Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Examining UK clinical trialists' experience, knowledge and confidence accessing and using Health Systems Data to support development of resources and support

Jemima Thompson
University College London
Sharon B Love
University College London
Marina Bobou
University College London
Fiona Lugg-Widger
Cardiff University
Robert Trubey
Cardiff University
Michael Robling
Cardiff University
Isla S Mackenzie
University of Dundee
Macey L Murray
macey.murray@ucl.ac.uk

University College London

Research Article

Keywords: Health Systems Data, Knowledge transfer exchange, clinical trials, professional education

Posted Date: November 13th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11059088/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: Competing interest reported. JT, SBL, MB, FLW, RT and MLM declare that they have no competing interests. ISM declares Institutional research grants from HDR UK, British Heart Foundation, AstraZeneca, Wellcome Trust, Tayside Institute for Cardiovascular Research (TICR), RTI Health Solutions, University of Bordeaux and NovoNordisk; honoraria from Amgen, Amarin, Novartis and NovoNordisk; support from Novartis, Menarini and AstraZeneca to attend meetings; participant of data monitoring committee for WAVE, AFFECT and CARE-UK; vice-chair of Scottish Heart and Arterial disease Risk Prevention (SHARP), and member of British and Irish Hypertension Society (BIHS) collaborative research working group.

Examining UK clinical trialists' experience, knowledge and confidence accessing and using Health Systems Data to support development of resources and support

Jemima Thompson[1,2], Sharon B Love[1,2], Marina Bobou[1,2], Fiona Lugg-Widger[2,3], Robert Trubey[2,3], Michael Robling[2,3], Isla S Mackenzie[2,4,] Macey L Murray[1,2,5]

[1]UCL Innovative Clinical Trials Unit, UCL ICTM, 2nd Floor, 90 High Holborn, London, England, WC1V 6LJ

[2]Transforming Data for Clinical Trials programme, Health Data Research UK

[3]Centre for Trials Research, Cardiff University, 7th Floor, Neuadd Meirionnydd, Health Park, Cardiff, Wales, CF14 4YS

[4]Medicines Monitoring Unit (MEMO Research), University of Dundee, Ninewells Hospital and Medical School, Dundee, Scotland, DD2 1GZ

[5]MRC Centre of Research Excellence in Clinical Trial Innovation, UCL ICTM, 2nd Floor, 90 High Holborn, London, England, WC1V 6LJ

Corresponding Author(s):

Dr. Macey L Murray: macey.murray@ucl.ac.uk

Dr Jemima Thompson: jemima.thompson.18@ucl.ac.uk

Abstract

Background

Health Systems Data (HSD) are routinely collected real-world health data that can supplement or replace traditional routes of clinical trial participant recruitment and data collection.

Gaps in trialists knowledge and experience are barriers to effective HSD implementation in trials. Provision of resources to strengthen trialists' knowledge and confidence may drive wider adoption of HSD-enabled trials across the UK.

This study sought to characterise current levels of HSD experience, knowledge and confidence across the UK clinical trials community, to guide the development of targeted resources and support.

Method

An online survey was conducted, asking UK-based clinical trials staff to outline their experience, knowledge and confidence accessing and using HSD. Dreyfus' five stage model of skill acquisition framed the analysis.

Results

There were 110 respondents from across the UK. Reported years' HSD experience ranged from none to over two decades. Knowledge varied but, on average, respondents were situated in the 'advanced beginner' stage of skill acquisition, though small pockets of expertise were identified. Confidence also varied; around 10% of respondents felt very confident. Levels of knowledge, experience and confidence were not always aligned, e.g. being very experienced did not necessarily correlate with feeling very confident. Mentorship and support group initiatives were positively received as ideas.

Conclusions

Resources should be tailored to meet varying learning needs and skill levels. Additionally, community experts should be identified to lead the development of guidance and learning materials, alongside

collaborative efforts to establish mentoring and peer-support networks, though concerns about capacity would need to be addressed.

Key words: Health Systems Data, Knowledge transfer exchange, clinical trials, professional education.

