



**Globally competitive
UK-wide data-enabled
clinical trials:
the time is now**

March 2026

Contents

Introduction	3
The trial recruitment challenge	6
Shortfalls with current recruitment processes	10
Creating a competitive UK-wide offering	15
Industry requirements	18
Service design	27
Targeting the right participants for screening	32
Implications for different therapeutic areas	37
Measuring success	41
Conclusion	44
Annex	48
References	55

Introduction



Commercial clinical trials are central to the UK's life sciences sector, critical for patient access to innovative treatments, and represent substantial economic activity and scientific advancement.

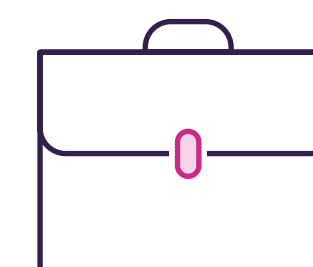
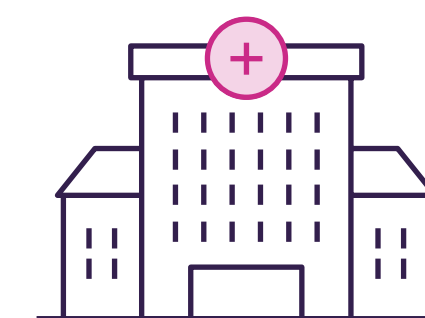
In 2022, industry clinical trials contributed £7.4 billion to the UK economy, generated £1.2 billion revenue for the NHS, and supported a total of 65,000 jobs across the UK.¹ Clinical trials offer patients the potential to access innovative treatments that may not be available through standard care, which can be life changing for those with limited options. Therefore, the UK's success in commercial clinical trials delivery directly impacts patient outcomes, economic prosperity and the UK's position as a global leader in life sciences.

In 2022 industry clinical trials contributed:



£7.4bn
for the UK
economy

£1.2bn
revenue
for the NHS



65,000
jobs across the UK

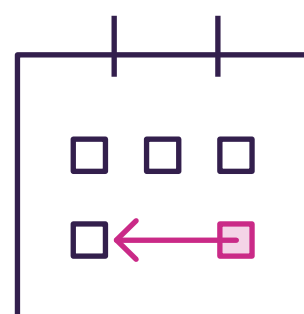
In the past decade, the UK was a global powerhouse in commercial trial delivery. However, in part due to the UK's leading role in delivering COVID-19 clinical trials, recovery was slower than other countries, resulting in a major decline in UK competitiveness.² Returning to 2017 trial activity levels would generate an added £3 billion for the economy, £485 million in NHS revenue, and 26,000 jobs.³

Returning to 2017 levels of industry clinical trials would generate an additional:

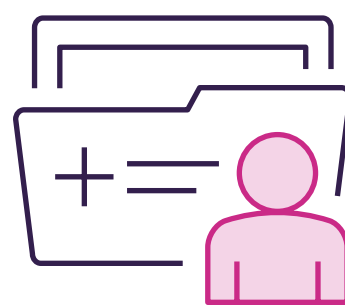


The Prime Minister's commitment to reduce trial set-up times to 150 days by March 2026,⁴ combined with the Life Sciences Sector Plan's goal to quadruple commercial trial participant recruitment by 2029,⁵ signals government's recognition of the opportunities and urgent need to restore the UK's competitiveness in commercial trial delivery. These commitments are supported by the £600 million investment in the Health Data Research Service (HDRS), in partnership with the Wellcome Trust, which should underpin a data-driven transformation of clinical research.⁶

Two government commitments:

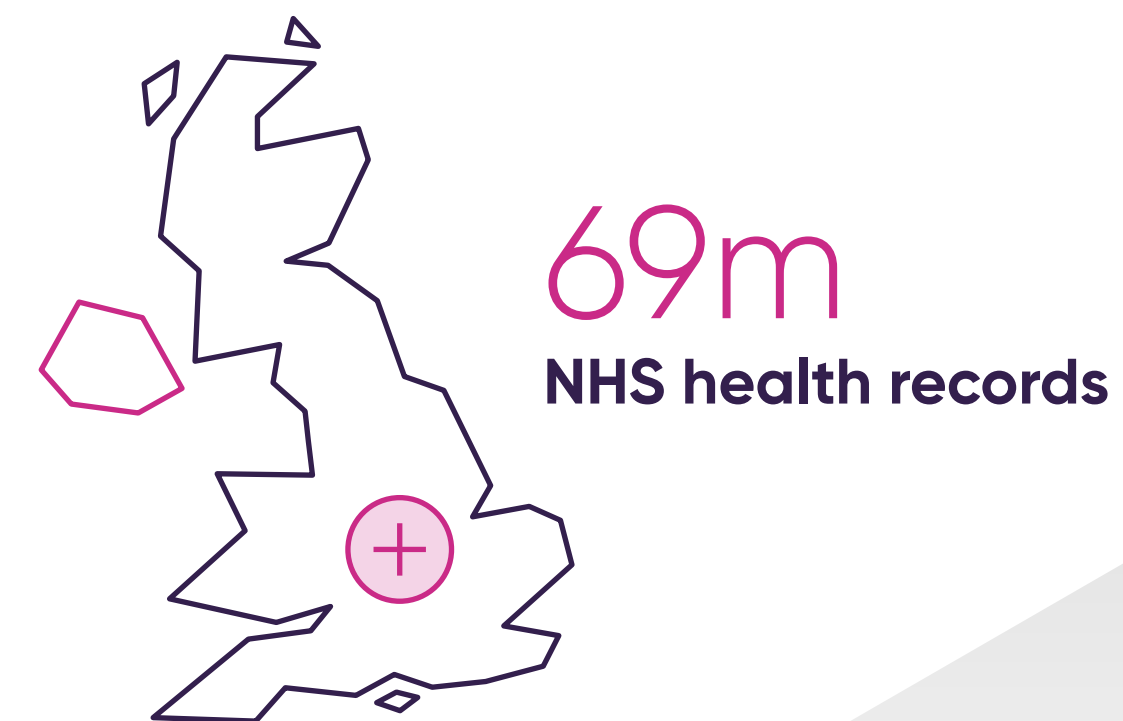


Reduce set-up times to **150 days**



£600m investment in Health Data Research Service

The UK has foundational strengths that competitors cannot easily replicate, including comprehensive longitudinal NHS health records covering 69 million people and a proven capacity to mobilise at scale, as shown through the COVID-19 vaccine trials. Yet the UK's potential to harness these unique assets to support commercial clinical trials delivery is still underutilised, with multiple barriers preventing systematic alignment of protocol requirements, NHS data assets and trial delivery systems.



This report examines challenges facing recruitment into UK commercial trials, the status of existing data-enabled clinical trial services, and issues with harnessing NHS data to support commercial trial delivery at scale. Based on extensive consultation with pharmaceutical companies, data service providers, research delivery networks and NHS trusts, the report makes the case for a new UK-wide data-enabled clinical trials function aligned to industry, delivery system and patient needs. A function designed to meet the rigorous requirements of commercial trials would, by its nature, serve the needs of non-commercial research, ensuring benefits across the entire UK clinical research ecosystem.

The report sets out a clear blueprint through seven recommendations to help inform and shape the design of the function, maximising the potential of the HDRS and strengths of the existing UK trials delivery infrastructure. The recommendations provide a pathway for translating the UK's structural advantages into predictable and efficient trial delivery.

The result would deliver:

- **the government's ambitions for UK commercial research**
- **secure international investment and;**
- **expand opportunities for UK patients to take part in trials of the latest innovative treatments**



