapeutic success.

The definition of an appropriate treatment outcome also depends on which perspectives participate in outcome selection. Patient involvement in core outcome set development has been associated with greater inclusion of life-impact domains [4]. This finding is consequential because a measurement system that routinely records biological control while omitting function, disruption, treatment work, or patient priorities can make treatment appear more successful simply because some dimensions have not been measured.

Health-economic and value-assessment methods already recognize that the value of treatment is multidimensional. The ISPOR Special Task Force described elements of value extending beyond conventional measures of health gain and cost [5]. Patient-preference methods similarly emphasize that attributes should be selected and interpreted in relation to the decision, population, and trade-offs under consideration [6]. These approaches create space for patient-relevant treatment characteristics, but they do not eliminate the need to distinguish the underlying constructs correctly before weighting them.

A particular classification problem arises when a treatment produces favorable biological change while the work required to obtain that change becomes unsustainable. Treating this state simply as “successful therapy” ignores feasibility. Calling it “treatment failure” also loses important information because biological activity may be substantial. A method is therefore needed that preserves the discordance itself.

This article addresses the question: How should health-outcome assessment classify treatment when biological markers improve while medication workload, monitoring demands, disruption, or treatment burden become unsustainable?

The methodological contribution is a distinction between biological response and sustainable treatment success that keeps biological or clinical response, symptom or functional response, treatment workload, experienced treatment burden, persistence feasibility, and patient-valued benefit analytically separate. The objective is not to create a universal burden threshold or additive score. It is to prevent an identifiable classification error: interpreting evidence of biological response as sufficient evidence that treatment is successful in a form that can be maintained and remains worthwhile to the patient.

Materials and Methods

Construct boundaries and unit of classification

The unit of classification is a treatment state observed over a clinically relevant time horizon, rather than a single laboratory result, adherence percentage, or burden score. Six constructs are distinguished.

Biological or clinical response refers to change in a disease-related physiological, laboratory, imaging, or clinical variable. The evidentiary meaning of that variable depends on whether it is itself patient-important or is being used as a surrogate.

Symptom or functional response refers to changes experienced by the patient or demonstrated through relevant functional performance. This domain is not assumed to follow biological response automatically.

Treatment workload refers to the tasks generated by treatment, including administration, monitoring, appointments, coordination, logistics, self-management, and related requirements.
Experienced treatment burden refers to the practical, emotional, social, financial, and daily-life consequences of carrying out that work.
Persistence feasibility refers to whether the required treatment can plausibly be continued over the horizon necessary to realize its intended benefit, given patient capacity, treatment demands, support, adverse effects, access, and competing responsibilities.
Patient-valued benefit refers to whether the resulting combination of effects and demands is considered worthwhile by the patient in the relevant decision context.
Treatment-burden measurement confirms the importance of keeping these constructs explicit. A scoping review of measures in multimorbidity identified multiple instruments with differences in content and psychometric evidence [7]. The Patient Experience with Treatment and Self-management measure captures several dimensions of treatment and self-management burden [8]. The Multimorbidity Treatment Burden Questionnaire provides a shorter validated instrument for multimorbidity [9], while brief treatment-burden measures demonstrate that lower measurement workload is also possible [10].
These instruments provide useful observations but do not establish a universal boundary between sustainable and unsustainable treatment. A high burden score may be accepted because treatment benefit is substantial. A moderate score may become unsustainable in a patient with declining capacity. Sustainability therefore cannot be assigned solely from an instrument threshold.
Medication adherence and persistence were also kept separate from treatment burden. The ESPACOMP Medication Adherence Reporting Guideline emphasizes distinctions among initiation, implementation, and persistence and the importance of defining how these behaviors are measured [11]. A systematic review of adherence assessment in healthcare databases for chronic obstructive pulmonary disease found substantial methodological heterogeneity in how implementation and persistence were operationalized [12].
Observed nonadherence may reflect excessive workload, but other explanations include adverse effects, financial barriers, access problems, intentional choice, beliefs about treatment, misunderstanding, treatment complexity, or measurement error. Adherence behavior is consequently treated as evidence requiring interpretation rather than as a direct measure of burden.

