CLARRIO.AI WHITE PAPER 02 · SCIENTIFIC PERSPECTIVES

Beyond the Clinical Trial

Longitudinal real-world evidence across the drug lifecycle

How sponsors, CROs and research teams can capture what happens to patients between visits, and turn it into evidence that is fit for its intended use.

Phase 2–3 development | Natural history | Post-authorisation research | Market access and HEOR

Readers RWE and epidemiology · clinical development · medical affairs · HEOR · CROs

Evidence base 23 sources **Reading time** about 30 minutes

47%

Pooled dropout across app-based interventions for chronic disease

Meyerowitz-Katz 2020, meta-analysis

48%

Of enrolled participants evaluable at 24 months with 80% annual retention and 75% outcome completion

Arithmetic model, Section 10

4×

The participants needed to halve the width of a confidence interval

Precision scales with √n

4

Levels of evidence use, each with different design and documentation needs

Framework, Section 06

EXECUTIVE SUMMARY

Evidence quality comes from design and measurement, not from more data

Randomised controlled trials remain the foundation for judging whether a medicine works and is safe. Their scheduled assessments and selected populations are strengths, but they describe patients at a handful of moments. Routine-care records describe encounters, not the weeks between them. In conditions where symptoms fluctuate, such as migraine, myasthenia gravis or the early months of a metabolic therapy, much of what matters to patients happens out of view.

Linking electronic health records, validated patient-reported outcomes, medication records, connected-device observations and context into one longitudinal record can close some of these gaps. It can also introduce new problems: device variability, participation bias, reporting burden and missing data that is not random. Whether the result is useful depends on the question being asked.

This paper proposes a four-level framework for evidence use: descriptive characterisation, exploratory association, fit-for-purpose measurement and comparative causal inference. For each level it sets out the design, measurement and documentation the level demands. It applies the framework to three use cases (migraine, rare-disease natural history with external comparators, and metabolic treatment utilisation) and follows evidence needs from Phase 2 to market access.

For economics, the paper avoids hypothetical savings. It offers an arithmetic "evidence yield" model for 500- and 2,000-participant programmes. At 80% annual retention and 75% completion of the planned outcome measures, fewer than half of enrolled participants contribute complete 24-month data. Raising both to 90% would cut the cost per evaluable participant by about a third at the same budget. Sponsors should measure these rates prospectively, not assume them.

Start from the question

Define the intended evidence use before choosing data sources or devices.

Measure the measurement

Track completion, retention, missingness and burden as primary operational outcomes.

Keep claims within the design

Descriptive data supports descriptive statements. Causal claims need causal designs.

About this paper. This is a scientific perspective for research professionals. It is not regulatory or legal advice and has not been reviewed or endorsed by any regulatory authority. References to FDA, EMA and ICH documents summarise publicly available guidance for context; readers should consult the current versions and their own advisers. All examples are hypothetical and contain no patient data. Clarrio.ai is not a medical device. It does not diagnose, predict disease, detect safety signals or recommend treatment.

---

01 · INDUSTRY CONTEXT

Most of the patient's experience happens between encounters

Trials, health records and claims each see part of the disease. None of them sees the daily course of symptoms, function and treatment.

Trial eligibility criteria, scheduled visits and controlled assessments give randomised trials their internal validity. The same features can limit how well a trial describes routine-care populations, or symptoms that rise and fall between visits. A migraine patient's attack frequency and disability vary day to day. In myasthenia gravis, fatigue and bulbar or limb weakness can fluctuate within a week. With metabolic therapies, tolerability and persistence change over months.

over months.

Electronic health records capture encounters, diagnoses, tests and prescriptions. Claims capture what was billed or dispensed. Registries capture disease-specific observations at defined times. Each source answers different questions, and none was built to describe the patient's experience between contacts with the health system.

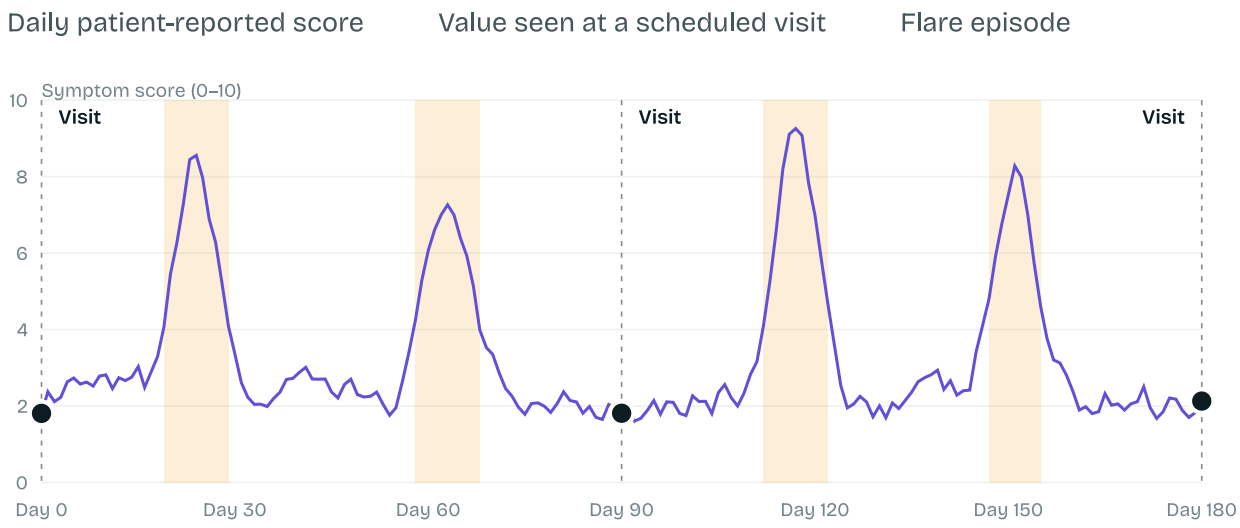
The difficulty is not only missing data. Different aspects of disease are recorded by different systems, at different times and for different purposes. Continuous patient-generated observations can narrow some of these gaps, but they bring their own measurement error and selection effects. The scientific question should determine the measurement strategy, not the other way round.

FIGURE 1

Quarterly visits record three values; the daily record shows four flares

CONCEPTUAL · SYNTHETIC

A fluctuating daily symptom score over six months, with what scheduled visits would capture



Synthetic illustration, not patient data. A fixed visit schedule can miss episodes entirely or catch them by chance. That is why recall-based measures and visit-only data may misstate burden in fluctuating conditions.

