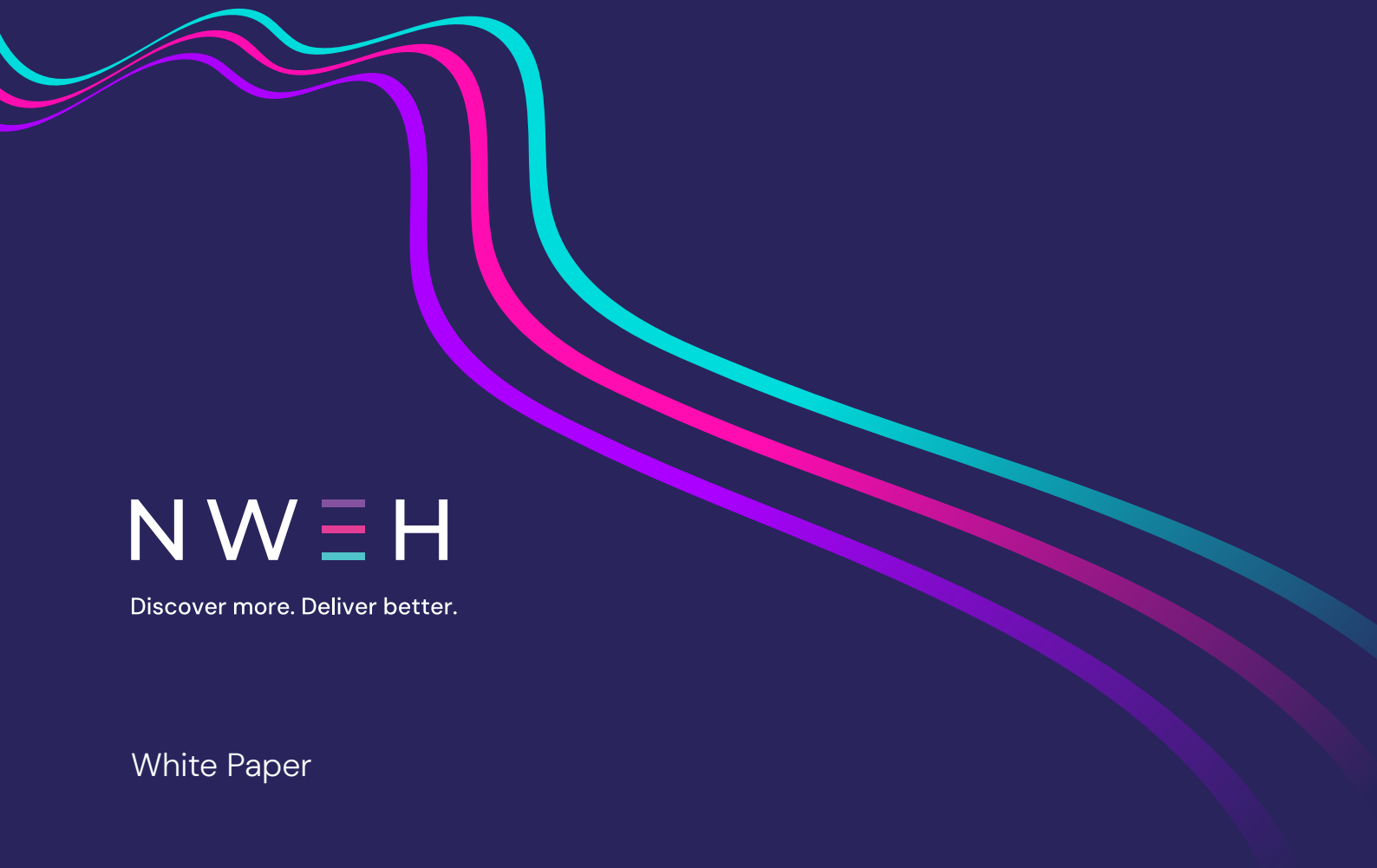
Beyond eSource

Redesigning Research Around Everyday Healthcare.

Why the future of clinical research lies in embedding research into routine healthcare.



NWEH

Discover more. Deliver better.

White Paper



Prepared by:

NWEH

Author:

Siobhan Duncan

Version:

1.0

Date:

November 2026

Beyond eSource
Redesigning research around everyday healthcare

This paper sets out why sponsors should move beyond electronic health record (EHR) to electronic data capture (EDC) transfer and towards governed reuse of routine healthcare data, with targeted study-specific capture only where it adds scientific, safety or regulatory value.