Evidence-linked classification procedure

Classification begins by recording each domain independently. Biological or clinical response is first characterized using disease-appropriate measures. Symptom or functional response is then evaluated using instruments that address the relevant patient experience or functional capability. Treatment workload and experienced burden are assessed directly rather than inferred from medication count. Persistence feasibility is interpreted over the horizon required for treatment benefit. Patient-valued benefit is assessed as a preference-sensitive judgment after these domains are described.
Measurement method is itself part of interpretation. Updated ISPOR guidance on patient-reported outcome measurement recommends examining whether collection-mode changes preserve comparability when modifications could affect response or instrument properties [13]. Functional outcomes require different logic. Recommendations for performance outcome assessments distinguish standardized demonstration of functional ability from patient-reported symptoms or lived experience [14]. Biomarkers, patient-reported outcomes, performance measures, burden instruments, and adherence records are therefore not treated as substitutes.
Biomarkers require an additional evidentiary step when used as surrogates. A scoping review of surrogate-endpoint definitions found heterogeneous terminology and guidance and reinforced the need for explicit justification of surrogate use [15]. Statistical methods for surrogate validation likewise evaluate whether effects on a surrogate predict effects on a target outcome across appropriate trials or datasets [16].
Accordingly, biomarker improvement is initially classified as biological response. It becomes evidence of broader patient benefit only when the marker is directly patient-important or when an adequately validated relationship supports that inference.
The procedure intentionally avoids summing all domains into one score. A treatment may produce strong biological response and poor sustainability, while another may be sustainable but clinically insufficient. Collapsing these states into a common average could obscure their differences.

Measurement discordance and disconfirmation rules

Discordance is defined as a clinically meaningful difference in direction or interpretation among outcome domains. Relevant patterns include biological improvement with worsening treatment burden, biomarker improvement without functional benefit, preserved disease control with deteriorating persistence feasibility, and high treatment workload that remains acceptable because benefit is substantial and the workload is time limited.

These states require context rather than automatic ranking. The interpretation depends on treatment duration, reversibility, available alternatives, disease severity, patient priorities, support, and consequences of discontinuation.

Three disconfirmation rules constrain the proposed method.

First, poor persistence cannot be attributed to treatment burden solely because both occur. Alternative explanations must be evaluated.

Second, biomarker improvement cannot be assumed to demonstrate patient-important benefit unless the biomarker is itself meaningful to patients or has adequate evidentiary support as a surrogate.

Third, treatment-work sustainability is methodologically useful only if its explicit assessment changes classification, interpretation, monitoring, treatment design, or decision making beyond information already available from established outcome measures.

No universal numerical threshold for sustainability is proposed. Such thresholds would require longitudinal validation and would likely vary by disease, regimen, treatment horizon, patient capacity, and preference.

Results and Discussion

Discordant response states across outcome dimensions

Evidence supports the proposition that treatment burden and patient-important outcomes can differ from biological or conventional clinical indicators, although the available literature does not support a universal causal pathway from treatment burden to adverse clinical outcomes.

In patients with multimorbidity, higher treatment burden has been associated with poorer health-related quality of life [17]. A UK primary-care study similarly found that burden was associated with factors including lower health literacy, financial difficulty, increasing numbers of long-term conditions, and medication complexity [18]. These relationships support direct measurement of treatment burden and patient capacity but do not establish that burden independently caused worse health.

Heart failure provides an important limiting case. Greater treatment burden has been associated with worse health status, but relationships with hospitalization and mortality were attenuated after adjustment for relevant characteristics [19]. Treatment burden should therefore not be converted into an independent prognostic surrogate without stronger evidence. Its importance lies in capturing an aspect of the treatment experience that biological measurements and conventional hard outcomes may not represent.

Cystic fibrosis illustrates how therapeutic progress can create new discordance. In the era of highly effective CFTR modulators, patients and clinicians have expressed substantial interest in reducing residual treatment workload [20]. Major biological improvement can therefore coexist with continued work from inhaled treatments, airway clearance, monitoring, and treatment organization.

This observation does not establish that residual therapies can safely be stopped. It establishes that biological effectiveness and treatment workload remain separate questions. The clinically relevant issue becomes whether disease control can be preserved while unnecessary or poorly tolerated work is removed.

These distinctions generate several recognizable outcome configurations.