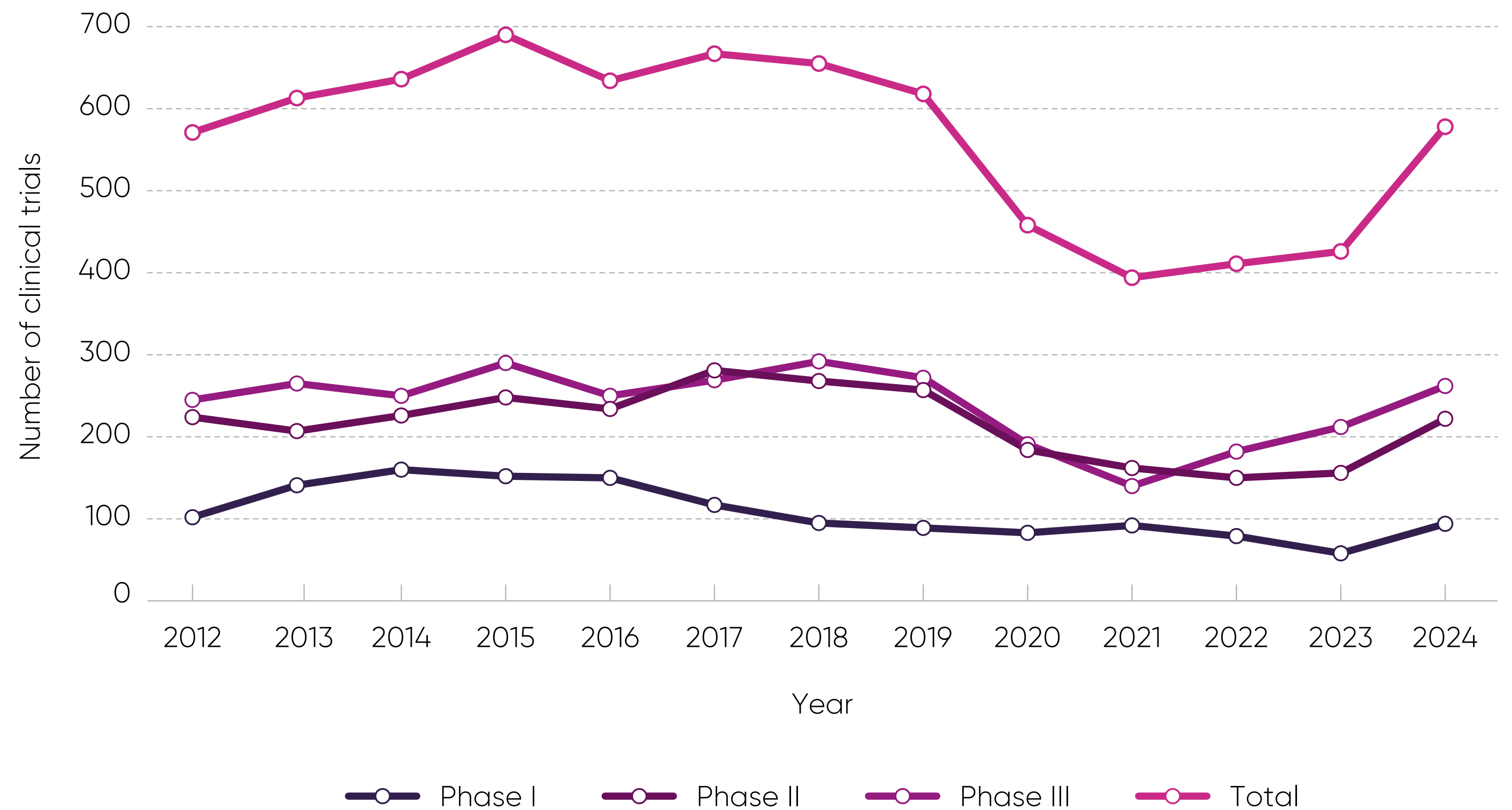
The trial recruitment challenge



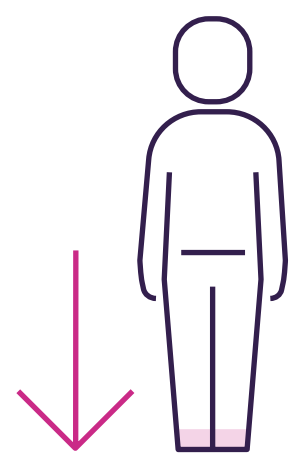


Following a period of post Covid pandemic decline in the number of UK-based commercial trials, there has been a resurgence in industry trials placed in the UK. In 2024, the number of commercial clinical trials initiated in the UK rose by 36 per cent above the previous year (Figure 1), with the UK climbing two places to sixth in global competitiveness ranking for all phase III trials initiated in 2024.² This increase is a positive indication, that if conditions are conducive for efficient commercial trial delivery, global companies are open to choosing the UK as a trial location.

Figure 1: Number of UK pharmaceutical industry interventional clinical trials initiated per year, by phase.²

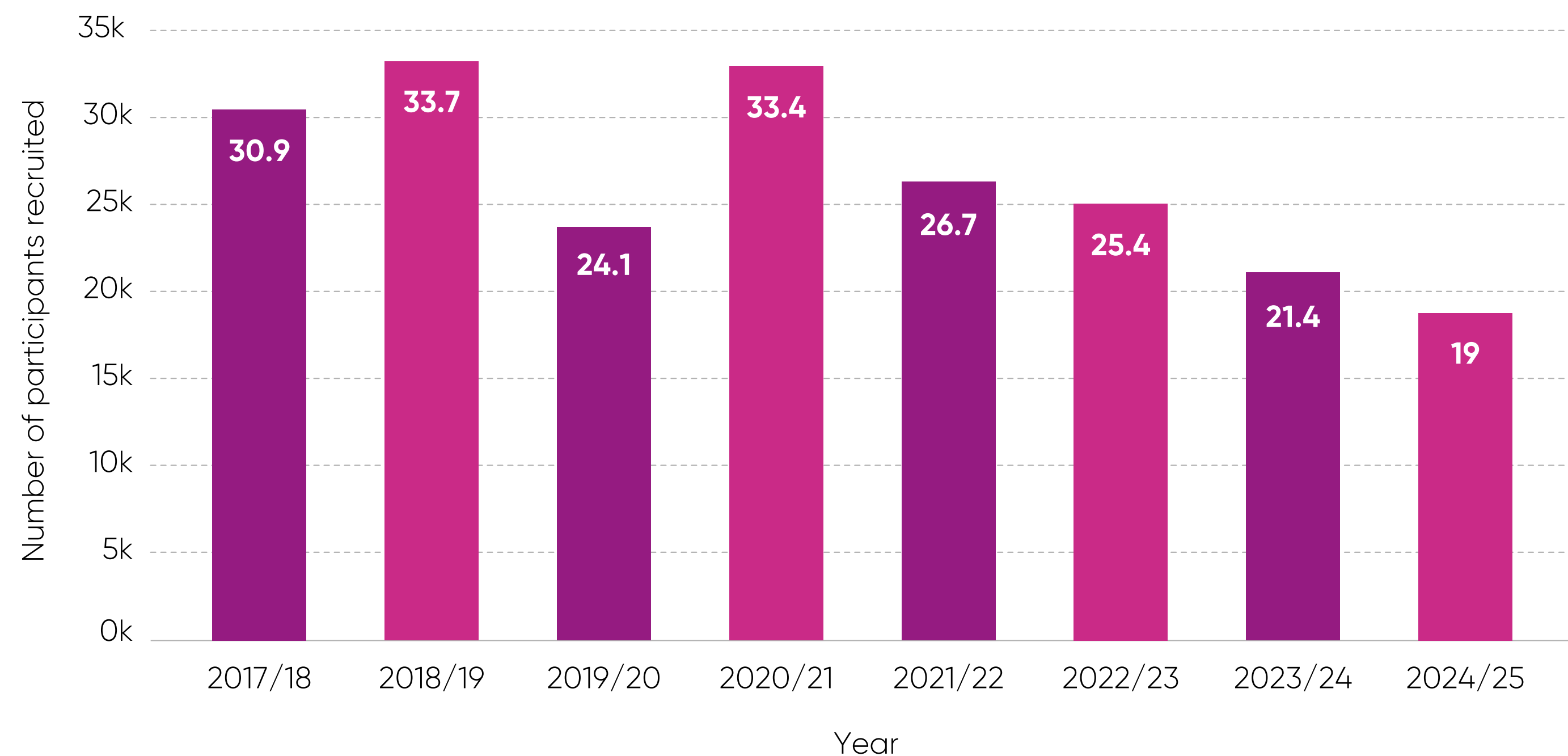


Since the Covid pandemic, the total number of participants in all interventional research studies in England in 2024/25 was almost double participation rates prior to the pandemic, demonstrating a clear willingness on behalf of patients and the public to take part in research studies. Unfortunately this overall increase was only observed for non-commercial studies, with participation in industry sponsored trials declining year on year to current levels, which are the lowest since 2017/18 (Figure 2).² As a result of this disparity, only 3.4 per cent of all participants in interventional research studies, are recruited to industry trials.



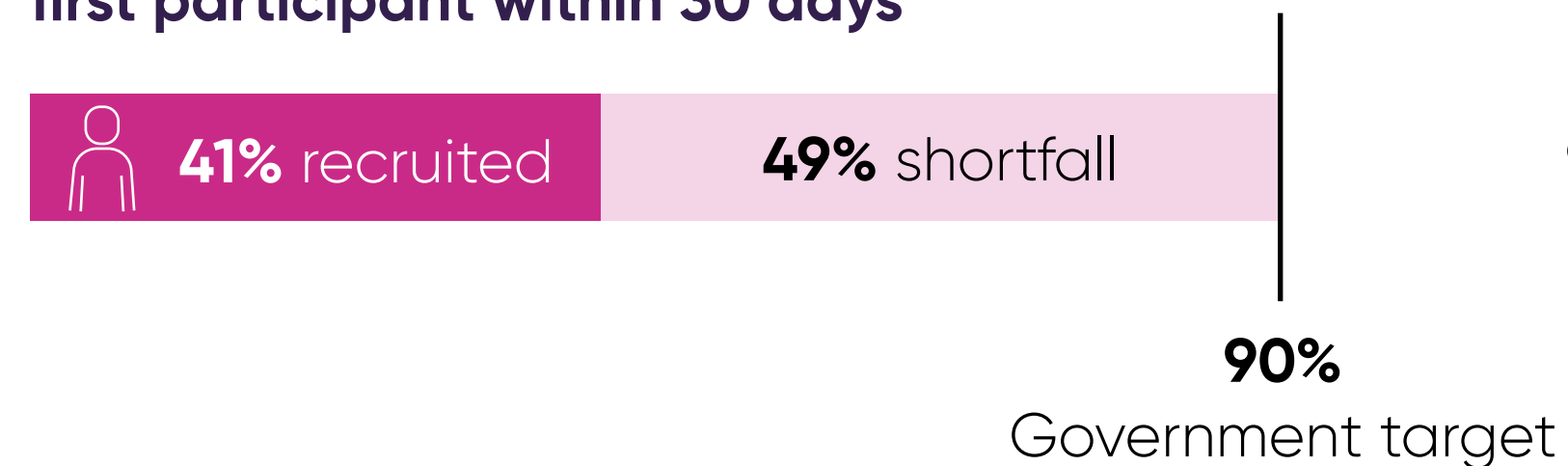
Only **3.4%** of all trial participants are in industry clinical trials

Figure 2: Number of participants recruited to interventional industry studies in the UK per year from 2017/18 to 2024/25.²



Between March 2024 and February 2025, just over a quarter of industry trials were open to recruitment within the 60-day government target, with a great variance in performance between sites. During this period, only 41 per cent of commercial trials recruited their first participant within 30 days of opening, a 49 percentage point gap from the government's 90 per cent target.^{2,7} These challenges are compounded by sites consistently setting uncompetitive recruitment targets, which are subsequently missed, despite strong patient willingness to participate and clear government commitments to bolster commercial trial activity.²

In 2024/25 commercial trials recruiting **first participant within 30 days**



Global competition to host commercial clinical trials is intensifying, creating urgency to capitalise on UK advantages. Spain increased industry phase I trial initiations by 133 per cent between 2022 and 2023, overtaking the UK as Europe's leading country for early-phase research.² China is continuing its rapid growth in activity and is on track to become the dominant country for trial initiations across all phases.

The rise in number of commercial trials and the overall increase in participants in research studies, represents a clear opportunity to reverse the declining participation in industry trials in the UK and offer more patients the chance to be involved in trials of the latest innovative therapies. Concerted action is therefore required to seize these opportunities and significantly boost recruitment and grow the UK's share of global commercial trials.



Shortfalls with current recruitment processes





Current recruitment methods are inefficient

Screen failure rates can reach 90 per cent in some therapeutic areas, meaning that out of one hundred participants contacted, only 10 are eligible for recruitment. This mismatch stems from manual chart reviews, local record searches and caseload estimates of potential participants that lack systematic data validation. From discussions with a range of pharmaceutical companies, current estimates are that roughly one in 10 UK sites recruit no participants, while 30–50 per cent recruit only one or two participants. This is incredibly costly and inefficient for everyone involved.

To secure trial allocations, UK affiliates of global pharmaceutical companies need credible feasibility estimates that accurately predict recruitment figures that are obtainable. Sponsors default to familiar sites and markets with proven reliability. It is therefore essential that the UK can consistently estimate a realistic number of participants eligible to enrol at sites.