02 · EVIDENCE LANDSCAPE

Each data source answers different questions

Each data source answers different questions

Regulators have described how real-world data may be assessed for particular uses. The common thread is relevance and reliability for a specific question.

TABLE 1

What each source typically captures

Typical coverage; actual content varies by system, country and study

Usually captured · Partly or inconsistently · ○ Rarely · Not captured

SOURCE	DIAGNOSES	PRESCRIBING	DISPENSING	SYMPTOMS BETWEEN VISITS	DAILY FUNCTION	L RES
Electronic health records			○	○	○	
Claims		○				
Disease registries			○	○		
PRO and diary apps	○	○	○			
Connected devices				○		
Pharmacy dispensing						

Authors' synthesis. Sources: FDA guidance on assessing EHR and claims data[9] and EMA's reflection paper on real-world data.[3]

Interest in real-world evidence has grown with the regulatory framework around it. The 21st Century Cures Act of 2016 directed FDA to develop a programme for evaluating real-world evidence in drug regulation,[22] and FDA published that framework in 2018.[4] Since then FDA and EMA have issued documents on assessing particular data sources, digital health technologies, externally controlled trials and clinical outcome assessments. The table below

lists the documents most relevant to this paper. It summarises their scope and is not an interpretation of what any agency would accept in a particular case.

TABLE 2

Selected guidance and good-practice documents

DOCUMENT	BODY	YEAR	STATUS	RELEVANCE HERE
Framework for FDA's Real-World Evidence Program[4]	FDA	2018	Framework	Defines RWD and RWE; outlines how FDA evaluates them
Rare Diseases: Natural History Studies for Drug Development[5]	FDA	2019	Draft	Design of natural-history studies
ICH E9(R1) estimands addendum[13]	ICH	2019	Final	Estimand framework, intercurrent events
Externally Controlled Trials for Drug and Biological Products[6]	FDA	2023	Draft	External comparators
Considerations for the Use of RWD and RWE[7]	FDA	2023	Final	Study design, data access, documentation
Digital Health Technologies for Remote Data Acquisition[8]	FDA	2023	Final	Selecting and verifying devices in clinical investigations
Assessing EHR and Medical Claims Data[9]	FDA	2024	Final	Data relevance, reliability, linkage
RWD in non-interventional studies, reflection paper[3]	EMA	2025	Final	European perspective on RWD studies
PFDD: Fit-for-Purpose Clinical Outcome Assessments[10]	FDA	2025	Final	Concept of interest, context of use, COA selection

Status as published by each body at the time of writing. Guidance represents an agency's current thinking and can be revised; check the current version before relying on it.

A data point is not a measure until it represents a defined concept

Outcome assessments are chosen for a concept of interest and a context of use. Data that happens to be available is not automatically fit for either.

FDA's patient-focused drug development guidance describes clinical outcome assessments by who or what provides the information.[10] Each type can contribute to a longitudinal record, and each has different strengths. Measures from digital health technologies add a fifth category. FDA's guidance on these devices expects sponsors to show that a device measures what it is meant to measure in the intended population.[8]

TABLE 3

Types of measure in a longitudinal record

TYPE	REPORTED BY	EXAMPLES IN THIS PAPER	STRENGTH	WATCH FOR
PRO	Patient	Headache diary, HIT-6,[14] MIDAS,[18] MG-ADL[20]	Captures symptoms and impact directly	Recall period, burden, licensing and translation
ClinRO	Clinician	Clinician-rated severity scores	Standardised clinical judgement	Only available at visits; rater variability
ObsRO	Observer, e.g. caregiver	Caregiver reports in paediatric rare disease	Useful when patients cannot self-report	Observer may not see all symptoms
PerfO	Standardised task	Timed walk, grip strength	Objective, repeatable	Effort, setting and protocol effects

DHT measure	Device	Step count, sleep estimate, heart rate	Continuous, low burden once worn	Device and firmware variability; non-wear; may not map to the concept

PRO, patient-reported outcome; ClinRO, clinician-reported; ObsRO, observer-reported; PerfO, performance outcome; DHT, digital health technology.

Context changes the answer

A wearable sleep estimate may be a reasonable descriptive covariate in a migraine cohort. The same estimate is not, without further evidence, a measure of migraine severity or treatment response.

04 · TECHNICAL AND OPERATIONAL CHALLENGES

Linking sources is the easy part. Aligning their meaning is hard.

Identity linkage

Records from a patient app, a hospital EHR and a pharmacy must refer to the same person, with consent for that linkage.

Time alignment

A prescription date is not an administration date. A symptom entered on Monday may describe Saturday.

Provenance

Each value needs its origin, instrument or device version, transformations and consent restrictions.

Device variability

Brands, models and firmware differ. A device change can look like a change in the patient.

Informative missingness

People stop reporting when they feel worst, or best. Gaps are rarely random.

Participation bias

Those with compatible phones, devices and time to report may differ from those without.

Safety information

Free-text and symptom reports can contain potential adverse events that need a defined handling route.

Consent and access

Consent for care, research use, linkage and future use must be explicit and traceable.

Technical interoperability does not establish scientific comparability. HL7 FHIR can move a laboratory value between systems,[21] but analysts still need to know the assay, units, reference range and whether the patient was fasting. The same applies to patient-reported and device data.

---

05 · WHY LONGITUDINAL MULTIMODAL DATA MATTERS

Repeated, linked observations turn snapshots into trajectories

Burden between visits

Prospective diaries reduce reliance on recall and describe how symptoms fluctuate, as headache-trial guidelines recommend for migraine.[2]

Context for change

Linking symptoms to medication changes, intercurrent illness and activity helps analysts interpret trajectories instead of single values.

Natural history

Repeated standardised assessments describe progression and variability in rare diseases where little longitudinal evidence exists.[5]

Treatment in daily life

Dispensing, patient-reported use and reasons for stopping describe persistence that prescription records alone cannot.