Executive summary

Clinical trials and healthcare have both become increasingly digital, but they still too often operate as parallel systems. U.S. Food and Drug Administration (FDA) guidance supports electronic source data when reliability, quality, integrity and traceability are maintained, and subsequent guidance addresses the use of EHR data in regulated clinical investigations. The strategic challenge is therefore no longer whether data can move electronically – it is whether sponsors can redesign trial operating models so that reliable routine healthcare data are reused with appropriate governance, and study-specific capture is reserved for data that cannot be derived safely or scientifically from care.

Across healthcare systems, clinicians record diagnoses, prescribe medicines, request laboratory tests, review referrals and document clinical events as part of ordinary care. When those same patients enter a clinical trial, much of that information is often collected again through a separate research process. The opportunity is not simply to connect records to trial systems, but to redesign research around routine healthcare data. The industry has therefore made progress, but not yet transformation.

For years, the sponsor question has been: how do we get EHR data into the EDC? That question is understandable. EHR-to-EDC workflows can reduce manual transcription, improve structured data-entry productivity and reduce errors in specific contexts. A 2025 time-controlled study, for example, reported 4,768 data points entered through EHR-to-EDC versus 3,023 manually, with errors reduced from 100 manual errors to one EHR-to-EDC error. The authors also caution that the findings were limited to structured labs and vitals, one site and one technology platform. The more strategic question is therefore different: which parts of this protocol genuinely require study-specific data capture, and which already exist in routine healthcare? (Patruno et al., Oct 2025)

This paper argues that the next generation of clinical research will not be defined by more digital collection alone. It will be defined by the governed reuse of routine healthcare data, targeted capture only where scientifically necessary, and operating models that embed research into care rather than layering research on top of it. Regulatory guidance and industry standards are increasingly supporting this direction, with greater emphasis on the appropriate use of real-world data, proportionate trial conduct and the interoperability needed to make healthcare data usable for research. The future of clinical research is not simply more digital; it is more integrated, more proportionate and designed around the way healthcare is actually delivered.

We solved the wrong problem

The industry has spent more than two decades digitising clinical trials. Paper case report forms became electronic case report forms (eCRFs) or EDC. Paper records became EHRs. Monitoring became increasingly remote. Trial data now move faster than they once did. Yet one of the most familiar inefficiencies in clinical research remains largely intact.

A patient visits their general practitioner (GP). Their diagnosis is recorded electronically. Their medication is prescribed electronically. Their laboratory results are returned electronically. Their referral is documented electronically. If that patient is enrolled into a clinical trial, much of the same information may then be entered again into a separate research system. Sometimes it is entered manually. Sometimes it is transferred electronically. Often it still requires reconciliation, query management, monitoring and transformation before it becomes usable evidence.

That duplication has become so normal that sponsors, contract research organisations (CROs) and sites often treat it as part of the cost of running trials. Paradigm Health/GlobalData research suggests that eSource familiarity and sponsor interest are rising, but that hybrid and EDC-centred models remain common, with integration complexity, site readiness and lack of standardisation still acting as barriers. The problem is not lack of digital ambition. It is that much of the industry still treats the EDC as the centre of the trial and asks healthcare data to feed it. TransCelerate makes a similar point from an industry consortium perspective: eSource adoption has been slow, particularly for registration trials, because operationalisation is difficult; as protocol-defined data increasingly come from patient records, devices and apps, interoperability and consistent data standards become paramount. Connecting systems matters, but connection alone is not transformation. (TransCelerate, 2026)

A different question

The established question, “How do we get EHR data into the EDC?”, assumes that the traditional trial workflow remains the organising architecture. It asks how routine healthcare data can serve an existing research process, rather than whether the process itself should be redesigned around data that already exist.

The better question is: “Which protocol data already exist in routine care, and what genuinely requires study-specific capture?” This changes the conversation. Healthcare data stops being merely an input to the EDC and becomes the foundation for a different operating model. Study-specific assessments still matter. Clinical judgement still matters. Participant safety still matters. Investigator oversight still matters. But duplication begins to lose its default status.