Current recruitment methods pre-screen for potential trial participants without using the right breadth or depth of data needed to remove those who are ineligible based on their medical history. Often available information sources are not up to date, despite timeliness being a key requirement for around 40 per cent of protocol eligibility

criteria.⁸ Coding errors and data-quality issues can further compound the problem, if staff at sites are not familiar with these issues.^{9,10} As a result, many patients are invited to undergo screening assessments, only to be told they are ineligible. In addition to creating frustration among patients who have been turned away, high screen failure rates place an added burden on already stretch NHS services. Meanwhile, other patients who may be eligible to enrol in a trial have not been identified and invited for screening.



1/10
UK sites recruit
no participants



30–50%
of sites recruit **only one**
or two participants



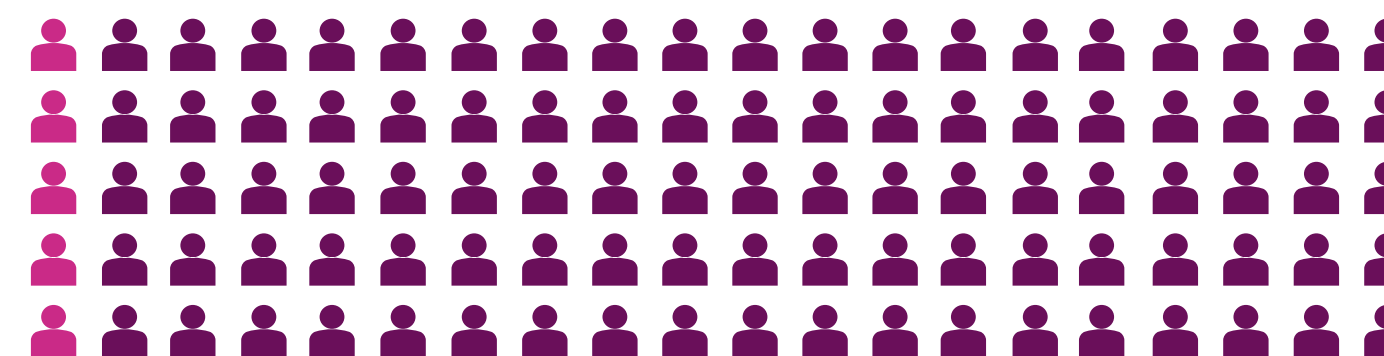
NHS data – an underused resource

The NHS holds comprehensive longitudinal health records on the UK's 69 million population, capturing demographics, diagnoses, prescribing, test results and secondary care episode data.^{11, 12, 13} Currently the potential of this data is not being harnessed to expedite clinical trials delivery at scale.

Approximately 90 per cent of healthcare contacts occur within primary care,^{14,15} where data on chronic disease management and some acute episodes, general practitioner (GP) prescribing, referrals and test results exist in searchable structured formats. This highlights both the importance and significant

opportunities offered by using longitudinal GP data to find individuals who are suitable for a clinical trial, based on their medical history.

Currently, fewer than 5 per cent of UK trials utilise routinely collected datasets to support delivery.^{16,17} This is partly because the UK data landscape is highly fragmented, with some patient data extracted by numerous databases, and some data only accessible to a limited number of researchers.¹⁸ As a consequence, the potential of primary and secondary care data to support targeted, scalable commercial trial recruitment remains untapped.



<5%
of trials use
health records



Fragmented approach using NHS data to support trial recruitment

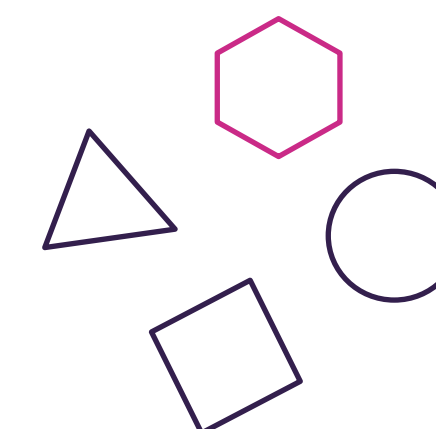
Multiple public and private data-enabled trial services exist in the UK. Each operates independently and serves different users or geographical locations based on data they have access to and their relationships with different parts of the trial delivery system.

National platforms, such as NHS DigiTrials in England, use secondary care datasets, including Hospital Episode Statistics (HES), to support recruitment for trials with broad inclusion criteria. However, lack of access to primary care data, lengthy agreement processes, and insufficient integration with trial delivery systems constrain its utility to support delivery of complex time-sensitive pharmaceutical industry trials.

The Speedy Patient Recruitment INto Trials (SPRINT) service run by the Clinical Practice Research Datalink (CPRD) is aimed at delivering commercial pharmaceutical trials. The service uses primary care data from 30 per cent of GP practices across the UK to model pharmaceutical industry protocols and works with its network of GP practices to deliver these trials. However, SPRINT's utility is constrained by lack of comprehensive GP coverage and limited access to secondary care data.

Accurate UK-wide feasibility assessment and support for locating the right participants who are suitable for recruitment regardless of their UK location requires access to both GP and secondary care data. The two government-funded services, CPRD SPRINT and NHS DigiTrials, draw on either primary or secondary data sources and so are unable to provide comprehensive geomapping of potential participants at trial sites across UK primary and secondary care settings.

Other national data platforms demonstrate what can be achieved at a smaller population scale. Wales's Secure Anonymised Information Linkage (SAIL) Databank links 100 per cent of Welsh secondary care records with 90 per cent of primary care data, alongside prescribing data, screening programmes and administrative datasets. SAIL operates within a geographically confined setting, and currently only offers trial feasibility services in Wales, which has limited impact for delivery of UK-wide industry trials.



Current services are **fragmented**

Regional services in England, like Northwest eHealth, have developed sophisticated integrated approaches, combining primary care data with secondary care data through their FARSITE and ConneXon platforms to support both feasibility and data-enabled recruitment. These services depend on individual GP practice engagement and work successfully at regional scale, but because they are regionally constrained, cannot be extended into the UK-wide coverage that commercial trial sponsors need.

Some local hospitals offer data-enabled feasibility and recruitment support services. Local services can access rich hospital data, including pathology reports, diagnostic imaging and hospital prescribing data. Despite their richness, these datasets are siloed within individual hospital systems and cannot be readily shared with other sites to support UK delivery. Invariably the datasets do not include GP data, which is vital

for understanding a patient's medical history and suitability for a trial.

Therapeutic area-specific registries deliver the depth and data quality necessary to enable feasibility assessments and participant identification for trials in these specific conditions. Therapeutic registries collect detailed clinical data from engaged patient groups, but this information is often not linked to other key NHS datasets, such as GP records, which provide a more comprehensive medical history. This 'deep but narrow' characteristic means most registries cannot scale beyond their specific disease focus and may require additional data to supplement patient identification. For instance, a registry that can find multiple sclerosis patients may not include information about their cardiovascular risk factors and will not include all medications a patient has been prescribed in primary care, which may exclude a patient from being eligible for a trial.



Creating a competitive UK-wide offering



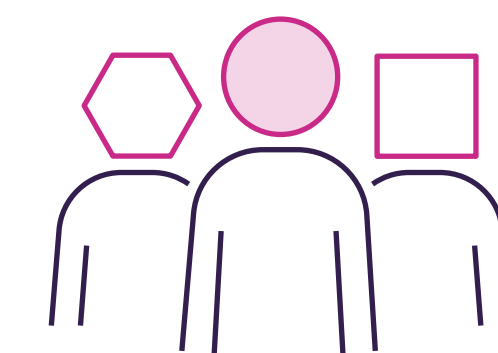


Clinical trials account for about 50 per cent of a pharmaceutical company's research and development spend.¹⁹ International competition to host commercial trials is growing as the widespread benefits trials bring to the economy and health outcomes are increasingly recognised.²⁰ Consequently, many countries are adopting a national or unified approach to incentivise industry investment.²¹

The potential for NHS data to accelerate trial delivery has been realised by multiple organisations that offer data-enabled services to different parts of the trial ecosystem. These services have grown organically or are privately funded with different governance arrangements, meaning it is not possible or practical to combine or retrofit their operations. Despite the enormous potential for NHS data to transform commercial trial delivery, to date there has been no UK-wide strategic approach to discerning how NHS data

could enhance commercial trial recruitment at scale, directly benefiting the NHS trial delivery infrastructure where sponsors place their trials.

Delivering trials that include participants representative of the patient population that will ultimately benefit from the medicine is key to developing robust evidence on the safety and efficacy of new medicines.¹⁸ Identifying representative populations through current recruitment methods creates substantial challenges. Using privacy-preserving eligibility searches of NHS records would not only accelerate the recruitment process but would support recruitment of a more representative cohort.



NHS data supports **more representative trials**

The £600 million investment in establishing an HDRS⁶ provides, for the first time, a realistic possibility of creating a compelling UK offering to expedite recruitment into phase II-IV commercial trials. Fundamental to achieving this ambition is understanding from a pharmaceutical industry perspective, where the greatest gains could be made and how UK national data assets could be seamlessly connected with the existing trials infrastructure to improve trial delivery.

The resulting recommendation from the ABPI and member pharmaceutical companies is to establish a dedicated UK-wide data-enabled clinical trial function that protects patient confidentiality and uses comprehensive NHS data to provide robust feasibility assessments and targeted participant recruitment to expedite trial delivery at scale across the UK.

1

Recommendation

Establish a UK-wide data-enabled clinical trials function built to meet industry needs.

The function should accelerate phase II-IV trial feasibility and recruitment to commercial interventional studies, operating at scale across all four nations. Investment in the HDRS provides the foundation for creating this capability.

Industry requirements





The pharmaceutical industry operates within global decision-making frameworks in which UK teams must compete for trial allocations against other markets. Understanding what commercial sponsors need to justify placing trials in the UK, and how a UK-wide data-enabled clinical trial service could deliver these requirements, is fundamental to realising the UK's competitive potential.