Each of these gains depends on the measure being fit for its purpose and on missing data being handled with care. A longitudinal design gives more data points. That only helps if completeness is measured and the analysis plan says in advance how gaps will be treated.[23]

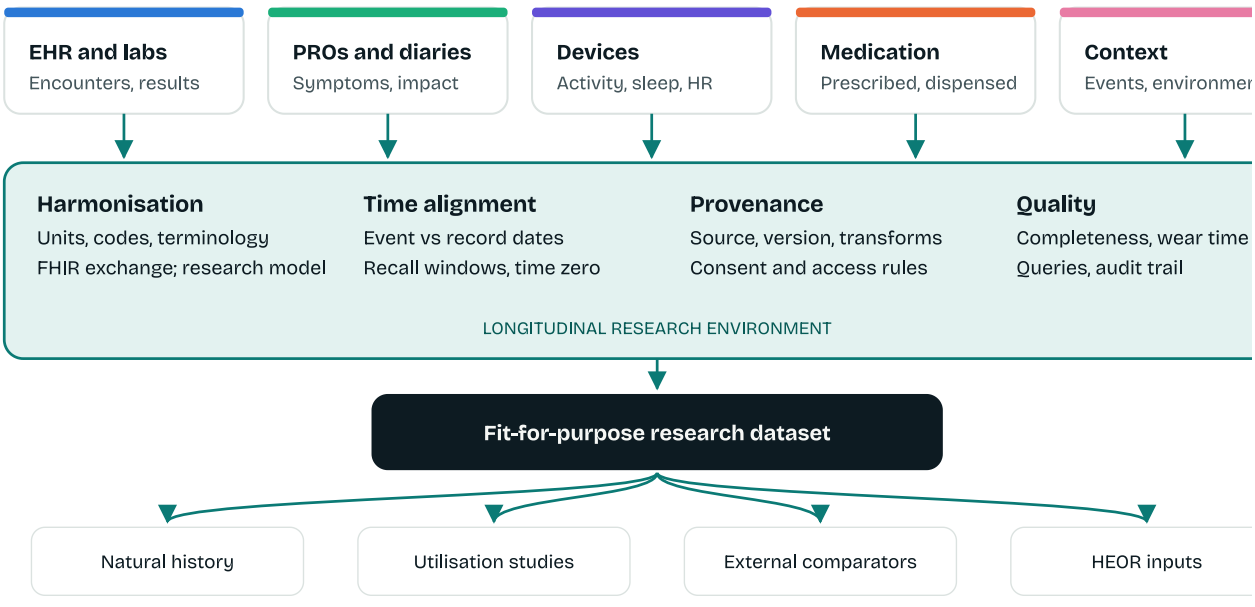
06 · PROPOSED FRAMEWORK

Four levels of evidence use, each with its own requirements

The same dataset can support some questions and not others. Naming the intended use first tells the team what the design, measures and documentation must look like.

FIGURE 2

From fragmented sources to a fit-for-purpose research dataset



Conceptual model of the information layers needed for interpretable longitudinal evidence. Hover a

Conceptual model of the information layers needed for interpretable longitudinal evidence. Hover a source for examples.

FIGURE 3

Requirements rise with the strength of the claim

Four levels of evidence use

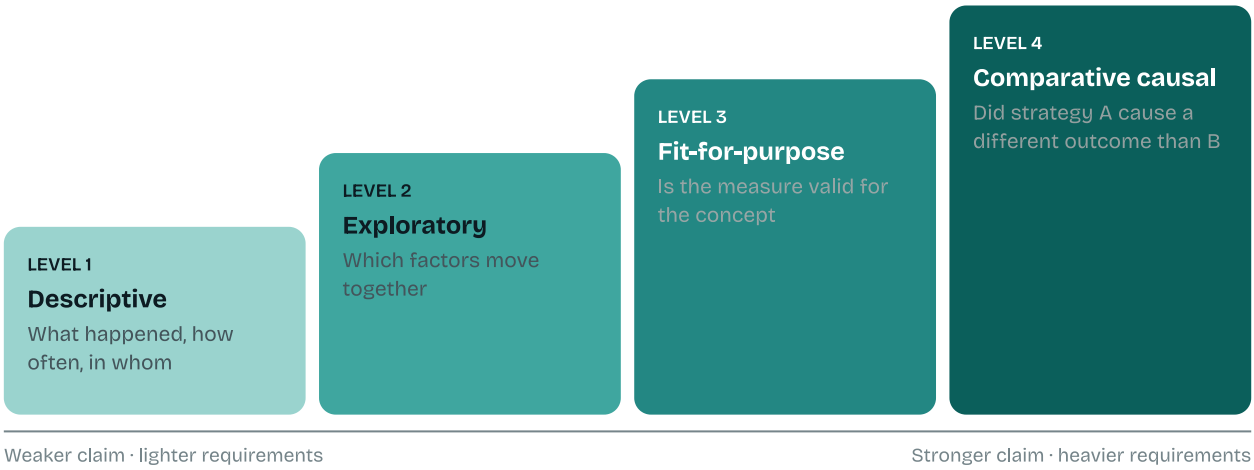


TABLE 4

Evidence interpretation: what the data can support

A study's evidentiary strength is set by its design and measurement quality

LEVEL	EXAMPLE QUESTION	DESIGN NEEDS	MEASUREMENT NEEDS	WHAT IT C
1 · Descriptive	How often do migraine days occur in this cohort, and how do they vary?	Defined population, follow-up and observation windows	Consistent definitions; completeness reported	Causes, t effects, g beyond tl
2 · Exploratory	Is irregular sleep associated with more attack days?	Pre-stated hypotheses; adjustment for obvious confounders	Covariates measured before outcomes where possible	That one causes th
3 · Fit-for-purpose measure	Does this diary-derived score reflect daily function in MG?	Measurement study in the target population	Content validity, reliability, ability to detect change[10]	Treatmen itself
4 ·	Did patients on	Target-trial	Validated outcome	Freedom

4 **Comparative causal**	Did patients on strategy A have fewer hospitalisations than on B?	Target trial emulation, estimand, time zero, bias analysis[1,11,12,15]	Validated outcome and exposure definitions; covariate history	Freedom unmeasu confound

Where data may be submitted to a regulator, sponsors should expect study-specific data assessment, documentation and early discussion with the agency.[7]

Where Clarrio.ai fits

Clarrio.ai provides an interoperable longitudinal environment that connects clinical records, patient-reported information, medication data and connected-device observations for research, with consent and provenance recorded alongside each value. It makes relevant observations accessible and traceable for study teams.