Table 1. Outcome configurations distinguishing biological response from sustainable treatment success

Outcome state	Observed treatment consequences			Sustainability judgment		Classification
Outcome configuration	**Biological or clinical response**	**Symptom or functional response**	**Treatment workload / experienced burden**	**Persistence feasibility**	**Patient-valued benefit**	**Methodological interpretation**
Convergent sustainable benefit	Improved or adequately controlled	Improved, preserved, or acceptable	Manageable in context	Feasible over required horizon	Favorable	Candidate sustainable treatment success

Biomarker-led discordance	Improved	Minimal, absent, or worse	High, disruptive, or escalating	Fragile or uncertain	Mixed or unfavorable	Biological response present; sustainable success not established
High-workload but accepted benefit	Improved	Improved or preserved	High but expected or accepted	Feasible with sufficient capacity/support	Favorable	High workload alone does not invalidate success
Sustainable but therapeutically insufficient state	Absent, inadequate, or worsening	Unchanged or worsening	Low or manageable	Feasible	Limited	Sustainability cannot compensate for inadequate therapeutic effect
Time-limited intensive treatment	Improved or expected to improve	Variable by treatment phase	High for a bounded period	Feasible for defined horizon	Favorable if trade-off accepted	Treatment horizon alters interpretation of workload
Persistence failure of uncertain cause	Variable	Variable	Known or unknown	Observed persistence poor	Unknown or mixed	Cause must be established before burden attribution

Note: These configurations are proposed methodological classifications. They are not validated clinical states, numerical scoring bands, or treatment recommendations.

Persistence feasibility changes the meaning of observed response

Persistence feasibility becomes important when continued exposure is necessary to maintain the biological effect. A regimen may achieve excellent laboratory or physiological results during intensive follow-up yet prove difficult to sustain once monitoring, travel, coordination, expense, adverse effects, or competing responsibilities accumulate. Conversely, demanding treatment may remain feasible when it is temporary, highly effective, adequately supported, and consistent with patient priorities.

Randomized evidence in cystic fibrosis demonstrates that treatment-work reduction can be tested empirically. In the SIMPLIFY trials, selected people receiving elexacaftor/tezacaftor/ivacaftor with relatively preserved lung function were randomized to discontinue or continue hypertonic saline or dornase alfa. Over six weeks, discontinuation was noninferior for change in percent-predicted FEV1 under the studied conditions [21].

This result does not justify generalized withdrawal of chronic therapy. It demonstrates a methodological principle: workload and biological response can sometimes be experimentally separated. If work is reduced while biological response remains stable, treatment sustainability may improve without a corresponding change on the biological axis.

Measurement remains complex. A systematic review of treatment-burden measurement in cystic fibrosis found that no single subjective or objective method captures all relevant dimensions [22]. Qualitative work after introduction of highly effective modulator therapy similarly suggests that treatment burden remains individualized and can persist or change form despite major clinical improvement [23].

Persistence feasibility should therefore be treated as a longitudinal property rather than a fixed attribute. It may improve after simplification, formulation changes, increased support, improved logistics, treatment familiarization, or recovery of patient capacity. It may deteriorate despite stable biological response if monitoring intensifies, adverse effects accumulate, treatment work increases, financial circumstances change, or competing obligations reduce capacity.

Because these changes can precede deterioration in biomarkers, outcome assessment that recognizes only biological response may continue to label treatment as fully successful after its real-world sustainability has begun to fail.