> **Clinical trials should collect only the information that healthcare cannot already provide and only after proving that routine data is fit for purpose.**

That qualification is essential. Routine data is not automatically research grade simply because it is digital. A study of pragmatic trials using EHR data found major concerns about reliability, missing or incomplete elements, variation across sites, and implausible or incorrect values; although 79% of surveyed trials conducted data quality assessments, fewer than a quarter did so before selecting a data source. The lesson for sponsors is not to abandon routine data. It is to govern it properly, test it early and design protocols around what the data can reliably support. (Raman, et al., 2022)

Three ways to run exactly the same study

This distinction matters because the same protocol can appear superficially similar across delivery models while placing very different demands on participants, sites and data-management teams. The comparison below uses one study to show where the operating model changes, not the scientific question.

The Same Trial, Three Ways

This comparison remodels an existing five-year Phase 4, multicentre, randomised, open label, interventional study currently delivered using ConneXon, as two alternative delivery models: a conventional site-based RCT and an EHR-to-EDC implementation.

The protocol, endpoints, eligibility criteria, and scientific objectives are held identical throughout. Only the method of capturing and managing study data changes.



Figure 1 Three ways to deliver the same clinical study. The protocol remains identical; only the operating model changes. Figures shown are illustrative planning estimates

ConneXon is NWEH's clinical research platform, designed to integrate routinely collected healthcare data with study-specific data and transform them into governed, analysis-ready datasets, including Study Data Tabulation Model (SDTM) outputs. It is used here as a practical example of the operating model explored in this paper.

To illustrate the difference, we used an existing real-world five-year, Phase 4, multicentre, randomised, open-label interventional obesity study delivered using ConneXon, and modelled how the same protocol could operate across three delivery models: a traditional randomised controlled trial (RCT), an EHR-to-EDC eSource model, and the ConneXon routine-data-to-analysis-ready model. The protocol, scientific objectives and endpoints remain the same; only the method of collecting, governing and managing data changes.

In the ConneXon routine-data-to-analysis-ready model, routine healthcare data, connected sources and targeted eCRFs are combined through governed mapping, provenance and transformation into analysis-ready outputs.

In the real-world trial comparison, face-to-face visits fall from 16 to four per participant, supplemented by 12 telehealth contacts; participant visit burden falls from 40 hours to 16 hours over five years; estimated site contact time falls from 32 hours to 13 hours; and a traditional Clinical Data Acquisition Standards Harmonization (CDASH)-aligned model requiring 28 eCRFs for manual transcription is compared with 13 targeted ConneXon eCRFs. These are directional planning estimates, not validated time-and-motion findings, and should therefore be used to frame hypotheses for sponsor evaluation rather than quantified savings claims.

One Study. Three delivery models

This comparison is not between three different studies. It is the same protocol, remodelled across three operating models. Scientific objectives, endpoints and eligibility criteria remain unchanged; only the data pathway changes

Conventional site-based RCT		eSource EHR to EDC		ConneXon Real World Evidence	
16	Number of face-to-face visits per participant*	**16**	Number of face-to-face visits per participant*	**4**	Number of face-to-face visits per participant*
40hr	Participant hours (visit burden, over 5 yrs)*	**40hr**	Participant hours (visit burden, over 5 yrs)*	**16hr**	Participant hours (visit burden, over 5 yrs)*
32hr	Site contact hours (incl. safety review)*	**~25hr**	Site contact hours (incl. safety review)*	**~13hr**	Site contact hours (incl. safety review)*
High	Manual transcription of 28 eCRF forms	**Reduced**	Manual transcription of 23 eCRF plus 10 auto eCRF forms	**Minimal**	Manual transcription 13 eCRF forms the rest goes direct to the SDTM

Figure 2 Illustrative operational comparison of three delivery models using the same five-year Phase 4 protocol.
**The figures are directional planning estimates and should not be interpreted as validated savings claims*

The comparison can be summarised as a shift in where trial work happens. In a traditional RCT model, the site recreates protocol data in research systems. In an EHR-to-EDC model, selected routine data are transferred into the EDC, reducing transcription but preserving many downstream hand-overs. In a routine-data-to-analysis-ready model, validated routine data, connected sources and targeted study-specific capture are governed together so that fewer fields need to be recreated and more data can move with provenance toward analysis-ready outputs.

Why the data pathway matters

The distinction between EHR-to-EDC and routine-data-to-analysis-ready delivery is subtle but strategically important. Many organisations describe studies as EHR-enabled because healthcare data populate an EDC system, reducing manual transcription and improving data quality. That is a meaningful improvement over traditional data entry, but the EDC remains an additional staging layer before data become analysis ready.