Protocol design, outcome selection, statistical analysis, causal inference, clinical interpretation, safety reporting and any regulatory interaction remain the responsibility of the sponsor and its study team. Clarrio.ai is not a medical device and does not diagnose, predict disease, detect safety signals or recommend treatment.

07 · USE CASES

Three conditions, three different evidence problems

MIGRAINE

Characterising burden between visits

RESEARCH QUESTION

How do monthly migraine days, acute medication use and disability vary over 12 months in routine care?

MEASURES

Prospective daily headache diary; HIT-6[14] and MIDAS[18] at set intervals; EHR diagnoses and treatment history. Sleep and activity as contextual covariates.

DESIGN AND ANALYSIS

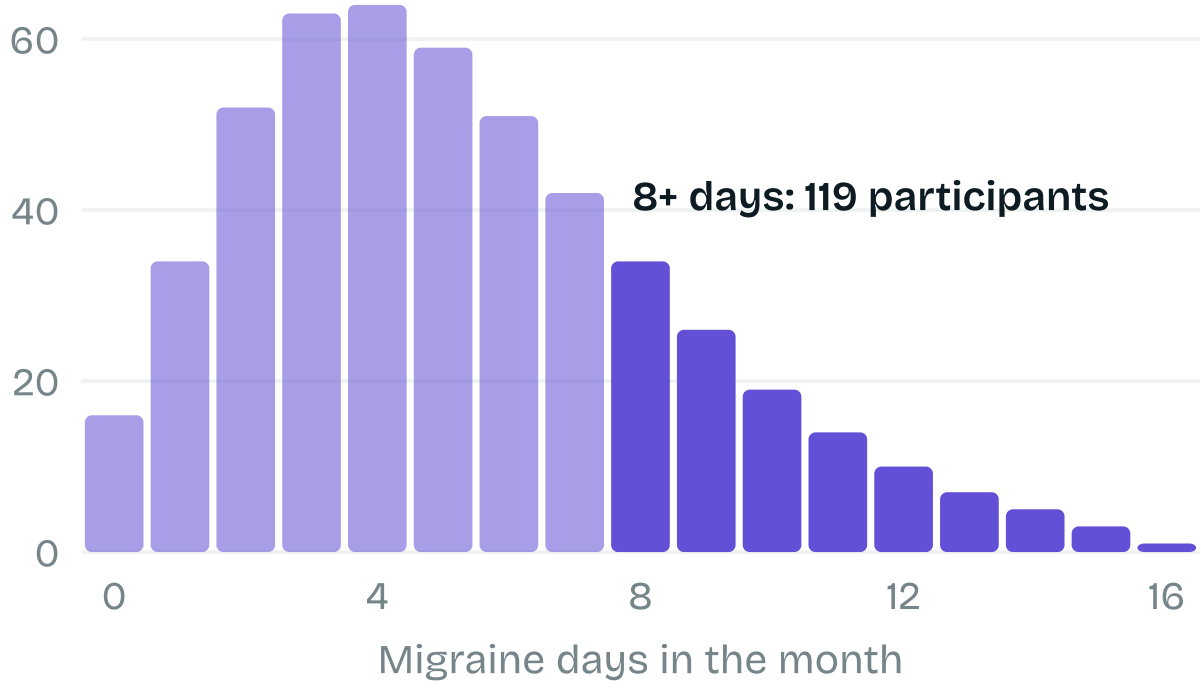
Prospective observational cohort. Diary definitions aligned with IHS trial guidelines.[2] Mixed-effects models with a pre-specified missing-data approach.

WHAT IT CANNOT SHOW

That a wearable sleep estimate measures migraine, that sleep causes attacks, or that any treatment works. Associations between sleep and attacks may reflect reverse causation.

Monthly migraine days, month 6

Synthetic cohort, n = 500



RARE DISEASE

Natural history and external comparators

RESEARCH QUESTION

How does daily function fluctuate in generalised myasthenia gravis, and could a cohort inform a future externally controlled study?

MEASURES

MG-ADL[20] at defined intervals and during flares; clinician-rated scores at visits; treatment and hospitalisation history. Device data as exploratory context only.

DESIGN AND ANALYSIS

Prospective natural-history cohort.[5] For comparator use, eligibility, baseline severity, prior therapies, calendar time and outcome definitions must match the trial.[6] Target-trial emulation, weighting and sensitivity analyses

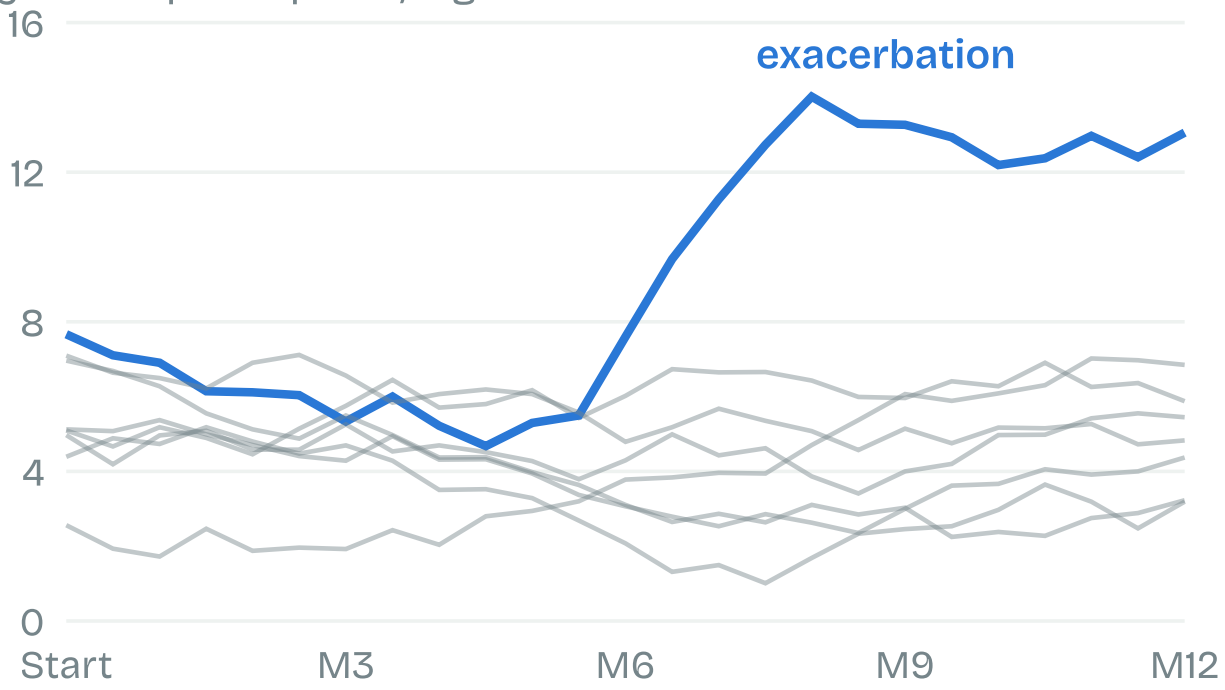
trial emulation, weighting and sensitivity analyses.