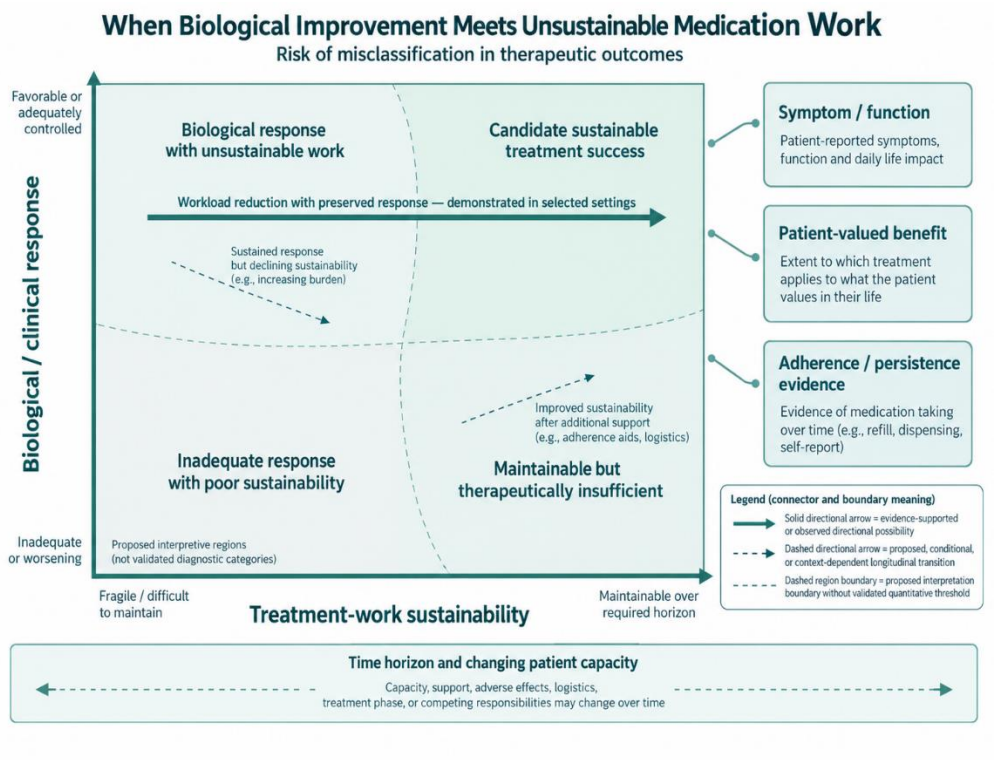

Figure 1. When biological response and treatment-work sustainability diverge

Measurement systems capture different treatment realities

Different measurement systems observe different components of treatment. A biomarker may provide a precise indication of pharmacological or biological activity while providing little information about the amount of work required from the patient. A treatment-burden questionnaire may detect disruption and organizational strain without identifying whether disease control is adequate. Dispensing records can show whether medication is obtained but may not explain motivation, ingestion, burden, or clinical response.

Patient-reported information can materially alter the clinical picture. In advanced cancer care, electronic symptom monitoring generated information associated with clinically consequential differences in management and outcomes [24]. The implication is not that patient-reported outcomes should replace biomedical monitoring. It is that different observation systems contain nonredundant information.

Core outcome set research demonstrates the same principle at a methodological level. A systematic review of outcome sets developed for routine care identified substantial variation in outcome selection and patient involvement [25]. The architecture of measurement determines which dimensions of treatment become visible.

Construct overlap should therefore be analyzed without assuming equivalence. A patient experiencing demanding medication work may later exhibit poor persistence. These observations may be related, but treatment burden describes the experienced demands of care, whereas persistence describes behavior over time. Poorer function may reduce capacity to perform treatment work, while treatment work may itself interfere with daily function. Cross-sectional concordance cannot establish the direction of these relationships.

Table 2. Measurement sources for biological response, treatment burden, persistence, and patient-important outcomes

Evidence source	Construct and timing		Direct reach and blind spot		Interpretation	
Measurement source	**Primary construct observed**	**Typical temporal form**	**What it establishes directly**	**Important blind spot**	**Main misclassification risk**	**Role in sustainable-success assessment**
Biomarker or physiological measure	Biological or clinical response	Single or repeated clinical measurement	State or change in measured biological variable	Treatment workload, acceptability, lived benefit	Biological change treated as complete success	Establish biological-response component

Symptom patient-reported outcome	Experienced symptoms	Repeated patient report	Patient-perceived symptom change	Objective mechanism and workload	Symptom improvement assumed to imply sustainability	Characterize patient-experienced clinical effect
Functional or performance assessment	Capability to perform relevant activities	Episodic or longitudinal	Functional performance under specified conditions	Treatment organization and subjective burden	Performance treated as equivalent to daily-life experience	Test translation of response into usable function
Treatment-burden instrument	Experienced healthcare demands and consequences	Cross-sectional or repeated	Patient-perceived burden	Exact causal source and biological effect	Burden score treated as treatment failure	Directly observe sustainability-related experience
Workload/task assessment	Amount and organization of treatment work	Daily, weekly, regimen-specific	Tasks, monitoring, coordination, and time demands	Whether workload is acceptable	Intensive treatment assumed unsustainable	Describe work generated by treatment
Dispensing or administration record	Medication implementation/persistence	Longitudinal	Observable acquisition or administration pattern	Motivation, burden, ingestion, preference	Nonpersistence treated as proof of burden	Identify behavioral signal requiring explanation
Preference or value elicitation	Importance assigned to attributes/trade-offs	Decision-specific	Patient valuation of defined benefits and burdens	Objective effectiveness	Preference treated as clinical outcome	Support preference-sensitive interpretation
Combined longitudinal assessment	Relationship among outcome domains	Repeated over treatment horizon	Whether domains converge, diverge, or change	Greater measurement complexity	Correlated outcomes double counted	Validate sustainable-success classification

Measurement should remain parsimonious. Adding every available instrument could itself increase workload. Domains should be measured when they can plausibly change interpretation, treatment design, or value assessment.