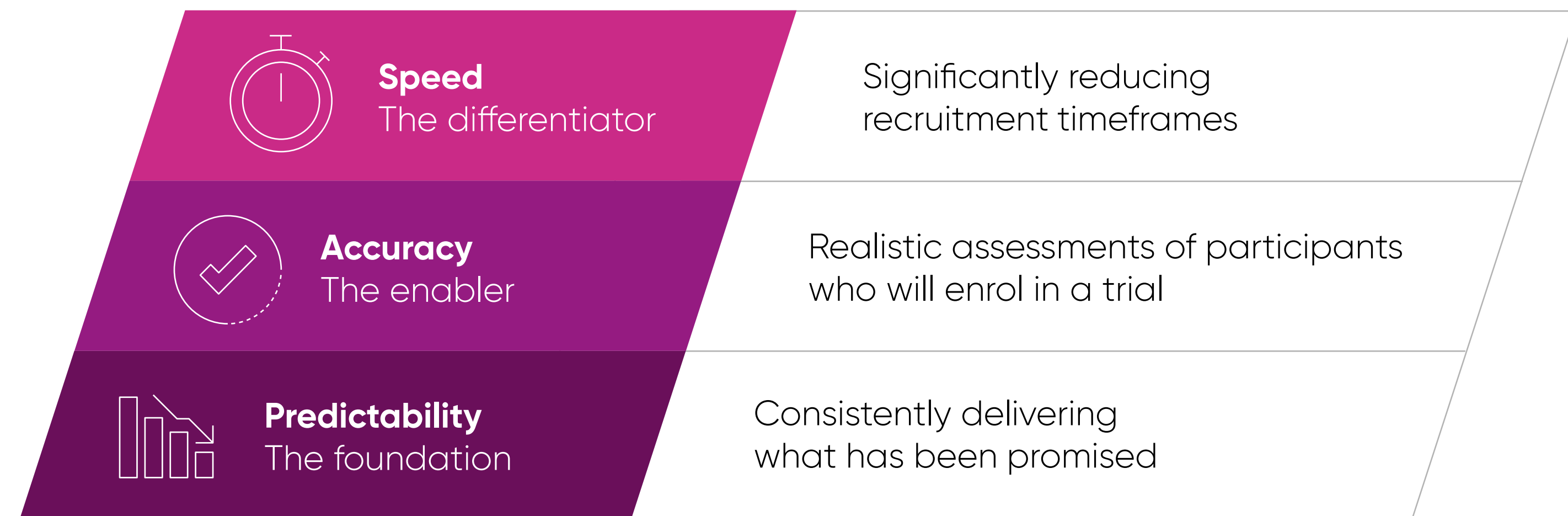
Drawing on direct engagement with the pharmaceutical industry and NHS trial infrastructure, this section defines the operating principles for UK-wide data-enabled clinical trials services that would support the key requirements for pharmaceutical industry trial delivery. Findings are based on a desk-based review of clinical trials performance data and academic literature on eligibility criteria; semi-structured interviews with 20 stakeholders spanning UK-wide data providers

(for example, CPRD, NHS DigiTrials, NorthWest eHealth, SAIL Databank, Research Data Scotland), research delivery networks and NHS trusts; and two Task and Finish Group workshops with ABPI members (see the Appendix for further detail).

Operating principles

Three themes consistently emerged as core operating principles for a data-enabled clinical trials function (see Figure 3).

Figure 3: Core operating principles of the data-enabled clinical trial function.



- Predictability.** Sponsors return to sites and markets they trust. Reliable delivery justifies ongoing investment in the UK and will help UK teams to secure larger trial allocations in the future. The service must consistently prove it can support sites across the UK to deliver what has been promised.
- Accuracy.** Feasibility assessments must provide an accurate estimate of eligible populations that will translate into participants who enrol in a trial. Current screen failure rates can approach 90 per cent in some therapeutic areas due to inaccurate feasibility assessments and inability to target the right patients to invite for screening. According to ABPI members, reducing screen failures to 50 per cent by more accurate feasibility assessments and participant identification would nearly halve screening costs and site burden. In addition, it would reduce patient time wasted in attending screening, only to be told they are ineligible for a trial, based on medical information already captured in their records.
- Speed.** Global teams need feasibility responses within two weeks, or they allocate trials elsewhere. Rapid participant recruitment will increase the likelihood of UK sites successfully delivering their trial quotas within faster timeframes. Providing each site with a pre-screened list of pseudonymised potential participants during the regulatory approvals process would speed up subsequent identification of eligible participants at sites, enabling recruitment to begin as soon as regulatory approval has been received.

2

Recommendation

Core operating principles must be predictability, accuracy and speed

Services must consistently deliver what is promised, provide accurate timely feasibility assessments based on clinical information that translates into participants enrolled, and speed up recruitment by providing sites with pre-screened lists of potential participants during the regulatory approvals process.

Minimum datasets required

According to ABPI member companies consulted, information contained within structured GP data linked to Secondary Uses Service (SUS)ⁱ data in England or equivalent national secondary care datasets in the devolved nations, encompass approximately 60 per cent of selection criteria in most industry trial protocols, excluding oncology and rare diseases. GP records capture 90 per cent of patient contacts with the healthcare system,^{14,15} and contain longitudinal data on chronic diseases and acute conditions, prescribing histories, laboratory test results, diagnoses and referrals. This information supports identification of patients who may be suitable for trials in many chronic conditions. However, for most trials, comprehensive GP data alone is insufficient and must be supplemented with national secondary care data. Together these datasets provide an

overview of patients' interactions with GP and hospital settings, enabling more reliable targeting of participants who may be eligible for a wide range of commercial trials.

The categories of key information captured within GP and SUS datasets is illustrated in Figure 4 below.

Figure 4: Key variables within the minimum viable dataset.

Structured GP data provides	Secondary Uses Service data adds ⁱⁱ
Demographics: e.g., age, sex, ethnicity, deprivation indices	Hospital admissions
Diagnoses: e.g., SNOMED codes for chronic conditions (diabetes, asthma, COPD, hypertension, cardiovascular disease)	Outpatient attendances and specialty
Prescribing: e.g., complete medication histories with dates, including dose and frequency where recorded	A&E presentations
Laboratory results: e.g., HbA1c, lipid panels, renal function, liver function, full blood counts	Procedures
Physiological measurements: e.g., blood pressure, BMI, smoking status	Hospital episode duration and timings
Immunisations and screening participation	

ⁱ The SUS serves as England's single comprehensive repository for healthcare data, enabling reporting and analysis that supports NHS service delivery. Most national level datasets in England, such as the Hospital Episodes Statistics, are extracted from SUS data.

ⁱⁱ Information captured within national secondary care data in each devolved nation will differ from the SUS.

In addition to extensive coverage of patients' medical histories and interactions with the health service, these two data sources exist in structured, queryable formats. For these datasets to be of value for facilitating recruitment into industry trials, the data need to be updated with sufficient frequency to align with a range of protocol selection criteria. As such, GP data will need to be updated weekly, as a minimum, while the current monthly SUS data updates, with validation typically lagging by approximately six months, should suffice to support data-enabled trials services.

At a minimum, comprehensive UK-wide GP data coverage, linked to national secondary care data and mortality data is required to provide meaningful feasibility assessments and to locate participants for commercial pharmaceutical trials based around delivery sites across the UK. Adopting this minimal essential dataset approach will lead to more targeted, efficient recruitment and significantly reduce costly and time-consuming screen failure rates.

3

Recommendation

The minimum dataset should include UK-wide GP data linked to Secondary Uses Service data in England, or equivalent secondary care data in devolved nations, and mortality data

Comprehensive UK GP data coverage, linked to national secondary care and mortality data, which is updated on a frequency that aligns with pharmaceutical trial protocol selection criteria, should form the minimum viable dataset. Together these datasets can encompass 60 per cent of protocol criteria (excluding oncology and rare diseases), which enables more reliable targeting of participants, reducing costly and time-consuming screen failures.



Essential services

In discussions ABPI member companies identified the following three services that should be provided to support UK-wide data-enabled clinical trials, based on access to timely UK-wide GP and secondary care data, that would accelerate commercial trial delivery and increase global competitiveness of the UK.

Service 1

Feasibility assessments with self-service option

Currently it is not possible to obtain accurate feasibility assessments of the number of patients across the UK who may be eligible for a trial, due to the inability to access comprehensive population-wide clinical data. Consequently, UK feasibility assessments for trials are either generated locally at a selection of sites from current caseload lists and subsequently collated, derived nationally using tools based on proxy indicators of real patient numbers or extrapolated from access to incomplete patient databases. Many services in the UK and internationally still rely on form-based requests to interrogate patient records, even for straightforward questions about the number of patients fitting within the inclusion and exclusion criteria of a trial.²²

By centrally matching comprehensive GP data linked to national secondary care data to key selection criteria in the trial protocol, it will be

possible to rapidly obtain more accurate UK-wide feasibility assessments. This will enable UK affiliates to bid for a UK trial allocation and be confident that these estimates will subsequently translate into patients enrolled. It is not uncommon for companies needing to submit feasibility assessments within a two-week timeframe, so accuracy combined with speed is vital. Equally important, by using this approach companies can more reliably assess whether there are sufficient patients in the UK to take part in a trial and not waste effort bidding for trials that will not deliver. More accurate feasibility assessments that reliably predict the number of participants who could enrol in a trial will also increase efficiency in the system by enabling NHS staff at sites and industry to focus their resources on trials that have the greatest likelihood of successful delivery.

In addition to a managed service offering, carried out by staff using pseudonymised GP and secondary care data, the feasibility service

should include a self-service option, enabling industry researchers to also directly run queries on these pseudonymised datasets. The use of pseudonymised data and privacy techniques ensures that patient identities are concealed from all researchers using the feasibility service.

Service 2



Geomapping and site selection

The UK could gain further competitive advantage by understanding where relevant potential participants are physically located across the UK and therefore which trial delivery sites may hold the greatest recruitment potential. Trial site selection could consequently be based not only on site's capacity and performance information, as is currently the case, but also on geomapping sites to locations with the highest relevant patient populations. Using a more evidence-based approach for site selection would drive efficiencies

by placing trials in areas with greater certainty of delivery. This would reduce costly site set-ups and wasted NHS staff time on sites that fail to recruit or only recruit very small participant numbers. Geomapping would be particularly advantageous for rare disease trials where locating sufficient eligible patients within site catchment areas is an even greater challenge.

Service 3



Recruitment pre-screening

Although there is increasing recognition of the value of using patient records in pre-screening potentially eligible participants, currently trial sites in secondary care must rely on access to local available hospital datasets, usually in the absence of GP records. The current practice of each site carrying out bespoke searches on incomplete datasets that lack most of patients' medical histories is inefficient and labour intensive.