ConneXon adopts a different approach: routinely collected healthcare data is validated, governed and transformed directly into analysis-ready datasets where appropriate, with targeted eCRFs used only where information cannot be derived from routine care. The result is not simply less data entry, but fewer process steps, reduced reconciliation, lower monitoring effort and a more efficient pathway from routine healthcare to usable clinical evidence.

Not all ‘EHR-enabled’ trials save the same time

Fewer hand-overs, less duplicate entry, faster to SDTM

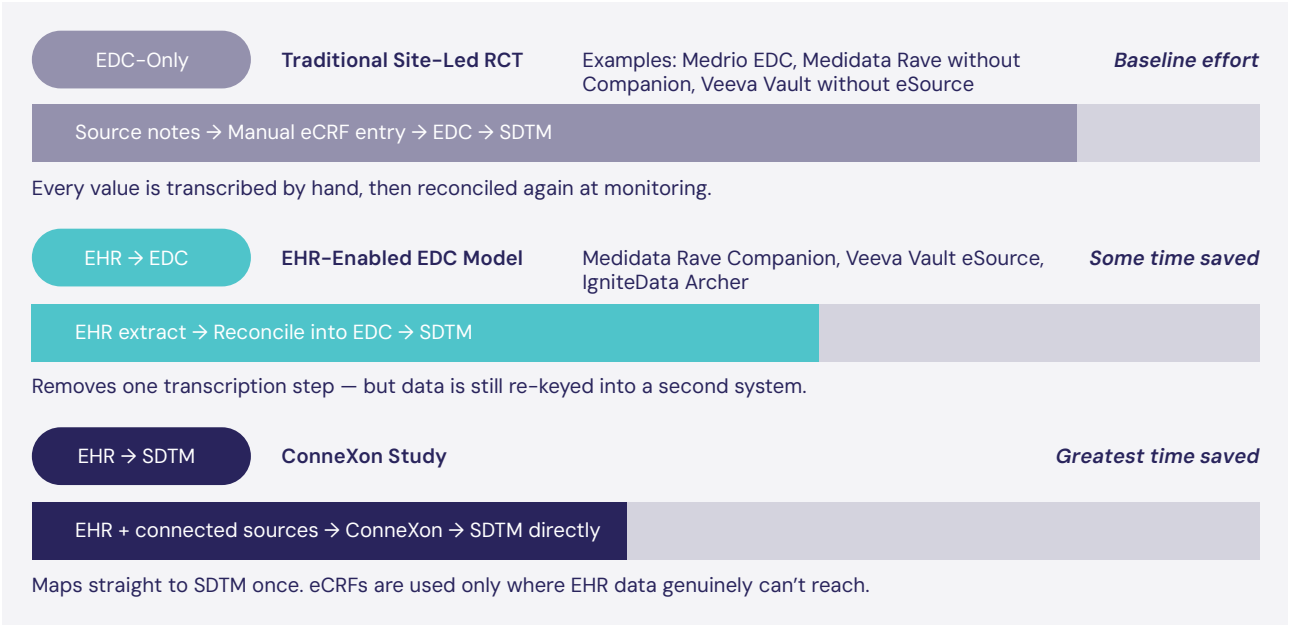


Figure 3 EHR-to-EDC reduces manual entry; ConneXon removes additional hand-overs by taking routine healthcare data directly to analysis-ready datasets.

When the operating model changes

The operating-model change is not simply that fewer forms are completed, or that some visits become remote. It is that research can fit more naturally around how participants already receive care, while investigators and coordinators spend less time documenting information that already exists. The goal is not automation for its own sake; it is removing work that adds little scientific value while preserving the judgement, oversight and safety controls that do.

A recent EHR-to-EDC study supports the narrow operational value of reducing manual structured-data entry, but it also shows the boundary of the current evidence base: the authors note reliance on structured data, standard identifiers such as Logical Observation Identifiers Names and Codes (LOINC) and a single-site context. That distinction matters for sponsor boards.

Validated evidence currently supports specific workflow benefits; broader claims about portfolio-wide cost, database-lock acceleration or monitoring reduction should be treated as measurable opportunities, not guarantees. (Patruno, et al., Oct 2025)

Decentralised and hybrid trial evidence points in the same direction, while still leaving important questions for sponsors to test. A 2023 systematic review of decentralised clinical studies found that 11 of 13 studies reported improved recruitment, seven reported positive retention outcomes and six reported diversity benefits or trends, while acknowledging that more comparable evidence is needed. FDA’s decentralised elements guidance recognises telehealth visits, in-home visits and local healthcare provider visits as ways for trial activities to occur remotely at locations convenient for participants, while stressing that regulatory requirements remain the same and that sponsors need appropriate oversight, data integrity and participant safety controls.