Competing explanations and boundary cases

High treatment workload does not automatically mean treatment has failed. Burden depends partly on patient capacity and context, including health status, finances, literacy, comorbidity, treatment complexity, and available support [18]. An intensive regimen may remain sustainable when the treatment horizon is short, the expected benefit is substantial, and sufficient support exists.

Reverse causation is also plausible. Poor health can increase the difficulty of carrying out treatment, producing higher reported burden even if the regimen itself has not changed. The attenuation of relationships between burden and hard outcomes after adjustment in heart failure demonstrates the importance of this possibility [19].

Nonpersistence has similarly heterogeneous causes. Cost, adverse effects, access limitations, inconvenience, treatment beliefs, competing responsibilities, perceived lack of benefit, and deliberate preference-sensitive decisions can all produce discontinuation. Persistence data should therefore initiate causal inquiry rather than terminate it.

Cystic fibrosis illustrates why neither high workload nor simplification should be interpreted automatically. Patients and clinicians have prioritized investigation of treatment simplification [20], and randomized trials have shown that selected therapies can be withdrawn over a defined period without meaningful short-term loss of pulmonary-function response in selected modulator-treated patients [21]. Yet measurement reviews continue to identify limitations in the coverage of treatment-burden instruments [22], and qualitative evidence demonstrates substantial heterogeneity in the experience of treatment after major therapeutic improvement [23].

The upper-left state in **Figure 1** should consequently not be interpreted as an automatic recommendation to discontinue therapy. It signals unresolved discordance. Additional information is required about the cause of burden, alternative treatments, reversibility, expected consequences of modification, patient priorities, and the time horizon over which the current workload must continue.

Similarly, adherence does not guarantee sustainability. Some patients maintain treatment through exceptional effort or substantial sacrifice. Apparent persistence can therefore coexist with a treatment state that is progressively eroding capacity.

Consequences for value assessment and comparative effectiveness

The principal implication is that evidence of therapeutic effect should be distinguished from evidence that the effect can be produced and maintained under conditions acceptable to the patient. Biological efficacy remains necessary in many therapeutic contexts. The classification problem emerges when biological response is allowed to define the entire treatment state.

This distinction is consistent with broader health-economic approaches. Value-assessment methods recognize that health technologies may affect dimensions extending beyond narrowly specified health gain and direct cost [5]. Patient-preference research likewise shows that the importance of benefits, risks, convenience, and related treatment attributes depends on the relevant population and decision [6].

The present proposal operates before these attributes are aggregated. The domains must first be observed accurately. A preference weight applied to a poorly defined or duplicated construct cannot repair the original measurement problem.

Formal quantitative benefit–risk methods emphasize careful attribute selection, weighting, dependence, and avoidance of double counting [26]. These requirements are especially relevant here because treatment burden, adherence, function, quality of life, and patient preference can overlap. An additive score that assigns separate points to highly correlated consequences could produce misleading precision.

The proposed approach therefore preserves the domains until their relationships are known. Biological response, treatment work, symptoms or function, persistence feasibility, and patient-valued benefit may subsequently inform benefit–risk assessment, comparative effectiveness, or health technology assessment, but they should not be collapsed merely because a single summary measure is convenient.

Stakeholder research suggests continued interest in value elements that extend beyond conventional evaluation [27]. Patient-centered health technology assessment also emphasizes integration of patient and caregiver perspectives into assessment processes [28]. Neither literature provides a validated universal value for treatment workload, and the present method does not attempt to create one.

Comparative effectiveness provides a particularly useful application. Two treatments may produce similar biological effects but impose different administration schedules, monitoring requirements, device use, travel, preparation, or disruption. Those differences may materially alter persistence and patient preference. Conversely, a more demanding treatment may remain preferred when it produces a larger or more important clinical benefit.

Sustainable treatment success therefore does not mean minimum workload. It means that the work required for treatment remains feasible and worthwhile relative to the benefit achieved.

Validation agenda, implementation requirements, and limits

Validation requires longitudinal measurement because sustainability is dynamic. A patient may initially maintain intensive therapy successfully and later lose capacity as treatment tasks accumulate, health deteriorates, support changes, or competing responsibilities increase. Biological measures can remain stable during this period, producing a clinically important interval in which conventional monitoring continues to report success while the feasibility of maintaining treatment is deteriorating.