Significant time and cost savings can be made by centralised pre-screening of linked GP and national secondary care datasets, such as SUS in England. For each trial, the data could be modelled to the industry protocol once, using methods that safeguard patient confidentiality, and the resulting pre-screened pseudonymised lists of potentially suitable participants could be used by sites for more detailed local clinical review. Supplying individual sites with a pre-screened list would not only reduce the burden of each site having to do searches for suitable individuals from scratch for every study, but it would expedite trial startup if the centralised lists were generated during the regulatory approval process. This would free up resources at sites, enabling NHS staff to focus on selecting the most appropriate participants and inviting them for screening immediately after the site receives regulatory approval for the study to commence.

4

Recommendation

Essential services should include centralised feasibility assessments, geomapping and site selection, and recruitment pre-screening services

Providing centralised services based on comprehensive pseudonymised population coverage enables more reliable feasibility predictions of trial participants and site selection, informed by where relevant patients are geolocated. Centralised pre-screening of potential participants for subsequent clinical review at sites reduces the burden on each site and increases efficiencies in trial startup processes and times.

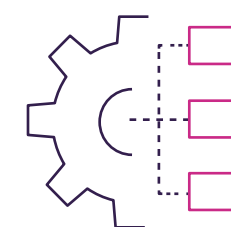
Additional service offerings

Additional services could be offered to supplement the essential three functions of the data-enabled clinical trials service.



Offering 1

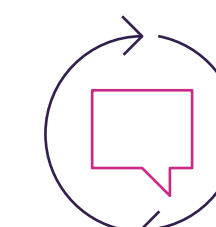
Protocol design consultation



A protocol design consultation service could offer sponsors evidence-based guidance on which protocol inclusion and exclusion criteria could be reliably modelled using pseudonymised, structured NHS data, which would require manual abstraction, and which are not available in clinical records. For some companies, there are limited opportunities to adjust pre-determined protocols that have been optimised for US healthcare delivery models. However, where there is an opportunity to shape protocol development to be more aligned with UK systems, this service offering, in conjunction with existing services that inform UK protocol design, could contribute to more efficient downstream feasibility assessments and rapid patient recruitment.

Offering 2

Feedback loop to support continuous service improvements



The ability to continually refine and improve services through a feedback loop will underpin successful evolution, ensuring services are fit for purpose to meet the needs of trial participants, delivery infrastructure and industry sponsors. Analysis of the utility of pre-screened, pseudonymised potential participants lists for facilitating identification of suitable participants at sites to invite for screening, and subsequent conversion rates of screened individuals into enrolled participants, should be carried out on an ongoing basis. Monitoring the translation rates of feasibility assessments into real-world participants screened and enrolled at sites will be crucial for validating and continually improving centralised modelling of pseudonymised national datasets against protocol inclusion and exclusion criteria.

Service design





UK-wide, centralised data-enabled function that integrates with the UK clinical trial delivery infrastructure

A new data-enabled clinical trial function must enhance the capability of the UK's existing clinical trials delivery infrastructure and not create a standalone data service operating in isolation from established clinical trial workflows and delivery systems. It should seamlessly interface with research delivery networks, UK Clinical Research Delivery Centres, site infrastructure and trial coordination mechanisms to increase efficiencies and boost recruitment at secondary and primary care trial delivery sites across the UK. The function should provide a unified and coordinated UK-wide offering, carrying out centralised data searches on pseudonymised NHS data, modelled against industry protocol criteria, to support subsequent clinical validation, patient identification and recruitment by clinical trial delivery sites.

This integrated approach contrasts with current fragmentation, where industry navigates multiple disconnected services, approaching CPRD SPRINT, NHS DigiTrials, regional services, therapeutic area registries and trials networks separately. While each offers effective services within its remit, coordination across different services is limited. Sponsors lose time navigating different interfaces, governance requirements and timelines, resulting in trial sites receiving data from multiple sources without sufficient context.

The HDRS has a mandate to reduce fragmentation and increase availability of, and access to, NHS datasets for research. Given that three of the four stated priority datasets that the HDRS will initially make available – GP data, national secondary care data and mortality data – are the proposed minimal datasets to support data-enabled clinical trials services, it makes logical sense for a data-enabled clinical trials function to operate within the HDRS. This UK-wide data-enabled

capability housed within the HDRS would, for the first time, provide a comprehensive overview of UK recruitment potential and offer a single engagement point for sponsors, trial coordinating infrastructure and individual sites seeking support from data-enabled approaches.

The proposed architecture for an end to end service design maintains appropriate information governance boundaries between centralised pseudonymised feasibility assessments, site selection and pre-screening in the HDRS and subsequent patient identification following clinical review at NHS sites (see Figure 5). Pseudonymised datasets would be centrally matched against industry protocol inclusion and exclusion criteria within the HDRS to determine trial feasibility, and to geomap where the highest number of potential participants may be located to help inform site selection. Following more in-depth

data modelling and clinical interpretation against complex protocol eligibility criteria, a pre-screened potential participant lists would be generated and passed on to selected delivery sites within the NHS. The identity of individual patients would only be known to NHS staff.

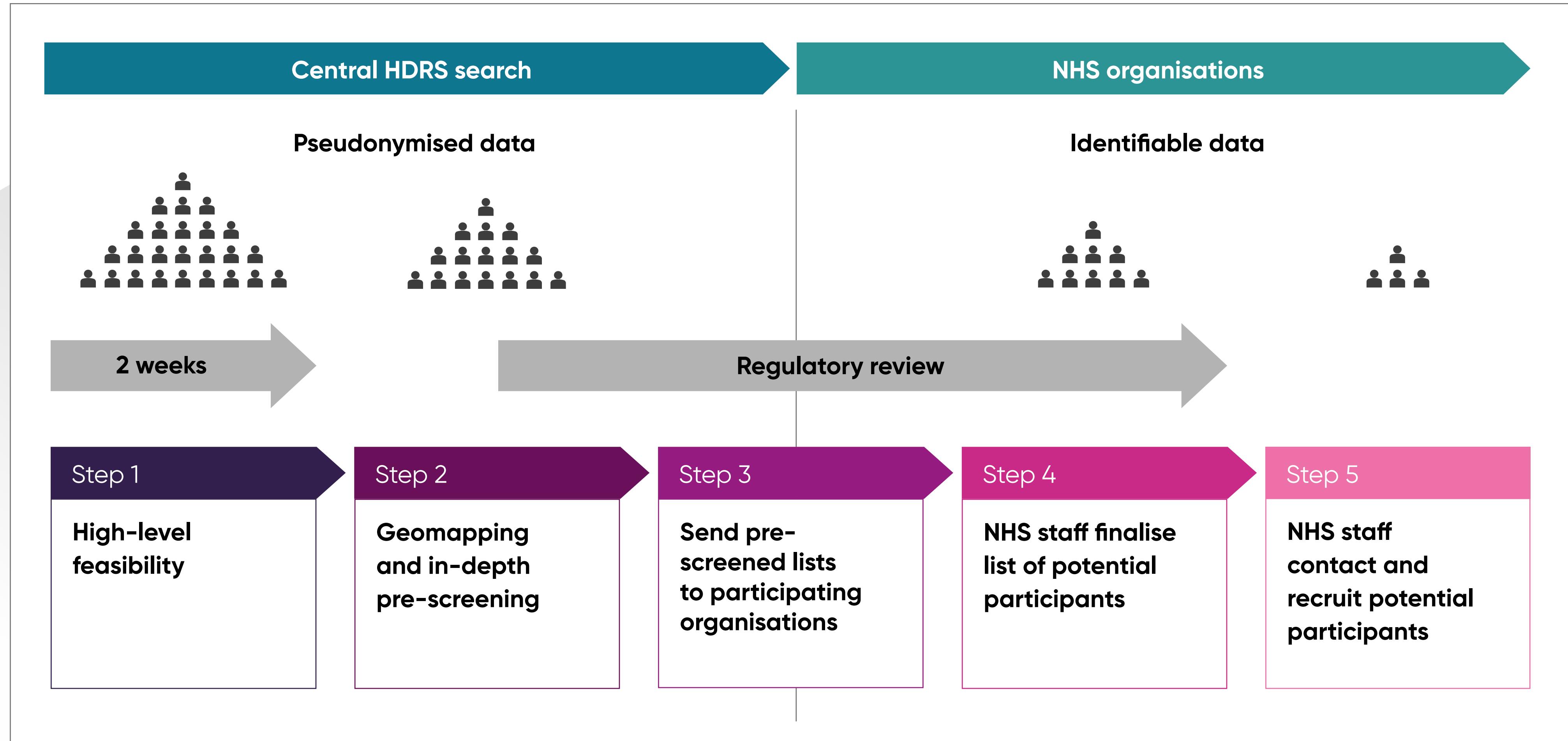
Trial sites would receive pre-screened, pseudonymised participant lists from the data-enabled trials function in the HDRS, accompanied by clear contextual information indicating which eligibility criteria have been modelled in data searches and which require further clinical review by local NHS site teams. The provision of pre-screened lists would negate the need for each delivery site to search available information from scratch, which would drive system efficiencies by saving valuable NHS resources and time. Local clinical teams at sites would then validate pre-screened lists and identify suitable individuals

to invite for screening when the regulatory approvals have been granted. Participants would be identified and contacted by appropriate people with a legitimate basis for access to identifiable health data. This process of matching patient eligibility criteria to their medical history prior to invitation reduces screen failures and frustration for patients, who are contacted only to subsequently find out they are not eligible based on existing information in their health records.

The centralised data-enabled clinical trial services provided by the HDRS should be appropriately staffed with researchers, such as epidemiologists and data scientists, as well as individuals with clinical knowledge who understand NHS clinical practice, coding conventions and data-quality issues, including when coding errors impact data interpretation.^{9,10,23,24}



Figure 5: UK-wide data-enabled clinical trials: from feasibility to recruitment.



5

Recommendation

Data-enabled services should be based within the HDRS and integrate with the existing UK clinical trial infrastructure to enhance system efficiencies and UK competitiveness

Providing centralised services based on comprehensive pseudonymised population coverage enables more reliable feasibility predictions of trial participants and site selection, informed by where relevant patients are geolocated. Centralised pre-screening of potential participants for subsequent clinical review at sites reduces the burden on each site and increases efficiencies in trial startup processes and times.