WHAT IT CANNOT SHOW

That device signals quantify neuromuscular severity, or that weighting removes unmeasured confounding or incompatible source data.

MG-ADL over 12 months

8 synthetic participants; higher = worse



METABOLIC

Treatment utilisation in daily life

RESEARCH QUESTION

How long do patients stay on metabolic therapies in routine care, and what reasons do they give for stopping?

MEASURES

Prescribing and dispensing records; patient-reported use, tolerability and reasons for interruption; weight and HbA1c from the EHR.

DESIGN AND ANALYSIS

Prescribing, dispensing and confirmed use treated as separate exposure definitions. Persistence and adherence defined before analysis. Comparisons of strategies need methods for time-varying confounding.

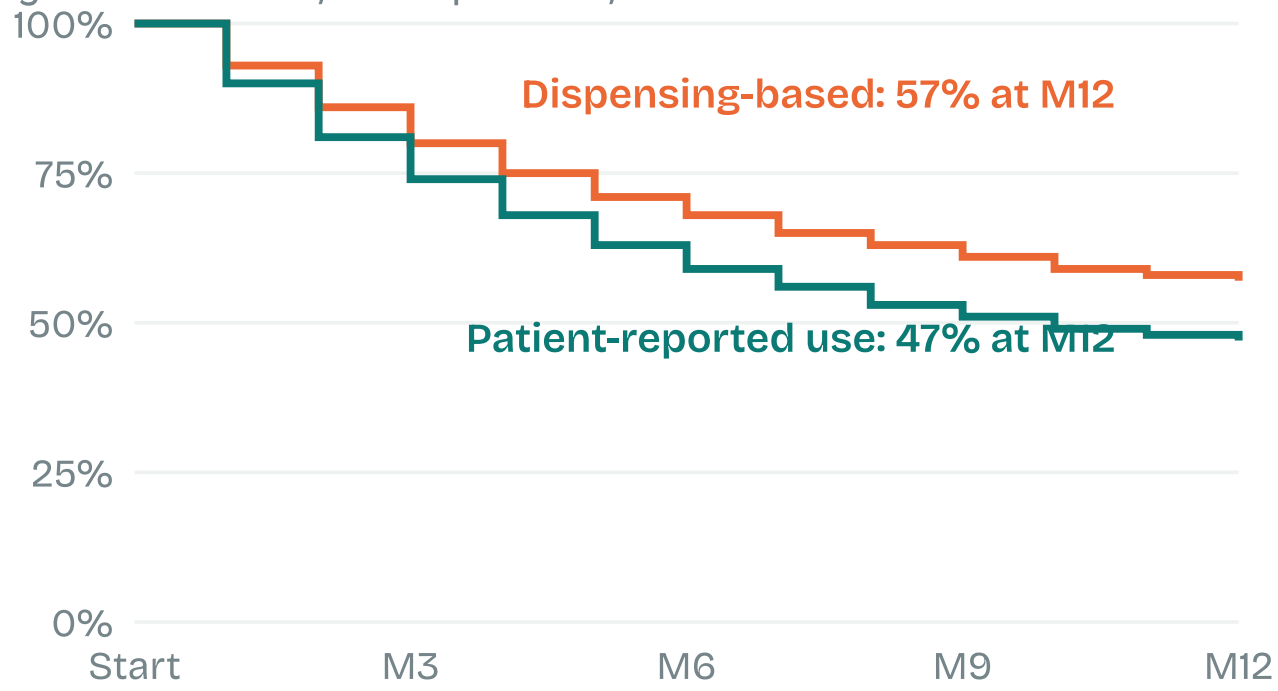
WHAT IT CANNOT SHOW

Economic value by itself. Market-access use needs validated outcomes, cost inputs, relevant comparators and transparent uncertainty.

relevant comparators and transparent uncertainty.

Persistence by exposure definition

Synthetic cohort; same patients, two definitions



08 · ACROSS THE LIFECYCLE

Evidence questions change from development to access

FIGURE 4

Where longitudinal real-world data can contribute

Typical questions at each stage; each still needs a design suited to its level of evidence use



Post-authorisation safety monitoring remains within the sponsor's established pharmacovigilance system. Patient-reported symptom capture complements that system and does not replace it.

TABLE 5

Who uses what

FUNCTION	TYPICAL DECISION	CONTRIBUTION OF A LONGITUDINAL RECORD
Clinical development	Endpoint and eligibility choices; sample size	Variability and event rates in the target population; feasibility counts
Epidemiology and RWE	Natural history; comparator feasibility	Linked, time-aligned data with provenance
Medical affairs	Evidence gaps after approval	Patient-reported burden and treatment experience in routine care
HEOR and market access	Inputs for value dossiers	Quality of life, utilisation and persistence data, where measures are validated
CROs and sites	Study operations	Remote collection between visits within the protocol; fewer manual transcriptions
Patients	Whether to take part, and how much to report	A record of their own experience; burden kept in proportion to the question