The implication is clear: sponsors should not evaluate digital trial models by the number of systems connected. They should evaluate them by the amount of unnecessary work removed, the quality of provenance preserved, the participant burden reduced and the speed with which reliable evidence becomes usable.

Technology does not reduce burden. Removing unnecessary steps does.

The value of digital transformation is not measured by the number of systems connected. It is measured by the amount of unnecessary work that disappears.

Routine healthcare becomes a research platform

The broader ambition is to change the relationship between healthcare provision and research, not just to make individual trials marginally more efficient.

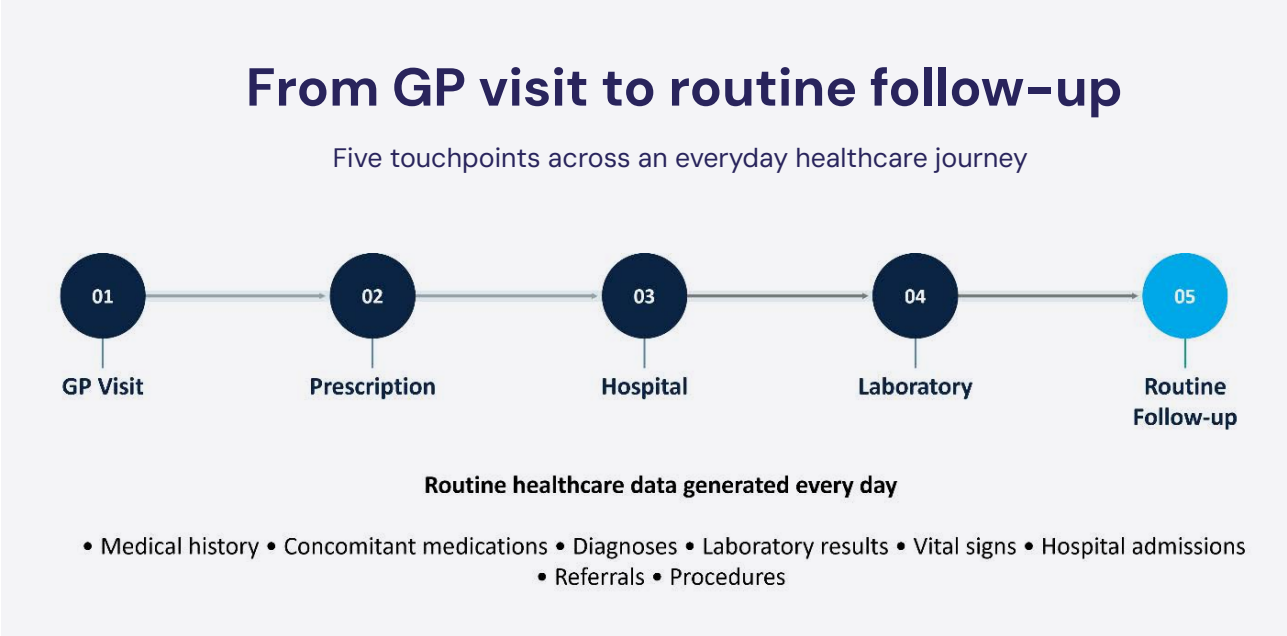


Figure 4 Routine care timeline. Illustrative sequence; timing and care setting may vary.

The Salford Lung Study remains one of the most memorable proofs of possibility because it showed how routine-care infrastructure can support regulated research without making research feel separate from care. It was designed as a pre-licence pragmatic Phase III randomised trial in asthma and chronic obstructive pulmonary disease (COPD), conducted in routine primary care, using electronic health records for near-real-time monitoring and data collection while minimising additional patient visits. A subsequent safety-methodology paper reported that the programme used linked data from primary care practices, two university hospitals and national databases; daily summaries alerted a specialist safety team to potential safety events, and the authors concluded that integration of multiple data sources may provide more comprehensive safety information than standard RCTs. (New et al., 2014) (Collier et al., 2017). The lesson from Salford is not that every trial should become a pragmatic trial. Sponsors will continue to need explanatory RCTs, tightly controlled endpoints, adjudication, central laboratories, imaging review, protocol-defined assessments and investigator oversight. The lesson is that routine-care infrastructure can support a broader evidence spectrum: pragmatic interventional studies, long-term follow-up, registry-based studies, post-authorisation safety studies, real-world evidence programmes and learning-health-system models.