1. Background

Health Systems Data (HSD) are a form of real-world data, routinely collected from patients' interactions with the health system. Information included in HSD often overlap with that collected about participants in Randomised Controlled Trials (RCTs) (1), such as cause of death, hospital admissions and diagnoses. This overlap means that HSD may be used to supplement or replace traditional routes of participant recruitment and bespoke trial data collection. This could improve recruitment and retention rates and trial operating costs (2,3). For example, a substudy in the Cluster Randomised Trial of PSA testing for Prostate Cancer (CAP) found that trial operational costs were substantially reduced when HSD was utilised over local hospital data (4). The RECOVERY trial reduced baseline operational costs to less than £100 per patient by using HSD, a 99% reduction compared to traditional data collection methods (5). These potential benefits mean that more UK-based RCTs seek to use HSD, aka become "data-enabled".

In recent years, the number of trial protocols stating intended HSD use rose from 47% prior to 2019 to 62% between 2019-2022 (6). Despite increased interest in data-enabled trials, there are some common barriers that affect the extent of HSD use. Complex, fragmented data access and data quality concerns are challenges widely discussed among the UK trials community (2,3,7–11). For example, the 2023 O'Shaughnessy report states that there are clear, technical barriers to HSD use in UK trials due to issues with interoperability and lack of common standards (12). This was echoed in the Sudlow Review of the UK health data landscape (9).

Alongside these technical, system-level challenges, limited knowledge and experience among trial staff remain important barriers to effective HSD implementation in clinical trials(2,13). These expertise gaps are partly attributable to the relatively recent application of HSD within trials and the absence of established, standardised guidance for its use (14). Evidence suggests that experience with HSD remains limited across the UK trials community (15). In a systematic review of successful data access requests (2013-2018), only 63% (N = 32) of UK registered CTUs had accessed data for at least one trial, while just 3% of HRA-approved trials in 2015 successfully obtained HSD (15). However,

there remains little empirical evidence describing the current levels of knowledge, confidence and experience in HSD use across the wider UK trials workforce.

The COMORANT-UK Delphi study further highlighted the need to address methodological and implementation evidence gaps, alongside investment in infrastructure, to support best practice in data-enabled trials (7). Of 47 challenges initially identified, seven were prioritised, many relating to training, service reorganisation, and methodological uncertainty. As uptake of data-enabled trials remains relatively low, many trial staff have limited practical experience of obtaining and using these data sources, which can lead to delays as teams navigate complex access processes and governance requirements (3,11). For example, the PRIMORANT study reported that, although trial staff possessed considerable expertise in their respective disciplines, many lacked the specialised skills required to access and manage HSD (8). To our knowledge, no previous study has quantified knowledge, confidence and experience relating to HSD use across the UK clinical trials community at a national level.

Clarity about trialists' knowledge, experience and learning needs provides a foundation for developing and sharing appropriately targeted resources and support. However, the level and type of support required is likely to vary according to an individual's role and intended use of HSD. Training for HSD-enabled trials should therefore be tailored to suit differing levels of expertise, evolving from foundational instruction to more complex, scenario-based learning. Mentoring and peer support may further promote skill development by enabling observation, guided participation, and constructive feedback (16,17). To our knowledge, this model has not previously been applied to examine knowledge, confidence and experience of HSD use within the UK trials community.

Health Data Research UK (HDR UK) is the UK's institute for health data science, whose work seeks to *"address the most pressing health research, and to provide training and education"(18)*. The HDR UK Transforming Data for Trials (TDfT) programme endeavours to improve how HSD is accessed, analysed and used in clinical trials by building a standardised framework of guidance, models and protocols, creating digital resources and providing targeted community training (19). As part of this programme,

knowledge gaps will be addressed through a suite of learning packages, hosted on the HDR UK Futures learning platform (20) and training courses, such as Data Utility Comparison Study (DUCkS) workshops (21). The programme also includes the development of a fourteen-station Routemap that complements the NIHR Clinical Trials Toolkit, and a case study catalogue of learnings; both aim to provide practical guidance for identifying, accessing and using HSD in trials (19).

Capturing details about the current state of knowledge and experience across the trials community can enable the design and delivery of the HDR UK training programme, workshops, learning materials and community support, by ensuring that they are targeted to key areas, pitched at appropriate levels and use suitable formats. This study was undertaken as the first UK-wide assessment of clinical trialists' knowledge, confidence and experience using HSD. It will generate evidence to directly inform national capability-building activities within HDR UK and the UK clinical trials ecosystem. Strengthening knowledge, gaining experience and building confidence may facilitate wider adoption of HSD-enabled clinical trials in a way that is tailored to the needs of the community.

2. Aim

This study sought to characterise current levels of experience, knowledge and confidence in accessing and using HSD across the UK clinical trials community, to guide the future development of targeted resources and support.