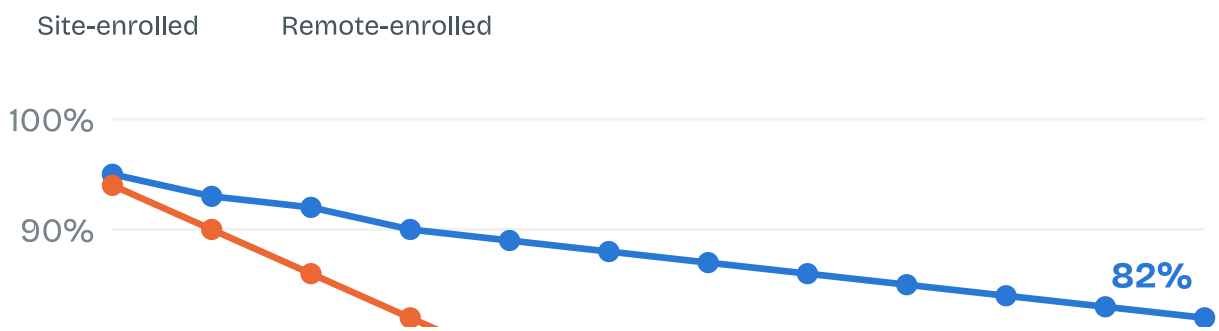
09 · ANALYTICAL OUTPUTS

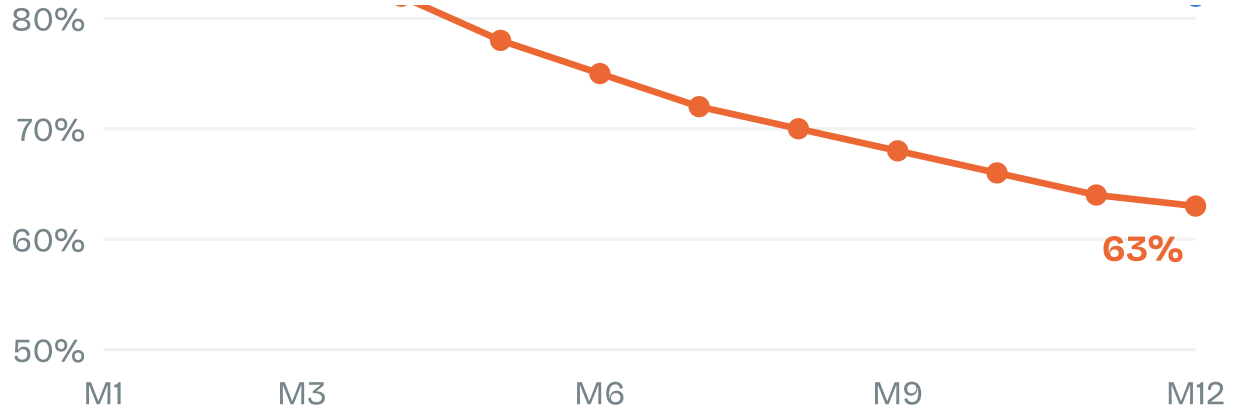
Outputs should report data quality before results

FIGURE 5

Observation coverage by month SYNTHETIC

Share of active participants with a complete weekly PRO, by month and enrolment channel



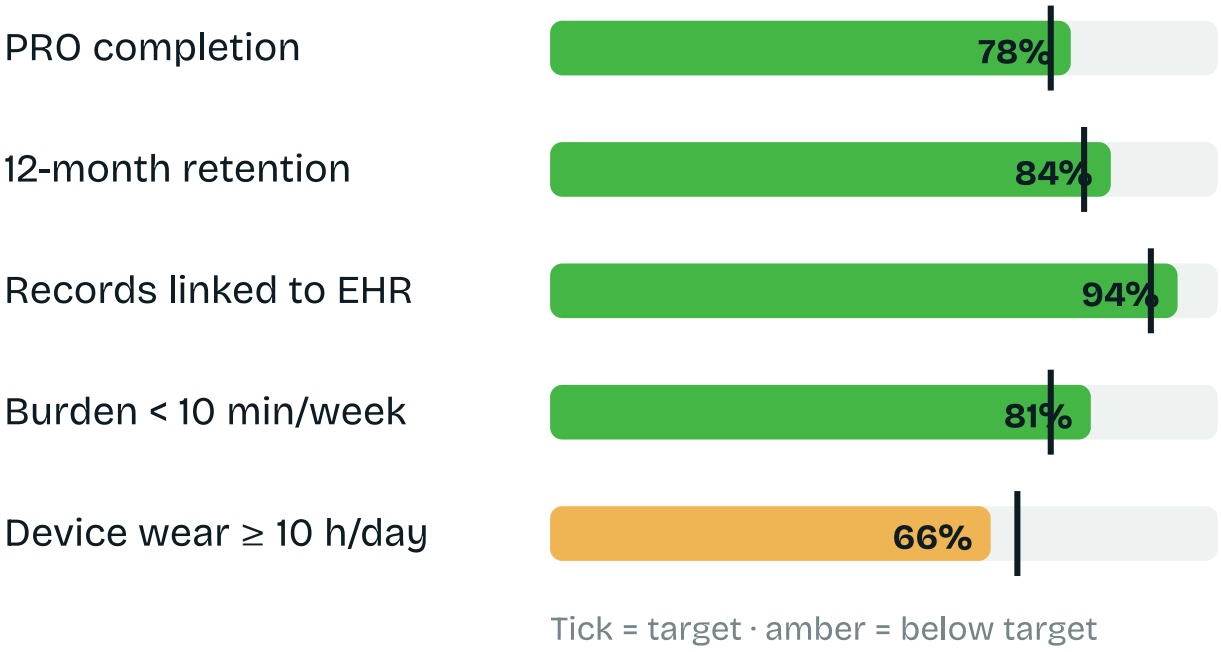


Different decline by channel signals possible differential missingness, which the analysis plan must address.

FIGURE 6

Data-quality summary SYNTHETIC

The first page of any interim data review



Thresholds are set in the protocol or data-management plan before enrolment.

- **Descriptive:** distributions of outcomes over time, by pre-specified subgroups, with numbers at risk.
- **Exploratory:** association estimates with confidence intervals, labelled hypothesis-generating.
- **Operational:** completion, retention, non-wear, query resolution time and

participant-reported burden.

- **Documentation:** data lineage, transformation logs and the versions of instruments and devices used.

---

10 · EVIDENCE ECONOMICS

The cost that matters is the cost per evaluable participant

This section uses arithmetic only. It assumes no improvement from any platform, assigns no clinical value to completeness and estimates no savings.