The opposite trajectory is equally important. Treatment workload may fall after simplification, formulation changes, home monitoring, logistical support, familiarization, or recovery from an acute treatment phase. Biological response may remain unchanged while the overall treatment state becomes easier to maintain.

Interventional studies offer particularly strong tests. If a treatment modification reduces patient work while preserving biological effect, a valid sustainability dimension should detect improvement without requiring movement in the biological outcome. The SIMPLIFY trials provide one example of this logic in a narrowly selected clinical context [21].

A longitudinal validation design should therefore measure several trajectories together rather than assuming that one represents the others. Repeated biological or clinical measures should be accompanied, where relevant, by direct assessment of treatment workload, experienced burden, symptom or functional outcomes, persistence behavior, and patient valuation. The purpose is not maximal measurement density; it is to observe whether clinically meaningful divergence develops.

This longitudinal logic is summarized in **Figure 2**.

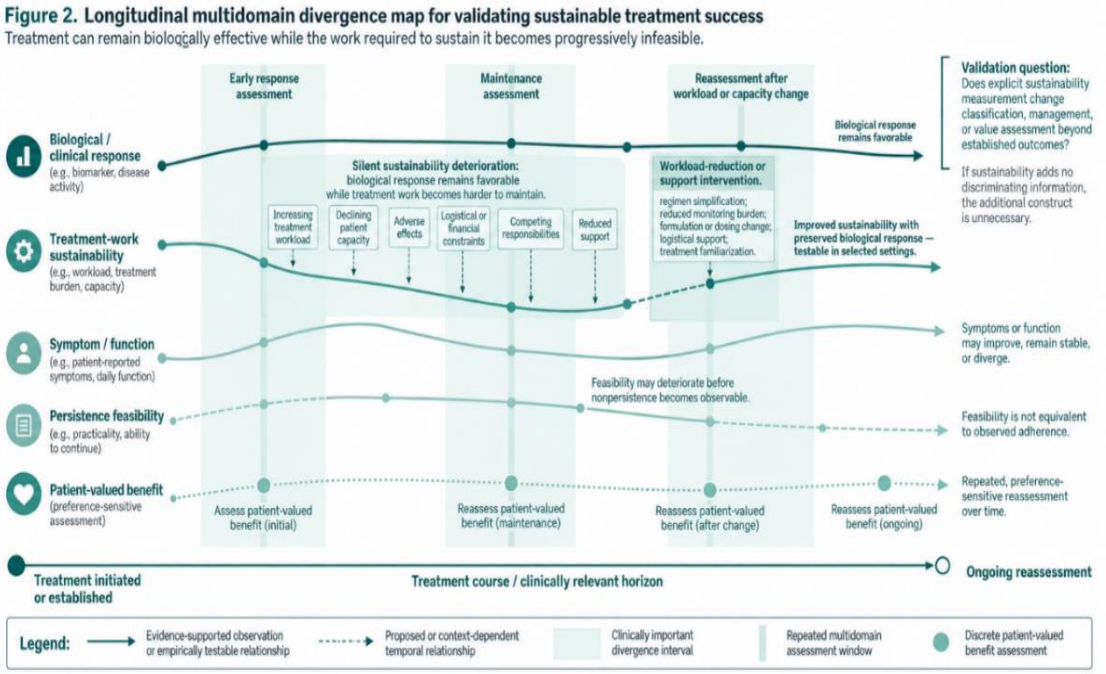

Figure 2. Testing whether biological benefit remains sustainable over time

Conclusion

Therapeutic success cannot always be inferred from biological improvement alone. A treatment may produce a favorable biomarker response while requiring work that becomes difficult to maintain, provides limited patient-important benefit, or places future persistence at risk. The reverse is also possible: demanding treatment may remain successful when its benefits are substantial, its duration is acceptable, and adequate capacity and support are available.

A more defensible assessment keeps biological response, symptom or functional response, treatment workload, experienced burden, persistence feasibility, and patient-valued benefit separate until their relationships are examined. Sustainability should not be inferred from adherence, and treatment failure should not be inferred from burden alone.