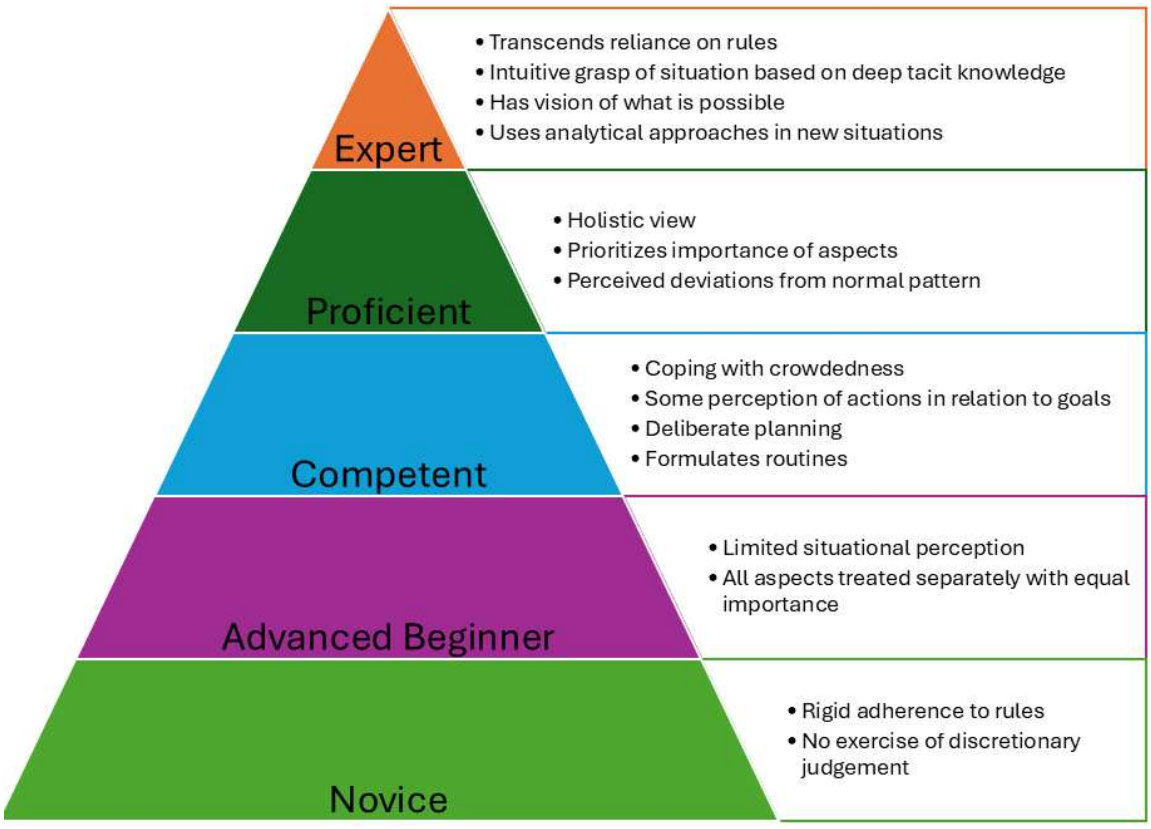
3. Method

3.1. Design

This exploratory study consisted of a survey of UK-based Clinical Trials Unit (CTU) staff and other members of the trials community for their perspectives and experience of HSD use. The online survey comprised a combination of open and closed questions, hosted on Microsoft Forms (22) (See Additional files 1).

Eight domains of knowledge and relevant aspects of HSD use within each domain were derived from the TDfT Routemap (19). Dreyfus' model of skill acquisition (23) was used to frame the analysis of the survey data. According to Dreyfus and Dreyfus (23), skill acquisition develops through five stages: novice, advanced beginner, competent, proficient, and expert. This framework provides a structured approach for understanding workforce capability. Progression through these stages is accompanied by changing learning needs. Closed questions relating to knowledge levels consisted of a five-point response scale, which mapped onto Dreyfus' model. Experience using HSD was captured using closed and open questions. Confidence was rated on a four-point scale, from not at all confident to very confident. Respondents also ranked their preferred learning formats from most to least preferred.

Figure 1: Dreyfus' five stages of skill acquisition (24)



3.2. Recruitment

Due to the exploratory nature of this study, a statistically representative sample size was not formally calculated. Rather, we focused on obtaining demographic diversity, in particular geographic, role and experience.