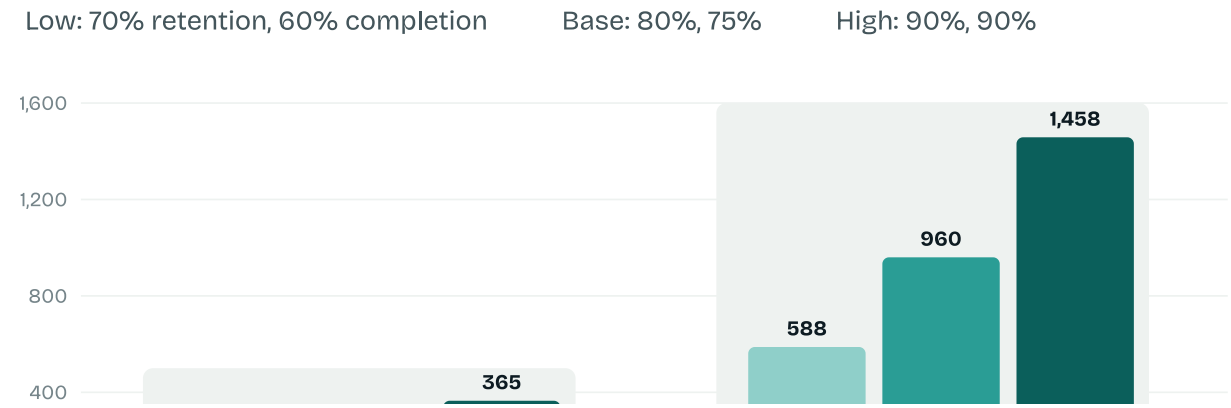
In a prospective observational programme, cost scales mostly with enrolled participant-time. The evidence produced scales with participants who contribute complete, analysable data. Two rates connect the two: **retention**, the share of participants still active after each year, and **completion**, the share of planned outcome assessments actually completed. Both should be measured from the first month and reported alongside every result.

For a 24-month programme, the share of enrolled participants with complete follow-up is approximately retention^2 × completion. The table applies this to 500 and 2,000 participants under three sets of rates.

FIGURE 7

At 80% retention and 75% completion, fewer than half of enrolled participants are evaluable

Evaluable participants at 24 months by programme size and scenario





Grey band = participants enrolled (the 2,000 band is cut at the axis top). Evaluable = N × retention² × completion. Illustrative rates; not measured outcomes of any programme.

TABLE 6

Evidence yield and precision

SCENARIO	ENROLLED	EVALUABLE AT 24 MO	SHARE	95% CI HALF-WIDTH*
Low (70%, 60%)	500	**147**	29%	±7.4 pts
Base (80%, 75%)	500	**240**	48%	±5.8 pts
High (90%, 90%)	500	**365**	73%	±4.7 pts
Low (70%, 60%)	2,000	**588**	29%	±3.7 pts
Base (80%, 75%)	2,000	**960**	48%	±2.9 pts
High (90%, 90%)	2,000	**1,458**	73%	±2.4 pts

*For an estimated proportion of 30%: 1.96 × √(0.3 × 0.7 / n). Cost index: base scenario = 100, at equal cost per enrolled participant.

FIGURE 8

Retention and completion multiply

Share of enrolled participants evaluable at 24 months

Annual retention →

Completion →	95%	90%	85%	80%	75%	70%	65%	60%
95%	34%	40%	47%	53%	61%	69%	77%	86%
90%	32%	38%	44%	51%	58%	65%	**73%**	81%
85%	31%	36%	42%	48%	54%	61%	69%	77%
80%	29%	34%	39%	45%	51%	58%	65%	72%
75%	27%	32%	37%	42%	**48%**	54%	61%	68%
70%	25%	30%	34%	39%	45%	51%	57%	63%
65%	23%	27%	32%	37%	42%	47%	53%	59%
60%	22%	25%	29%	34%	38%	43%	49%	54%

	60%	65%	70%	75%	80%	85%	90%	95%
55%	20%	23%	27%	31%	35%	40%	45%	50%
50%	18%	21%	24%	28%	32%	36%	41%	45%

Darker cells yield more evaluable participants per enrolled participant. Moving from the base cell (80%, 75%) to (90%, 90%) cuts cost per evaluable participant by about a third at the same budget.

How to use this

- Set target retention and completion in the protocol, with a minimum below which the analysis plan changes.
- Report both rates monthly, by enrolment channel and subgroup.
- Count participant-reported burden as an outcome. High burden predicts dropout.[16]
- Value operational changes only after they are measured. A gain in completion is an improvement in evidence capture, not a health benefit.

Precision is not validity

A larger evaluable sample narrows confidence intervals under the model's assumptions. It does nothing to correct selection bias, confounding or a poorly chosen measure.

INTERACTIVE

Estimate evidence yield for your programme

Enter your own rates and cost. The example cost is a placeholder, not a benchmark.

Participants enrolled 500

Annual retention 80%

Completion of planned outcome assessments 75%

Cost per enrolled participant-year (your figure) $1,500

EVALUABLE AT 24 MONTHS

240

SHARE OF ENROLLED

48%

95% CI HALF-WIDTH

±5.8 pts

PARTICIPANT-YEARS

900

PROGRAMME COST

$1.35M

COST PER EVALUABLE

$5,625

Participant-years = N × (1 + retention). CI half-width for a 30% proportion. Costs use only the figure you enter.