The proposed distinction is a methodological specification for identifying discordant treatment states, not a validated score or universal threshold. Its value is empirically falsifiable. If treatment-work sustainability adds no clinically meaningful information beyond existing outcome measures, it should not become an additional assessment construct. If it identifies states in which biological improvement masks deteriorating feasibility or alters comparative and patient-centered decisions, then sustainable treatment success provides information that biological response alone cannot supply.

Acknowledgments: None

Conflict of Interest: None

Financial Support: None

Ethics Statement: None

References

1. Lee JE, Lee J, Shin R, Oh O, Lee KS. Treatment burden in multimorbidity: An integrative review. BMC Prim Care. 2024;25(1):352. doi:10.1186/s12875-024-02586-z
2. Spencer-Bonilla G, Quiñones AR, Montori VM; International Minimally Disruptive Medicine Workgroup. Assessing the burden of treatment. J Gen Intern Med. 2017;32(10):1141-5. doi:10.1007/s11606-017-4117-8

3. Christensen R, Ciani O, Manyara AM, Taylor RS. Surrogate endpoints: a key concept in clinical epidemiology. J Clin Epidemiol. 2024;167:111242. doi:10.1016/j.jclinepi.2023.111242
4. Dodd S, Gorst SL, Young A, Lucas SW, Williamson PR. Patient participation impacts outcome domain selection in core outcome sets for research: An updated systematic review. J Clin Epidemiol. 2023;158:127-33. doi:10.1016/j.jclinepi.2023.03.022
5. Lakdawalla DN, Doshi JA, Garrison LP Jr, Phelps CE, Basu A, Danzon PM. Defining elements of value in health care—a health economics approach: An ISPOR Special Task Force report [3]. Value Health. 2018;21(2):131-9. doi:10.1016/j.jval.2017.12.007
6. Bridges JFP, de Bekker-Grob EW, Hauber B, Heidenreich S, Janssen E, Bast A, et al. A roadmap for increasing the usefulness and impact of patient-preference studies in decision making in health: A good practices report of an ISPOR Task Force. Value Health. 2023;26(2):153-62. doi:10.1016/j.jval.2022.12.004
7. Mendoza-Quispe D, Perez-Leon S, Alarcon-Ruiz CA, Gaspar A, Cuba-Fuentes MS, Zunt JR, et al. Scoping review of measures of treatment burden in patients with multimorbidity: Advancements and current gaps. J Clin Epidemiol. 2023;159:92-105. doi:10.1016/j.jclinepi.2023.05.013
8. Eton DT, Yost KJ, Lai JS, Ridgeway JL, Egginton JS, Rosedahl JK, et al. Development and validation of the Patient Experience with Treatment and Self-management (PETS): A patient-reported measure of treatment burden. Qual Life Res. 2017;26(2):489-503. doi:10.1007/s11136-016-1397-0
9. Duncan P, Murphy M, Man MS, Chaplin K, Gaunt D, Salisbury C. Development and validation of the Multimorbidity Treatment Burden Questionnaire (MTBQ). BMJ Open. 2018;8(4). doi:10.1136/bmjopen-2017-019413
10. Eton DT, Linzer M, Boehm DH, Vanderboom CE, Rogers EA, Frost MH, et al. Deriving and validating a brief measure of treatment burden to assess person-centered healthcare quality in primary care: A multi-method study. BMC Fam Pract. 2020;21(1):221. doi:10.1186/s12875-020-01291-x
11. De Geest S, Zullig LL, Dunbar-Jacob J, Helmy R, Hughes DA, Wilson IB, et al. ESPACOMP Medication Adherence Reporting Guideline (EMERGE). Ann Intern Med. 2018;169(1):30-5. doi:10.7326/M18-0543
12. Vauterin D, Van Vaerenbergh F, Vanoverschelde A, Quint JK, Verhamme K, Lahousse L. Methods to assess COPD medications adherence in healthcare databases: A systematic review. Eur Respir Rev. 2023;32(169):230103. doi:10.1183/16000617.0103-2023
13. O'Donohoe P, Reasner DS, Kovacs SM, Byrom B, Eremenco S, Barsdorf AI, et al. Updated recommendations on evidence needed to support measurement comparability among modes of data collection for patient-reported outcome measures: A good practices report of an ISPOR Task Force. Value Health. 2023;26(5):623-33. doi:10.1016/j.jval.2023.01.001
14. Edgar CJ, Bush EN, Adams HR, Ballinger R, Byrom B, Campbell M, et al. Recommendations on the selection, development, and modification of performance outcome assessments: A good practices report of an ISPOR Task Force. Value Health. 2023;26(7):959-67. doi:10.1016/j.jval.2023.05.003
15. Manyara AM, Davies P, Stewart D, Weir CJ, Young AE, Wells V, et al. Definitions, acceptability, limitations, and guidance in the use and reporting of surrogate end points in trials: A scoping review. J Clin Epidemiol. 2023;160:83-99. doi:10.1016/j.jclinepi.2023.06.013
16. Baker SG, Lassere MND, Lo WP. Surrogate endpoint metaregression: Useful statistics for regulators and trialists. J Clin Epidemiol. 2024;175:111508. doi:10.1016/j.jclinepi.2024.111508
17. Gebreyohannes EA, Gebresillassie BM, Mulugeta F, Dessu E, Abebe TB. Treatment burden and health-related quality of life of patients with multimorbidity: A cross-sectional study. Qual Life Res. 2023;32(11):3269-77. doi:10.1007/s11136-023-03473-3
18. Morris JE, Roderick PJ, Harris S, Yao G, Crowe S, Phillips D, et al. Treatment burden for patients with multimorbidity: Cross-sectional study with exploration of a single-item measure. Br J Gen Pract. 2021;71(706). doi:10.3399/BJGP.2020.0883
19. Smith JL, Killian JM, Shippee N, Eton DT, Montori VM, Strand J, et al. Burden of treatment in patients with heart failure. J Am Heart Assoc. 2025;14(10). doi:10.1161/JAHA.124.039695
20. Gifford AH, Mayer-Hamblett N, Pearson K, Nichols DP. Answering the call to address cystic fibrosis treatment burden in the era of highly effective CFTR modulator therapy. J Cyst Fibros. 2020;19(5):762-7. doi:10.1016/j.jcf.2019.11.007