The survey took place between 23-Sep-2025 and 31-Oct-2025. These date parameters allowed enough potential participants time to respond, whilst also taking into consideration October school holidays, which can disrupt work routines. Recruitment ended when the research team felt that the sample was sufficiently diverse.

To be eligible, respondents needed to self-identify as currently working in clinical trials in the UK, in any role and with any level of HSD experience. Respondents were recruited via several channels, including direct emails with named individuals who had previously consented to be contacted by TDfT (via their Community Insight Group), emails or newsletters/updates to relevant trial groups, such as the UK Trial Managers' Network (UKTMN), UK registered CTU network, Trials Methodology Research Partnership working groups and the Royal Statistical Society. Links to the survey were also included in some of the study team email signatures.

Respondents self-selected to participate. Prior to completing the survey, respondents accessed a Participant Information Sheet and were required to provide written informed consent (Additional file 1). Respondents were offered a small incentive to complete the survey, which was the option to be entered into a prize draw upon completion on the survey with a chance of receiving either a £50 gift voucher, HDR UK branded merchandise, such as mugs and pens, or some sweet treats. This study was approved by the UCL Research Ethics Committee [project ID: 1697].

3.3. Analysis

Quantitative data were analysed in Stata v19.0 (25). Descriptive statistics are reported including frequencies, percentages and means. Levels of knowledge were mapped onto Dreyfus' 5-stages of skill acquisition (23). Qualitative data from open questions was analysed using a qualitative descriptive content analysis approach (26). Data were single coded by JT, by hand and findings discussed with MLM.

4. Results

A total of 121 people consented. However, 11 did not meet eligibility requirements, therefore 110 respondents completed the survey.

4.1. Demographics

Respondents worked in various roles across clinical trials, though the majority held Trial Management or Statistician roles. All four UK nations were represented and were broadly proportional to population distribution across the four nations. Respondents were from 50 institutions across the UK (Table 1).

Table 1: Respondent demographics

Job Category (N=110)	**N (%)**
Trial Management	36 (32.7%)
Statistician	27 (24.6%)
Clinical	20 (18.2%)
Data Management	16 (14.6%)
Other	11 (10.0%)
Country (respondents) (N=109)	**N (%)**
England	89 (81.7%)
Scotland	13 (11.9%)
Wales	6 (5.5%)
Northern Ireland	1 (0.9%)
Institutions across English regions (N=50)	
London	14 (28%)
South-East England	6 (12%)
Yorkshire and Humber	5 (10%)
West Midlands	4 (8%)
East Midlands	3 (6%)
South-West England	3 (6%)
North-East England	2 (4%)
East of England	2 (4%)
North-West England	1 (2%)
Institutions across Scotland, Wales, Northern Ireland (N=10)	
Scotland	8 (16%)

Wales	1 (2%)
Northern Ireland	1 (2%)

4.2. Experience

Around two-thirds of respondents had HSD experience and almost three-quarters of those were currently working on a trial that was planning to or was actively accessing/using HSD. A small number of respondents (7 [11%]) had over 20 years' HSD experience (table 2).

Table 2: **Respondents' years of experience using HSD**

Ever accessed/used HSD (N=109)	**N (%)**
Yes	65 (60%)
No	44 (41%)
Currently working on a trial planning to or currently accessing/using HSD (N=65)	**N (%)**
Yes	46 (72%)
No	19 (29%)
Time involved with HSD (N=65)	**N (%)**
<1 year	2 (3%)
1-9 years	39 (61%)
10-19 years	16 (25%)
>20 years	7 (11%)

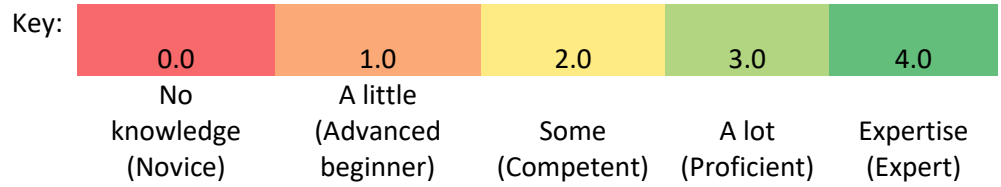
4.3. Knowledge

Figure 2 shows that, when interpreted using Dreyfus' model, respondents' average knowledge across HSD domains was generally situated at the advanced beginner level. None of the groups averaged at proficient or expert levels. Clinical staff reported the highest overall knowledge, particularly in trial design and management, approvals, and data management and analysis, although their average knowledge remained within the advanced beginner stage.