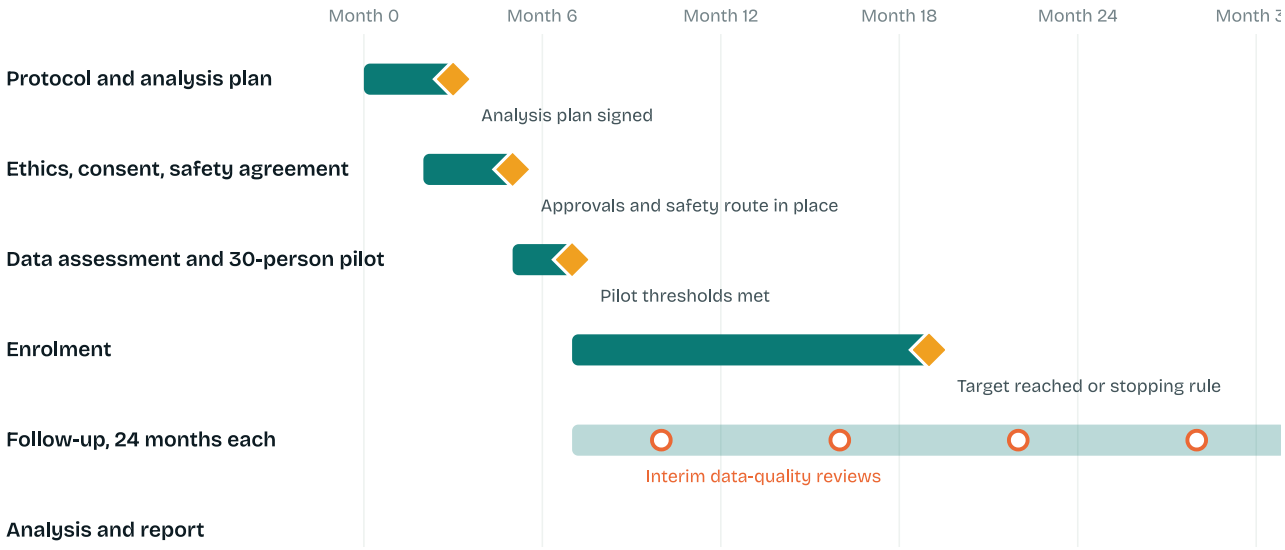
11 · ILLUSTRATIVE IMPLEMENTATION

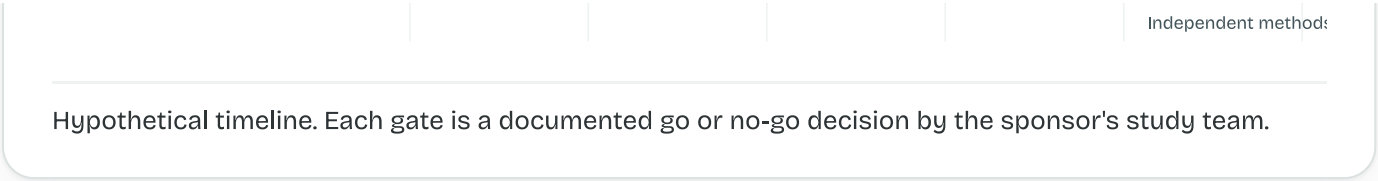
A 500-participant natural-history programme, with gates

A hypothetical prospective cohort in generalised myasthenia gravis, 24 months of follow-up per participant.

FIGURE 9

Phases and decision gates





Hypothetical timeline. Each gate is a documented go or no-go decision by the sponsor's study team.

The pilot phase answers practical questions before full enrolment: whether participants complete the diary at the planned frequency, how much time it takes them, whether device wear is sufficient, and whether linked records arrive with the expected completeness. Thresholds for proceeding are agreed in advance and recorded.

---

12 · QUALITY, BIAS, PRIVACY AND SAFETY

Good governance is designed in, not added after enrolment

12.1 Data quality

- Instrument versions, licences and translations recorded for every PRO.
- Device model and firmware stored with each observation; comparisons made within device where possible.[8]
- Non-wear distinguished from zero activity; no silent imputation.
- Data lineage and transformation logs kept for audit and reproduction.

12.2 Bias

- **Selection:** digital participation favours some groups. Compare participants with eligible non-participants.
- **Informative missingness:** plan sensitivity analyses for missing-not-at-random scenarios.[23]
- **Confounding by indication** and time-varying treatment decisions in comparative questions.
- **Quantitative bias analysis** where unmeasured confounding is plausible.[15]

12.3 Privacy, consent and security

12.3 Privacy, consent and security

Health data is special-category data under GDPR and may be protected health information under HIPAA. Consent should separate care, research use, record linkage and future use, and allow withdrawal. Controls include encryption, role-based access, audit logs and documented data residency. Sponsors qualify vendors through their own procedures, including any assessment of electronic-records requirements for clinical investigations.

12.4 Safety information

Patient-reported symptoms and free text can contain potential adverse events. Before enrolment, the sponsor's pharmacovigilance function should define what is reviewed, by whom, how quickly, and how cases are forwarded into its established safety system. Exploratory symptom data does not replace that system, and the platform does not perform safety-signal detection.

> **Planning and reporting templates help**
>
> Structured templates such as STaRT-RWE[19] and the ISPOR–ISPE good-practice recommendations[1] make design choices explicit and reproducible. Visual study diagrams make time zero and windows clear to reviewers.[17]

---

13 · LIMITATIONS AND OPEN QUESTIONS

What this framework does not settle

Limitations

- The four-level framework is a planning aid. It has not been evaluated as a tool.
- Device-derived measures in migraine and myasthenia gravis are largely exploratory.
- The evidence-yield model ignores timing of dropout and assumes completion is independent of retention.
- Use cases are hypothetical and contain no study results.

Open questions

1. What diary frequency balances completeness against burden in each condition?
2. How should differential missingness by enrolment channel be handled in natural-history comparators?
3. Which device measures, if any, can be shown to be fit for purpose for fluctuating weakness?
4. How can consent for linkage be made understandable without lowering participation?

---

14 · RECOMMENDATIONS AND CONCLUSION

Seven practices for credible longitudinal evidence

1 · Name the evidence use first

Choose the level (descriptive to causal) before choosing sources or devices.

2 · Use established measures

Prefer validated PROs and ClinROs; treat device data as context until shown otherwise.

3 · Pre-specify definitions

Exposure, persistence, outcomes, windows and time zero, before the data arrives.

4 · Track yield monthly

Retention, completion and burden, by channel and subgroup.

5 · Plan safety handling

Agree the adverse-event review route with pharmacovigilance before enrolment.

6 · Keep lineage

Provenance, versions and transformations for every value.

7 · Engage early

where data may support a regulatory question, discuss the plan with the agency early.

> **Continuous collection alone will not define the future of real-world evidence. Selected, valid and well-governed observations will.**

Longitudinal multimodal data can help close the gap between clinical encounters and the lived experience of disease. It can inform natural history, treatment utilisation, patient-centred outcomes and carefully designed observational studies, and it remains complementary to randomised trials. Its value depends on fit-for-purpose measurement, provenance, representativeness and design.

REFERENCES

Sources

1 Berger ML, et al. Good practices for real-world data studies of treatment and/or comparative effectiveness: recommendations from the joint ISPOR-ISPE Special Task Force. *Pharmacoepidemiol Drug Saf*. 2017;26(9):1033–1039. doi:10.1002/pds.4297

2 Diener HC, Tassorelli C, et al. Guidelines of the International Headache Society for controlled trials of preventive treatment of migraine attacks in episodic migraine in adults. *Cephalalgia*. 2020;40(10):1026–1044. doi:10.1177/0333102420941839

3 European Medicines Agency. Reflection paper on use of real-world data in non-