Targeting the right participants for screening





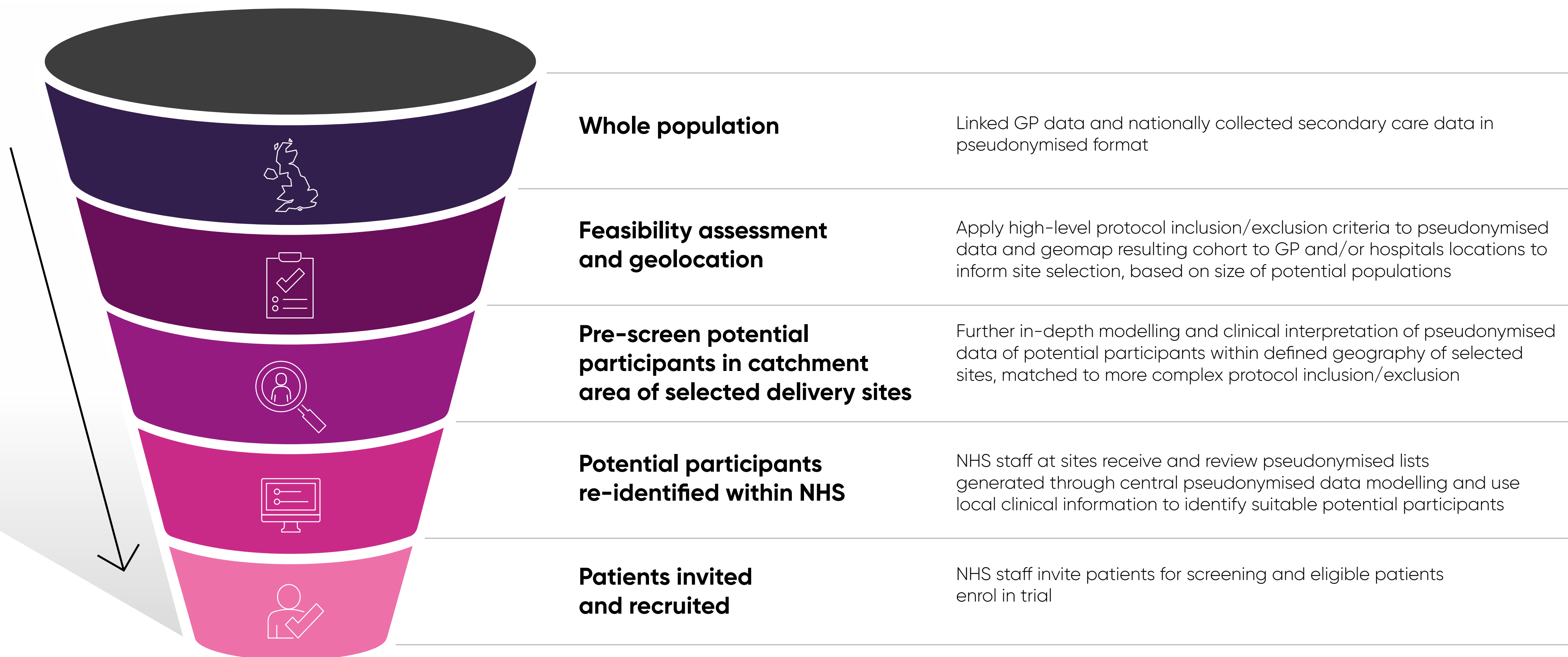
The efficiency funnel: from protocol to recruitment

Current methods for identifying participants to recruit to a trial often creates overwhelming volumes of false positives at the screening stage. Not only is this costly, but it is time consuming for all involved. The goal should be to increase efficiency by more effectively identifying suitable patients prior to inviting them to take part in a trial.

The operating model for the data-enabled clinical trials function should therefore follow a funnel efficiency approach, modelling NHS data to protocol inclusion and exclusion criteria where possible. Systematic application of information within the medical record will progressively narrow the population pool to a more targeted group of individuals with a higher likelihood of being eligible to participate in a trial, before inviting individuals for recruitment screening.



Figure 6: Efficiency funnel from feasibility to recruitment.



Developing trial protocol archetypes

While there is clear potential for improving recruitment efficiency and inclusivity by mapping trial eligibility to NHS data sources, the reliability of meaningfully translating the array of NHS data codes into protocol inclusion and exclusion criteria is not well understood. Existing data-enabled services have gained a wealth of experience interpreting pharmaceutical industry protocols. However, this knowledge largely remains siloed within these organisations.

To retain system knowledge and avoid unnecessary duplication of effort, it is recommended that a mapping exercise is carried out characterising applicability of NHS data to common phase II-IV selection criteria in the trial protocol. The aim would be to generate reusable protocol archetypes, organised by therapeutic area, documenting the criteria that can be reliably interpreted from coded data in the minimum recommended datasets, which could be extracted from specialist hospital data sources, and which would only be obtained at the screening stage. The mapping should capture diagnostic coding accuracy, data completeness and generate metadata on data quality, data velocity and known limitations of NHS data by therapeutic area, in particular in relation to common protocol temporal requirements.

This reusable evidence base would transform current methods of bespoke data modelling, which currently generate novel algorithms requiring clinical interpretation for every trial protocol, into a system that matches against validated archetypes, significantly reducing the workload for each new study. For instance, for cardiovascular trials that consistently require recent lipid panels, the archetype would capture which laboratory systems record this, with what completeness, at what update frequency.



6

Recommendation

Create a reusable resource of protocol archetypes by therapeutic area

To understand which eligibility criteria commonly used in industry phase II-IV trial protocols can be reliably mapped to NHS data, a mapping exercise across different NHS data sources should be carried out to generate protocol archetypes by therapeutic area. Creation of a central repository characterising the utility of NHS datasets for interpreting trial protocols would significantly reduce time taken to model each new protocol.



Implications for different therapeutic areas





Therapeutic areas most suited to modelling protocols against NHS datasets

The utility of using NHS data to facilitate clinical trial recruitment will vary by therapeutic area depending on which datasets the most relevant clinical information is recorded in, whether or to what extent salient information is captured, and how accessible these data sources are to researchers. Not all therapeutic areas will be equally suited to a data-enabled approach that is predicated on central modelling of national NHS datasets against protocol criteria, supplemented by local data sources at delivery sites.

Trials in therapeutic areas that most lend themselves to this two-tiered approach align with conditions where a large proportion of the

protocol criteria are routinely recorded in GP and national secondary care datasets. Many of the protocol inclusion and exclusion criteria for conditions commonly managed in primary care, such as cardiovascular, metabolic and respiratory conditions, can be mapped to diagnoses, prescribing, test results and interventions recorded in GP data, supplemented by national secondary care data. These two datasets can contribute to neurology and immunology protocol pre-screening, however, some of the trial selection criteria for these conditions depend on access to specialist hospital datasets that currently cannot be accessed centrally. For these trials, there would need to be greater emphasis on site delivery teams consulting relevant local data sources to augment the initial pre-screens carried out on national data held within the HDRS. The situation for oncology trials is different. Key eligibility

selection criteria for oncology trials are entirely dependent on information within datasets held in specialist oncology centres or individual trusts.

Understanding applicability of different datasets for supporting trials in various therapeutic indications, informed by development of trial protocol archetypes, will define the scope of where data-enabled clinical trials services can realistically support commercial trial recruitment. Annex Table 1 provides an overview of how modelling NHS data via the proposed centralised approach might support trial pre-screening in different therapeutic areas.



Oncology trials require a different approach

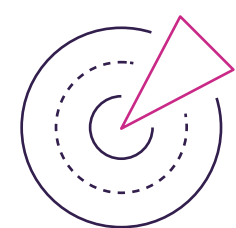
Oncology trial screen failure rates often reach 90 per cent because matching patients to the complex trial eligibility criteria requires access to data held within hospitals and specialist oncology centres in multiple different NHS data sources. These can include histopathology reports confirming tumour type and grade, radiology reports documenting tumour progression, reports recoding biomarkers present, and prescribing data containing treatment sequencing. This information is outside the scope of GP and national secondary care data sources and, consequently, a centralised modelling approach using these datasets has limited utility for oncology trial recruitment. Annex Table 2 provides further detail on oncology trial specific data requirements.

The most recent ABPI annual clinical trials report highlighted that oncology was the leading therapeutic area among UK pharmaceutical industry clinical trials, accounting for three in every 10 phase II and III trials initiated in 2024. Significant gains in UK competitiveness could be made through more effective identification of eligible patients for these trials. Due to the challenges of collating information from the range of specialist data sources required to support oncology protocol pre-screening, a dedicated strategic approach is needed to inform the specification for an oncology data-enabled clinical trials offering.

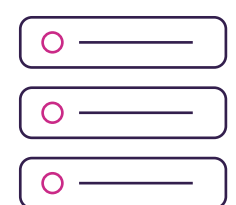
To understand the potential for data-enabled services to improve recruitment into commercial oncology trials, a network of four to five major tertiary oncology centres across the UK should initially be formed. The aim of the network would be to develop a standardised data framework for

broader rollout, by defining which datasets are of most value and what standards should be used, and to practically test data-sharing approaches across sites.

The proposed criteria for selecting network centres could be:



concentrated trial-eligible populations: major tertiary centres that have a high through-put of cancer patients who could be eligible for a wide range of trials



holding necessary data sources: centres holding relevant histopathology, radiology, genomics, and treatment datasets locally



existing research capabilities: already research active with clinical academics leading on numerous industry oncology trials.

7

Recommendation

Establish a dedicated oncology network across the four nations to explore the potential for a replicable, data-enabled approach for improving recruitment into oncology trials

Recruitment to oncology trials depends on accessing essential information in specialist hospital datasets. To understand whether oncology trials could benefit from a data-enabled approach, major tertiary oncology centres across the UK should work together to develop and test a data-sharing framework aimed at improving identification of cancer patients eligible for trials.

Measuring **success**





Factors that will determine impact

The primary driver for establishing a dedicated function to support data-enabled clinical trials at scale, is to increase UK competitiveness by improving efficiency of recruitment into commercial trials. Evaluating the impact of data-enabled services should therefore be based on measures of timely and accurate feasibility assessments and subsequent rapid, targeted participant recruitment by the trials delivery system. The following performance indicators are proposed:

- **Feasibility assessments within two weeks, based on an overview of whole population national datasets.** Provision of feasibility assessments within two-week timeframes will improve the UK chances of being selected in global allocations of trials. Accurate as well as timely assessments will be critical to instilling confidence among global pharmaceutical companies in choosing the UK as a location. To evidence whether the data-enabled feasibility service is improving quality of feasibility predictions, the accuracy of estimates generated by this service that subsequently translate into real patients enrolled in trials, compared with current baseline estimates using incomplete or proxy data, should be measured.