The European Medicines Agency (EMA) real-world evidence programme is moving in the same direction. The EMA states that greater use of real-world data and evidence can improve the evidence base for benefit-risk decisions, and Data Analysis and Real World Interrogation Network (DARWIN EU) provides evidence on the use, safety and effectiveness of medicines from real-world healthcare databases, with key figures showing around 250 million patients providing data in Europe in 2026, around 110 studies delivered since 2022 and around 40 data partners as of 2026.

Regulators are moving in the same direction

> Regulators are not asking sponsors to be more digital for its own sake. They are asking for evidence that is reliable, traceable, proportionate and inspection ready.

The FDA's electronic-source guidance promotes electronic capture while emphasising reliability, quality, integrity and traceability from electronic source to regulatory submission. FDA's EHR guidance recognises EHR systems as platforms maintained by healthcare providers and used to deliver care, and its 2024 RWD guidance gives sponsors considerations for using EHR and claims data in clinical studies to support regulatory decisions on effectiveness or safety.

The International Council for Harmonisation (ICH) E6(R3) reinforces the same direction of travel. The EMA summary states that E6(R3) introduces provisions applicable across trial types and settings, encourages risk-based and proportionate approaches, and includes Annex 2 for non-traditional interventional trials, pragmatic trials, decentralised trials and trials incorporating real-world data sources. The Clinical Data Interchange Standards Consortium (CDISC) and HL7 (Health Level Seven) have also created practical standards to make it easier for healthcare data to be mapped into regulatory submission formats.

The constraint is therefore no longer technology alone. It is governance, quality and trust. i~HD's eSource data-flow work makes this explicit by setting out multiple EHR to EDC integration patterns and noting that implementation depends on architecture, governance, technical readiness, operational fit and regulatory alignment. Its scalability framework goes further: moving from pilots to sustainable adoption requires institutional and regulatory foundations, operational excellence, collaboration and evidence generation, and future-readiness. (i~HD_eSource_Data_Flow_Pattern, 2025) (i~HD_Scaling_eSource_in_Clinical, 2025)



What this means for sponsors

Sponsors should move beyond asking whether EHR data can be used and instead ask sharper questions during protocol design.

Sponsors should ask:

- Which protocol data already exist in routine care?
- Which fields require targeted eCRFs?
- Which outcomes require adjudication or study-specific assessment?
- Which data sources are structured, complete, timely and reliable enough for their intended use?
- Which activities should remain human because they involve judgement, safety interpretation or participant-facing discussion?

This is not a call to automate away clinical responsibility. Eligibility confirmation, clinical judgement, prescribing, safety review, study-specific measures and investigator oversight remain essential. FDA's decentralised guidance similarly emphasises sponsor and investigator responsibilities, safety monitoring, data management plans, data origin, data flow, visit type, location, date and data-originator documentation. (U.S. Department of Health and Human Services, 2024).

The sponsor operating model should therefore shift in four ways:

1. **Design protocols around data provenance, not just visit schedules.**
   Each field should be classified as routine source, targeted eCRF, patient-reported, device-derived, adjudicated or not collectible without redesign.

2. **Validate routine data before committing to it.**
   Fitness-for-use assessment should happen early enough to change protocol design, not after data gaps are discovered mid-study.

3. **Redefine monitoring value.**
   When provenance is preserved and automated checks are credible, oversight can focus more on critical-to-quality factors, exceptions, safety signals, missingness and data lineage.

4. **Measure transformation by work removed.**
   Useful sponsor metrics include duplicate fields avoided, manual-entry hours reduced, query volume, endpoint completeness, safety-event timeliness, percentage of protocol fields populated from validated routine sources, participant travel burden, retention and representativeness.

The evolution of clinical research

Taken together, these changes represent a broader evolution in clinical research, from digitising existing processes towards embedding research within everyday healthcare.