21. Mayer-Hamblett N, Ratjen F, Russell R, Donaldson SH, Riekert KA, Sawicki GS, et al; SIMPLIFY Study Group. Discontinuation versus continuation of hypertonic saline or dornase alfa in modulator treated people with cystic fibrosis (SIMPLIFY): Results from two parallel, multicentre, open-label, randomised, controlled, non-inferiority trials. Lancet Respir Med. 2023;11(4):329-40. doi:10.1016/S2213-2600(22)00434-9
22. Altabee R, Mwamba MJ, Turner D, Davies G, Abbott J, Simmonds NJ, et al. Measurement of treatment burden in cystic fibrosis: A systematic review. J Cyst Fibros. 2025;24(3):434-45. doi:10.1016/j.jcf.2024.11.005
23. Butcher JL, Siracusa C, Grabowski H, Yablon D, Ertman B, McWilliams E, et al. "Time out for an hour, every day, my whole life": Understanding treatment burden in cystic fibrosis in the era of elexacaftor/tezacaftor/ivacaftor (ETI). J Cyst Fibros. 2025;24(6):1130-40. doi:10.1016/j.jcf.2025.07.014
24. Basch E, Deal AM, Dueck AC, Scher HI, Kris MG, Hudis C, et al. Overall survival results of a trial assessing patient-reported outcomes for symptom monitoring during routine cancer treatment. JAMA. 2017;318(2):197-8. doi:10.1001/jama.2017.7156
25. Kearney A, Gargon E, Mitchell JW, Callaghan S, Yameen F, Williamson PR, et al. A systematic review of studies reporting the development of core outcome sets for use in routine care. J Clin Epidemiol. 2023;158:34-43. doi:10.1016/j.jclinepi.2023.03.011
26. Tervonen T, Veldwijk J, Payne K, Ng X, Levitan B, Lackey LG, et al. Quantitative benefit-risk assessment in medical product decision making: A good practices report of an ISPOR Task Force. Value Health. 2023;26(4):449-60. doi:10.1016/j.jval.2022.12.006
27. McQueen RB, Inotai A, Zemplenyi A, Mendola N, Németh B, Kaló Z. Multistakeholder perceptions of additional value elements for United States value assessment of health interventions. Value Health. 2024;27(1):15-25. doi:10.1016/j.jval.2023.09.2910
28. Slejko JF, Lavelle TA, Vandigo J, Escontrías OA, Schoch SC, Oehrlein EM. Achieving patient-centered value/health technology assessment: Recommendations from a multistakeholder eDelphi panel. Value Health. 2025;28(10):1472-80. doi:10.1016/j.jval.2025.06.014