- **Reduced proportion of non-recruiting sites by geomapping sites with the highest recruitment potential.** Reducing the number of sites that fail to recruit any participants or recruit far fewer than their projected numbers will lead to cost savings for sponsors and will reduce time wasted by the NHS setting up sites that do not deliver. Comparing the proportion of non-recruiting sites in current and recent studies to trials where sites are selected through evidence-based geomapping will measure the performance of the geomapping service.
- **More efficient and targeted recruitment through centralised evidence-based pre-screening.** Supplying delivery sites with a pre-screened lists of pseudonymised participants who may be suitable for a trial should enable NHS staff at these sites to concentrate their time on reviewing a short-list of patients. This two-

tiered approach drawing on existing information within NHS data should lead to fewer screen failures among those invited for screening, resulting in a higher proportion of successfully enrolled participants. Conversion rates of invited participants to those invited for screening to similar types of trials using current methods can be used to evaluate the data-enabled pre-screening service performance. The approach of using centralised pre-screening searches should also facilitate faster trial set-ups, enabling sites to more rapidly identify suitable participants to invite for screening as soon as they receive regulatory approval. Time taken post-regulatory approval to recruit the first patient, as well as time to reach the target number of patients, compared with current timeframes could be measured to determine the data-enabled pre-screening service's contribution to expediting commercial trial delivery.

Additional benefits

Other benefits that a data-enabled clinical trials function should bring include time saved by NHS staff at each site carrying out pre-screening from scratch, and increased inclusivity and representativeness in clinical trials due to the ability to review all potential suitable trial participants. These benefits may be more difficult to quantify, as relevant metrics are not routinely collected. The Health Research Authority, government and trial delivery partners are currently developing metrics to assess diversity and inclusion in trials. When implemented, these metrics may provide a suitable baseline to assess whether data-enabled approaches contribute to the shared goal of increasing representativeness and inclusion in trials.



Conclusion





Government has recognised the immense value that industry clinical trials bring to patients, the NHS and the economy and the need to make the UK a more appealing and reliable place to deliver industry global trials. As a consequence, ambitious government targets have been set to reduce commercial trial set-up times and significantly increase commercial trial recruitment. To achieve these goals, a series of measures are being implemented to streamline processes, reward good practice and increase performance transparency across the delivery system. Undoubtedly, these actions will lead to demonstrable improvements that should make the UK a more attractive destination for global industry.

However, for the UK to seriously compete with preferred European markets in Spain, France and Germany, globally with China and the US, and with emerging markets such as South Korea and Canada, a transformational change is required.

Fortunately, the UK has an opportunity to enact this transformational change by strategically aligning the two major government commitments to establish an HDRS and to improve the environment for delivering commercial clinical trials. NHS records contain longitudinal medical information on the UK's population. If successfully implemented, creation of a data-enabled clinical trials capability that harnesses this globally unique NHS asset, offers the UK a competitive advantage in accelerating recruitment into commercial trials via the UK's established trial delivery infrastructure.

Public trust in responsible use of health data in research is paramount. The service design proposed for data-enabled clinical trials is protects patient confidentiality, with only staff at participating NHS organisations being able to access identifiable patient data. The proposed central data feasibility assessments and recruitment pre-screening services would

be carried out on pseudonymised data within the HDRS, where researchers do not have access to identifiable patient information. Use of pseudonymised data to support clinical trials should be part of HDRS planned discussions with patients, the public and healthcare professionals, which are aimed at building trust in transparent use of patient data for research.

Although the rationale for this proposal is to create industry-standard data-enabled services that will improve speed and predictability of industry trial delivery, the services do not have to exclusively support commercial studies. Industry clinical trials must comply with the highest regulatory standards and deliver to globally competitive timeframes. In setting the bar high to establish an operational standard required to fulfil industry's requirements,

the resulting service capabilities should also be fit for purpose to support the needs of non-commercial trials, resulting in benefits for the whole UK trial ecosystem.

The seven recommendations outlined in this report to drive UK-wide data-enabled clinical trials capability, would unlock the potential of NHS data to significantly increase efficiency of recruitment into commercial clinical trials. Critical to success is centrally collected GP data, as is the need to seamlessly integrate the data service with the existing trial infrastructure. Implementing the full suite of proposed UK-wide data-enabled services is a major endeavour and will take time to be fully operational. In the meanwhile, a modular stepwise approach, building on existing data-enabled trials services or national data services, in particular

those that have access to GP data given its criticality for matching to industry protocols, should be adopted to inform end-to-end service design and workflows.

The requirements have been developed by the pharmaceutical industry and from discussions with numerous stakeholders, including therapeutic conditions that would be most suited to a data-enabled approach to recruitment and where additional scoping, such as for oncology trials, would need to be done. With participation in commercial clinical trials in the NHS at its lowest point in seven years and the HDRS currently formulating its business plans, the time to act on these recommendations is now.

Next steps

Establishing a service design group

The critical next step is to convene a UK-wide service design group (SDG) bringing together government representatives from the four nations, the HDRS team, key parts of the clinical trials delivery system and the pharmaceutical industry to define how the delivery of data-enabled commercial trials at scale could be achieved. This collaborative design process must start with the problem to be solved – rapid, targeted recruitment into commercial trials across the UK.

The SDG would have five core objectives:

- 1 **Validate the proposed design with clinical research delivery systems** to confirm that it meets real-world needs for facilitating rapid feasibility assessment, site selection and recruitment across the UK.

- 2 **Set out the operational approach** and workflows needed to connect HDRS data searches with trial coordination and delivery systems and recommend governance structures with clear responsibility for implementation.

- 3 **Recommend a phased roadmap** to guide prioritisation of set-up and implementation of the key data-enabled services, including access to comprehensive UK-wide GP data linked to national secondary care and mortality data to form the minimum service dataset.

- 4 **Expand the Clinical Practice Research Datalink (CPRD) as a minimum viable product to demonstrate proof of principle** building on its existing industry-facing trial services and linking secondary care data with CPRD's UK-wide GP data to provide quick wins and inform UK service design.

- 5 **Define success metrics** translating the key performance indicators into measurable targets.

Convening the SDG in the first half of 2026 would signal a strong commitment to translating the UK's structural advantages into coordinated strategic actions aimed at significantly improving predictability and efficiency of commercial trial delivery. Successful delivery of data-enabled clinical trials would reverse declining recruitment, secure international investment and deliver on government commitments. Importantly, creating services that is aligned with industry requirements will finally realise the potential of NHS data to increase the UK's competitive advantage and reinstate its position as a global leader in clinical trials delivery.

Annex

Methodology

The research for this report was conducted for the ABPI by Newmarket Strategy and employed a mixed-methods approach combining desk-based literature review, stakeholder interviews, and participatory workshops to examine the NHS data landscape and challenges in feasibility and recruitment for clinical trials.

The literature review encompassed desk research on the NHS data infrastructure, systematic examination of academic papers focusing on inclusion and exclusion criteria in trial protocols, and analysis of participant recruitment challenges. This was supplemented by a review of grey literature on the UK's clinical trial landscape and associated recruitment difficulties, providing context on the operational environment within which electronic health records might be deployed.

Primary data collection comprised semi-structured interviews with 20 stakeholders from UK-wide data providers including SAIL Databank, Research Data Scotland, NHS DigiTrials, and CPRD. Additional perspectives were gathered from research delivery networks and NHS trusts to capture both the national infrastructure view and local operational realities.

ABPI member engagement

Two Task and Finish Group workshops were conducted with ABPI members to validate findings and co-develop recommendations. The first workshop focused on analysing inclusion and exclusion criteria from several trial protocols to identify which trial types would be most amenable to using electronic health-record data for participant identification and exploring data requirements for a data-enabled clinical trial

recruitment service. The second examined ABPI members' existing experience with feasibility and recruitment services and defining service requirements for a new data-enabled clinical trial functionality, including how it might integrate with existing infrastructure.

Findings were presented through four sessions to ABPI data and clinical trials member groups to gather feedback on findings followed by recommendations and structured follow-up discussions.

Interviewees

The views expressed in this report are those of the ABPI. Interviews were conducted as part of the background research and the perspectives of individual interviewees may differ from the positions presented here:

- ◆ **Damian Bowler**, Head of Commercial Business Development, Innovation Research & Life Sciences Strategy, NHS England
 - ◆ **Dr Will Brown**, Consultant Neurologist, Cambridge University Hospitals
 - ◆ **Michael Chapman**, Director of Data Access & Partnerships, Data & Analytics, NHS England
 - ◆ **David Dodd**, Senior Engagement Lead, Data Access Partnerships, NHS England
 - ◆ **Prof Phil Evans**, National Associate Director of Health and Care, NIHR Research Delivery Network
 - ◆ **Sarah Fallon**, Network Operations Director, NIHR Research Delivery Network
 - ◆ **Hilary Fanning**, Senior Responsible Owner of the Data for Research and Development Programme, NHS England
 - ◆ **Martin Gibson**, Chief Medical Officer, NorthWest EHealth
 - ◆ **Dr Alison Hamilton**, Safe Haven Manager, NHS Greater Glasgow and Clyde
 - ◆ **Laura Hobbs**, Head of NHS DigiTrials & Data Operations, NHS England
 - ◆ **Prof Kamaraj Karunanithi**, Consultant Haematologist, University Hospitals of North Midlands
 - ◆ **Andy Rees**, NHS DigiTrials & Research Products Operations Manager, Innovation Research & Life Sciences Strategy, NHS England
 - ◆ **Dr David Shukla**, GP & Health and Care Research Lead for Primary Care (West Midlands), NIHR
 - ◆ **Jon Smart**, Chief Operating Officer and Head of Programmes, SAIL Databank
 - ◆ **Prof Matthew Sydes**, Head of Data-Driven Clinical Trials, Innovation Research & Life Sciences Strategy, NHS England
 - ◆ **Ming Tang**, Chief Digital and Information officer (Interim), NHS England
 - ◆ **Prof Andrew Ustianowski**, Interim RDN Executive Director, NIHR Research Delivery Network
 - ◆ **Tim Williams**, Head of Interventional Research, CPRD
- 

Table 1: Overview of how using a centralised NHS data modelling approach might suit different therapeutic areas.