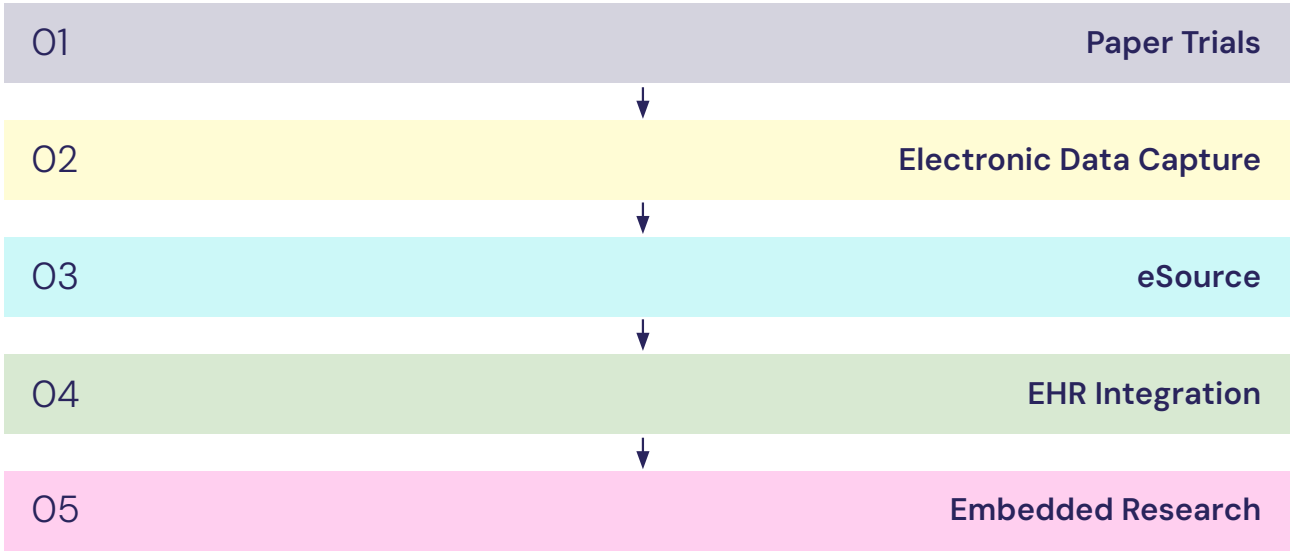


Figure 5 The evolution of clinical research operating models. The progression from paper-based trials to EDC, eSource and routine healthcare data illustrates a shift from digitising research processes towards embedding research within everyday healthcare.

The practical test for sponsors is therefore simple: if a protocol field already exists in routine care, can it be used with sufficient provenance, quality and oversight for its intended purpose? If the answer is yes, the protocol should avoid recreating it. If the answer is no, the protocol should specify the targeted capture, adjudication or clinical judgement required. That discipline is what turns routine healthcare data from a digital feed into a governed research asset.

Conclusion

Imagine a participant visiting their GP. The consultation contributes to their healthcare, and it also contributes to a clinical trial. Not because another eCRF has been completed or another database has been updated, but because care and research are finally speaking the same language.

That is the opportunity before pharmaceutical sponsors. The future of clinical research is not simply more digital. It is more integrated, more proportionate, more participant-centred and more closely aligned with the way healthcare is actually delivered.

Electronic health records made digital healthcare possible. eSource made electronic capture more practical. The next step is more ambitious: to make routine healthcare a governed, trusted and natural foundation for clinical research.

That is what it means to move beyond eSource.

About NWEH

NWEH sits at the intersection of healthcare, data, technology and clinical research.

We have built a unique infrastructure that combines healthcare data, technology, trusted NHS partnerships and research delivery expertise to enable clinical research within routine care.

By connecting patients, healthcare professionals, researchers and life sciences organisations through a single ecosystem, we help make research more accessible, evidence more representative and innovation more closely aligned to real-world healthcare.

For more than 15 years, we have helped to redefine where and how clinical research happens – bringing healthcare and research closer together to improve outcomes for patients, healthcare systems and industry.

Our mission is simple

To make clinical research part of everyday healthcare, for everyone.

We believe access to clinical research and innovation should be embedded within routine care, not limited by geography, awareness or opportunity.

By connecting healthcare, data and research, we are creating a future where participation in clinical research becomes a normal part of healthcare delivery, and access to innovation becomes the norm rather than the exception.

Building the clinical trials layer of everyday healthcare.