Knowledge reached the competent stage in only a few areas. This was most evident for informed consent (transparency and the right to withdraw) among trial management, data management, and clinical staff, and for de-identification and anonymisation among data management staff.

Novice knowledge was observed across several domains, particularly site feasibility and selection, protocol development, and data management. Knowledge of the RECORD guidelines was especially low, while understanding of synthetic data remained at the lower end of the novice stage across all job categories. Statisticians were most frequently classified at the novice level, particularly for approvals and contracts, Section 251 approvals, and managing participants' mental capacity within informed consent (details in Additional file 2).

Figure 2: Mean knowledge of HSD across domains by job category



HSD Domains	Job category					
	Trial Management	**Data Management**	**Statistician**	**Clinical**	**Other**	**Total**
Trial Design and Management	1.4	1.3	1.1	1.7	1.6	**1.4**
Approvals	1	0.9	0.6	1.4	1.3	**1**
Informed Consent	1.7	1.6	1.1	1.8	1.4	**1.5**
Site feasibility and selection	1	0.7	0.6	1.4	0.6	**0.9**
Contracts and agreements	1.5	1.3	0.9	1.4	1.2	**1.3**
Protocol development	1.1	0.8	1.1	1.7	0.8	**1.1**
Data management and analysis	0.5	1	0.9	1.3	0.6	**0.8**
Archiving and data sharing	1.3	1.6	1.1	1.4	0.9	**1.3**
TOTAL	**1.2**	**1.2**	**1**	**1.5**	**1.1**	**1.2**

A small number of respondents reported expertise in 40/46 (87%) aspects across all eight HSD domains. Notably, despite low overall knowledge of protocol development, particularly the RECORD guidelines, expertise was reported by a small number of respondents across multiple job categories (details in Additional file 3).

Expertise was absent in 6/46 (13%) aspects across the eight domains. The largest gap was in Contracts and agreements, where no respondents reported expertise in 5/8 aspects. The only areas in which expertise was reported were the importance of contracting, Data Protect Impact Assessments (DPIAs), and data sharing agreements, primarily among data management staff and statisticians. Within Data management and analysis, no respondents reported expertise in the use of synthetic data. Expertise in the remaining aspects was largely confined to data management staff and statisticians (table 3).

Expertise was reported across all job categories but was least common among trial management and other staff. Clinical staff reported expertise across all aspects of Trial design and management, Approvals, Informed consent, and Site feasibility and selection. Interestingly, while expertise was reported in all aspects of Trial design and management, none of those reporting expertise held trial management roles. Expertise in Archiving and data sharing was reported predominantly by data management staff.

[insert Table 3: HSD Expertise by job category]

4.4. Confidence

Confidence using HSD was mixed, with around a third of respondents reporting feeling not at all confident (Fig. 3).

Figure 3: Confidence accessing/using HSD by job category



For the 63 (58%) respondents who reported not feeling at all confident or slightly confident, five key reasons for their lack of confidence accessing/using HSD were identified (table 4).

Table 4: Reasons for lack of confidence using HSD

Category	Example quotes
Timeliness of data access/receipt	"[I'm] not confident we can deliver to proposed timelines as these seem change so much." ***(Respondent 122, clinical, slightly confident)***
Concerns/queries regarding data quality	"In my experience NHS data is often incomplete or unreliable for use in clinical trials." ***(Respondent 39, data management, slightly confident)***
The complexity involved in gaining access to/using the data	"Using HSD is incredibly complicated due to overlapping legislation and every organisation you can think of having their own rules." ***(Respondent 43, data management, slightly confident)***
Lack of experience	"I don't really know about HSD" ***(Respondent 124, clinical, not at all confident)***
Previous negative experiences	"Our previous use of routine data in a large, impactful cancer trial created a lot of issues, some of which are still ongoing. I would take a [word missing] of convincing to use routine data again" ***(Respondent 25, trial management, not at all confident)***

4.5. Knowledge – Confidence – Experience

Table 5 shows that those with HSD experience were generally more confident than those with none. Regardless of the length of HSD involvement, confidence levels varied.