Therapy area and service fit	What a centralised function will be able to handle well	What a centralised function will struggle with
<p>Cardiovascular and metabolic disease: excellent service fit</p>	<ul style="list-style-type: none"> • Basic eligibility criteria: diabetes diagnosis, hypertension, age, sex, ethnicity • Laboratory measurements: lipid panels, HbA1c values, eGFR, liver function tests • Primary care medications: statins, ACE inhibitors, metformin, with dates and dose where recorded • Smoking status and BMI measurements • Previous cardiovascular events coded in GP records 	<ul style="list-style-type: none"> • Disease control status: ‘controlled’ versus ‘uncontrolled’ hypertension or diabetes inconsistently coded despite being critical eligibility factors; requires inference from recent measurements and medication adjustments rather than explicit flags • Family history: critical eligibility factor for many cardiometabolic trials but buried in unstructured notes; systematic capture would significantly improve participant identification • Recent acute events: monthly SUS updates may miss cardiovascular events occurring within narrow timeframes; protocols specifying ‘MI within 12 weeks’ face not what patients took or continued • Medication adherence: prescribing data shows what was prescribed, not what patients took or continued

Respiratory conditions:
good service fit

- **Diagnosis codes** for asthma, COPD, with reasonable accuracy
- **Spirometry values:** FEV1, FVC, peak flow measurements
- **Inhaler prescriptions:** frequency, type, dose escalation patterns
- **Oral corticosteroid prescriptions** indicating exacerbation treatment
- **Hospital admissions** for respiratory causes via SUS data
- **Acute exacerbations within specific timeframes:** protocols specifying 'exacerbation within eight weeks' face challenges; monthly data updates may miss recent events, and exacerbation definition varies (hospital admission versus increased treatment versus symptom based)
- **Severity classification:** 'severe asthma' requires multiple criteria including treatment history and exacerbation frequency; these can be approximated but not definitively classified without clinical judgment
- **Smoking cessation dates:** smoking status captured but quit dates often imprecise, problematic for trials requiring specific time since cessation

Neurology:**moderate service fit**

- **Diagnosis codes** for major neurological conditions, although accuracy and timeliness are variable
 - **Disease duration** (calculable from diagnosis date)
 - **Certain medication histories:** disease-modifying therapies, symptomatic treatments
 - **Hospital admissions** and specialist appointments via SUS data
 - **Comorbidities and exclusionary conditions**
- **Cognitive assessment scores:** MMSE, FAQ, MoCA captured in narrative rather than structured format; these prove critical as outdated scores determine screen failures but extracting them requires manual chart review; standardised capture would transform neurology recruitment
 - **Imaging data:** MRI and CT staging information exists but resides in hospital systems with variable accessibility; some available through SUS-linked datasets like Discover NOW, others remain siloed
 - **Functional assessments:** EDSS scores for MS, UPDRS for Parkinson's typically exist only in specialist clinic letters, not extracted to structured fields
 - **Previous neuropsychiatric conditions:** psychosis history or severe depression often determines exclusion but may be incompletely coded if not currently active
 - **Treatment response patterns:** whether patients responded to previous therapies, why treatments stopped, requires interpretation beyond prescribing dates

Immunology:

moderate service fit

- **Basic diagnosis codes:** lupus, rheumatoid arthritis, Crohn's disease, ulcerative colitis
- **Some laboratory markers:** inflammatory markers (CRP, ESR), antibodies if ordered in primary care
- **Biologic prescriptions:** TNF inhibitors, IL-6 inhibitors via specialist prescribing data where captured
- **Comorbidities and common exclusions:** infections, malignancies, pregnancy
- **Disease-specific criteria:** specific antibody profiles (anti-dsDNA, anti-CCP), disease activity scores (DAS28, SLEDAI, Harvey-Bradshaw Index) exist only in specialist centre records, not nationally aggregated datasets
- **Precise staging:** mild versus moderate versus severe disease classification requires specialist assessment not consistently reflected in coded data
- **Treatment histories:** complex biologic sequences, reasons for switching, treatment failures require interpretation beyond what prescribing dates reveal
- **Joint counts, skin assessments, endoscopic findings:** clinical examination findings rarely coded systematically

Table 2: Data required to support oncology trial eligibility criteria that are not held within national GP or secondary care datasets.

Histopathology reports: tumour type, grade, receptor status (ER, PR, HER2), specific mutations – residing in individual trust pathology systems

Radiology reports: tumour size, progression, metastases documented in trust imaging systems

Genomic biomarkers: BRCA, EGFR, ALK, PD-L1 expression – National Genomic Test Directory exists but disconnected from clinical and treatment data

Treatment sequencing: SACT captures systemic anti-cancer therapy but 'lines of therapy' definitions remain notoriously difficult; protocols vary, making standardisation impractical

Staging information: Somerset Cancer Register holds diagnosis and staging separately from treatment and genomic data



References

1. The Association of the British Pharmaceutical Industry, 'The value of industry clinical trials to the UK - extended report', December 2024, available at www.abpi.org.uk/publications/the-value-of-industry-clinical-trials-to-the-uk-extended-report
2. The Association of the British Pharmaceutical Industry 'UK industry clinical trials: translating actions into impact', December 2025, available at www.abpi.org.uk/publications/uk-industry-clinical-trials-translating-actions-into-impact
3. The Association of the British Pharmaceutical Industry, 'The value of industry clinical trials to the UK', September 2024, available at www.abpi.org.uk/publications/the-value-of-industry-clinical-trials-to-the-uk
4. Department of Health and Social Care, Prime Minister's Office, Department for Science, Innovation and Technology, Office for Life Sciences, Office for Investment, 'Prime Minister turbocharges medical research', April 2025, available at www.gov.uk/government/news/prime-minister-turbocharges-medical-research
5. Department for Business and Trade, Department of Health and Social Care, Department for Science, Innovation and Technology, Office for Life Sciences, 'Life Sciences Sector Plan', July 2025, available at www.gov.uk/government/publications/life-sciences-sector-plan
6. Diwakar V, Vickers L, 'Health Data Research Service – unlocking the potential of health and care data to transform lives', NHS England, August 2025, available at www.england.nhs.uk/blog/health-data-research-service-unlocking-the-potential-of-health-and-care-data-to-transform-lives
7. Atkins V, May P, Morgan E, Matheson M, 'Full government response to the Lord O'Shaughnessy Review into commercial clinical trials', December 2023, available at www.gov.uk/government/publications/government-response-to-the-review-into-commercial-clinical-trials/full-government-response-to-the-lord-oshaughnessy-review-into-commercial-clinical-trials
8. Ross J, Tu S, Carini S, Sim I, 'Analysis of eligibility criteria complexity in clinical trials', Summit on Translational Bioinformatics, 2010, pp46–50
9. Dong H, Falis M, Whiteley W, Alex B, Matterson J et al., 'Automated clinical coding: what, why, and where we are?' NPJ Digital Medicine, 2022, 5, p159
10. Tsopra R, Peckham D, Beirne P, Rodger K, Callister M et al., 'The impact of three discharge coding methods on the accuracy of diagnostic coding and hospital reimbursement for inpatient medical care', International Journal of Medical Informatics, 2018, 115, pp35–42
11. Goldacre B, Morley J, 'Better, broader, safer: using health data for research and analysis', Department of Health and Social Care, April 2022, available at www.gov.uk/government/publications/better-broader-safer-using-health-data-for-research-and-analysis
12. Department of Health and Social Care, 'Data saves lives: reshaping health and social care with data', June 2022, available at www.gov.uk/government/publications/data-saves-lives-reshaping-health-and-social-care-with-data
13. Brooks S, 'The NHS data challenge: how to unlock insights from a sea of data', Health Service Journal, June 2024, available at www.hsj.co.uk/technology-and-innovation/the-nhs-data-challenge-how-to-unlock-insights-from-a-sea-of-data/7037336.article
14. McCarthy M, 'Reimagining the NHS must focus on restoring general practice', BMJ, 2024, 387, q2295
15. Baker M, Ware J, Morgan K, 'Time to put patients first by investing in general practice', British Journal of General Practice, 2014, 64, pp268–9
16. Lensen S, Macnair A, Love SB, Yorke-Edwards V, Noor NM et al., 'Access to routinely collected health data for clinical trials – review of successful data requests to UK registries', Trials, 2020, 21, p398
17. Sydes MR, Barbachano Y, Bowman L, Denwood T, Farmer A et al., 'Realising the full potential of data-enabled trials in the UK: a call for action', 2021, BMJ Open, 11, e043906

18. Zhang J, Morley J, Gallifant J, Oddy C, Teo JT et al., 'Mapping and evaluating national data flows: transparency, privacy, and guiding infrastructural transformation', 2023, *Lancet Digital Health*, 5, e737-48
19. Pharmaceutical Research and Manufacturers of America, 'PhRMA 2025 Annual Membership Survey', July 2025, available at www.psrma.org/resources/2025-phrma-annual-membership-survey
20. The European Federation of Pharmaceutical Industries and Associations, 'The economic impact of industry clinical trials across Europe', February 2026, available at www.efpia.eu/media/zdzg0bey/the-economic-impact-of-industry-clinical-trials-across-europe.pdf
21. The Association of the British Pharmaceutical Industry, 'Creating the conditions for investment and growth: pharmaceutical industry investment competitiveness framework', September 2025, available at www.abpi.org.uk/publications/creating-the-conditions-for-investment-and-growth
22. Callahan A, Polony V, Posada JD, Banda JM, Gombar S et al., 'ACE: the Advanced Cohort Engine for searching longitudinal patient records', *Journal of the American Medical Informatics Association*, 2021, 28, pp1468-79
23. Mahbubani K, Georgiades F, Goh EL, Chidambaram S, Sivakumaran P et al., 'Clinician-directed improvement in the accuracy of hospital clinical coding', *Future Healthcare Journal*, 2018, 5, pp47-51
24. Pankhurst T, Evison F, Atia J, Gallier S, Coleman J et al., 'Introduction of systematized nomenclature of medicine-clinical terms coding into an electronic health record and evaluation of its impact: qualitative and quantitative study', *JMIR Medical Informatics*, 2021, 9, e29532



About the Association of the British Pharmaceutical Industry

The ABPI exists to make the UK the best place in the world to research, develop and access medicines and vaccines to improve patient care.

We represent companies of all sizes which invest in making and discovering medicines and vaccines to enhance and save the lives of millions of people around the world.

In England, Scotland, Wales and Northern Ireland, we work in partnership with governments and the NHS so that patients can get new treatments faster and the NHS can plan how much it spends on medicines. Every day, our members partner with healthcare professionals, academics and patient organisations to find new solutions to unmet health needs.

The Association of the British Pharmaceutical Industry

A company limited by guarantee registered in England & Wales number 09826787

Registered office 2nd Floor Goldings House, Hay's Galleria, 2 Hay's Lane, London, SE1 2HB

XXX-XXXX-XXXX