References

Collier et al., 2017. Monitoring safety in a phase III real world effectiveness trial: use of novel methodology in the Salford Lung Study. Pharmacoepidemiology and Drug Safety, 26(3), pp. 344-352.

i~HD_eSource_Data_Flow_Pattern, 2025. eSource Data Flow Patterns: EHR-to-EDC Integration Models for Clinical Trials. [Online]
Available at: www.i-hd.eu/uploads/eSource/Publications/Annex-1_e-Source_Patterns_Final_June-10.pdf

i~HD_Scaling_eSource_in_Clinical, 2025. Scaling eSource in Clinical Trials: A Framework of Factors and Dimensions for Successful EHR2EDC Implementation. [Online]
Available at: https://www.i-hd.eu/uploads/eSource/Publications/2.-e-Source_Scalability_Final_Nov-26.pdf

New, J. P., Bakerly, N. D., Leather, D. & Woodcock, A., 2014. Obtaining real-world evidence: the Salford Lung Study. Thorax BMJ, 69(12), pp. 1152-4.

Paradigm Health, 2026. How eSource adoption is saving time and costs in clinical research, s.l.: Clinical Trials Arena.

Patruno, A., Panzarella, M. O., Buckley, M. et al, Oct 2025. Evaluating the Impact of Electronic Health Record to Electronic Data Capture Technology on Workflow Efficiency: a Site Perspective. JAMIA Open, 8(5).

Raman, S. R., O'Brien, E. C., Hammill , B. G et al, 2022. Evaluating fitness-for-use of electronic health records in pragmatic clinical trials: reported practices and recommendations. Journal of the American Medical Informatics Association, 29(5), pp. 798-804.

TransCelerate, 2026. transceleratebiopharmainc.com/initiatives/esource/. [Online]
Available at: https://www.transceleratebiopharmainc.com/initiatives/esource/ [Accessed 2026].

U.S. Department of Health and Human Services, 2024. Conducting Clinical Trials With Decentralized Elements Guidance for Industry, Investigators, and Other Interested Parties. [Online]
Available at: https://www.fda.gov/media/167696/download

The Same Clinical Study. Three Ways to Collect Data. Very Different Outcomes.

The protocol, endpoints and scientific objectives are identical. Only the data pathway changes.

ROUTINELY COLLECTED HEALTHCARE DATA

Already exists in everyday care

- Concomitant medications
- Lab results (haematology, chemistry, etc.)
- Vital signs
- Medical history & diagnoses
- Procedures & interventions
- Hospital admissions & ER visits
- Adverse events & outcomes
- Imaging & results
- Demographics

This information is generated as part of routine care, not because of the trial.

1 TRADITIONAL RCT (EDC-ONLY)

Manual collection



SDTM Dataset
(Data cleaned, transformed and mapped)

What this means

- Every value is transcribed by hand
- High risk of errors and missing data
- Extensive reconciliation and queries
- Highest burden for sites & participants

- Highest manual effort
- Highest error & query rate
- Highest site & participant burden
- Longest time to dataset

2 EHR → EDC (EHR-ENABLED EDC MODEL)

Populate EDC with EHR data



SDTM Dataset
(Data cleaned, transformed and mapped)

What this means

- Removes manual transcription
- Fewer data entry errors
- Still requires EDC reconciliation & queries
- Monitoring effort remains substantial

- **Moderate** manual effort
- **Reduced errors** vs. manual entry
- **Reduced** site & participant burden
- **Faster** time to dataset

3 EHR / SOURCES → SDTM (CONNEXON REAL WORLD EVIDENCE)

Direct to analysis-ready dataset with minimal EDC capture

A. ROUTINE DATA FLOW (MAXIMISED)



EHR / connected sources
(healthcare systems, labs, registries, devices)

ConneXon
maps & standardises with provenance and quality checks

SDTM Dataset (analysis-ready)
Directly mapped to SDTM

B. TARGETED EDC CAPTURE (MINIMISED)



Study-specific assessments

eCRF / EDC
(only where EHR doesn't cover)

Mapped into ConneXon / SDTM

What this means

- No EDC hand-off for routine data – goes straight to SDTM
- Automated provenance & data lineage
- Exception-based review, not full verification
- Lowest burden for sites & participants

- **Lowest** manual effort
- **Lowest** errors & queries
- **Lowest** site & participant burden
- **Fastest** time to dataset

THE BENEFIT

- Fewer steps
- Less rework
- Better data quality
- Lower burden for patients and sites
- Faster, more efficient studies

Same study. Same endpoints. Same scientific integrity.

The difference is removing unnecessary steps so that healthcare data can deliver maximum value with minimum burden.

ConneXon is the slickest, most efficient path to evidence.

Illustrative comparison based on a five-year Phase 4, multicentre, randomised, open-label study. Figures are directional planning estimates and not validated savings claims.

NWEH

Discover more. Deliver better.

nweh.co.uk