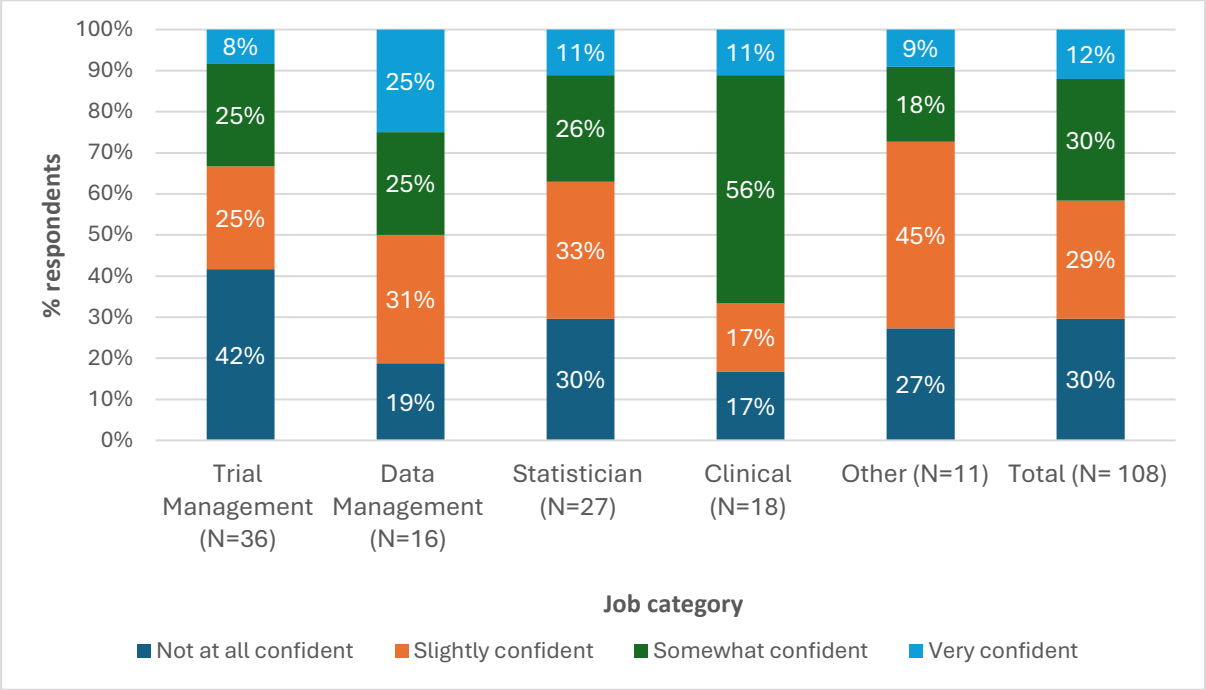
Of 12 respondents that had used HSD and were very confident, 7 (58.5%) reported having expertise in specific areas of HSD use. Six of 32 (15.6%) somewhat confident respondents had expertise in some areas, as had 7/32 (22.6%) slightly confident respondents.

Table 5: Confidence classified by experience of accessing/using HSD

Experience	**Confidence N (%)**				
	Not at all confident	**Slightly confident**	**Somewhat confident**	**Very confident**	***Total***
Ever used HSD (N=109)					
Yes	2 (3.1%)	23 (35.4%)	28 (43.1%)	12 (18.5%)	*65 (59.6%)*
No	30 (68.2%)	9 (20.5%)	4 (9.1%)	1 (2.3%)	*44 (40.4%)*
Total	**32 (29.4%)**	**32 (29.4%)**	**32 (29.4%)**	**13 (11.9%)**	***109 (100%)***
Time involved with HSD (N=59)					
<1 year	0 (0.0%)	1 (50.0%)	0 (0.0%)	1 (50.0%)	*2 (3.4%)*
1-9 years	10 (25.6%)	10 (25.6%)	15 (38.5%)	4 (10.3%)	*39 (66.1%)*
10-19 years	4 (28.6%)	3 (21.4%)	4 (28.6%)	3 (21.4%)	*14 (23.7%)*
>20 years	1 (25.0%)	1 (25.0%)	2 (50.0%)	0 (0.0%)	*4 (6.8%)*
Total	**15 (25.4%)**	**15 (25.4%)**	**21 (35.6%)**	**8 (13.6%)**	***59 (100%)***

Confidence and knowledge scores were re-scaled to allow for direct comparison between the two variables. Figure 4 shows that those in data management roles had the largest gap between their overall knowledge and confidence, whilst those in trial management were more closely aligned in terms of their perceived confidence and overall knowledge.

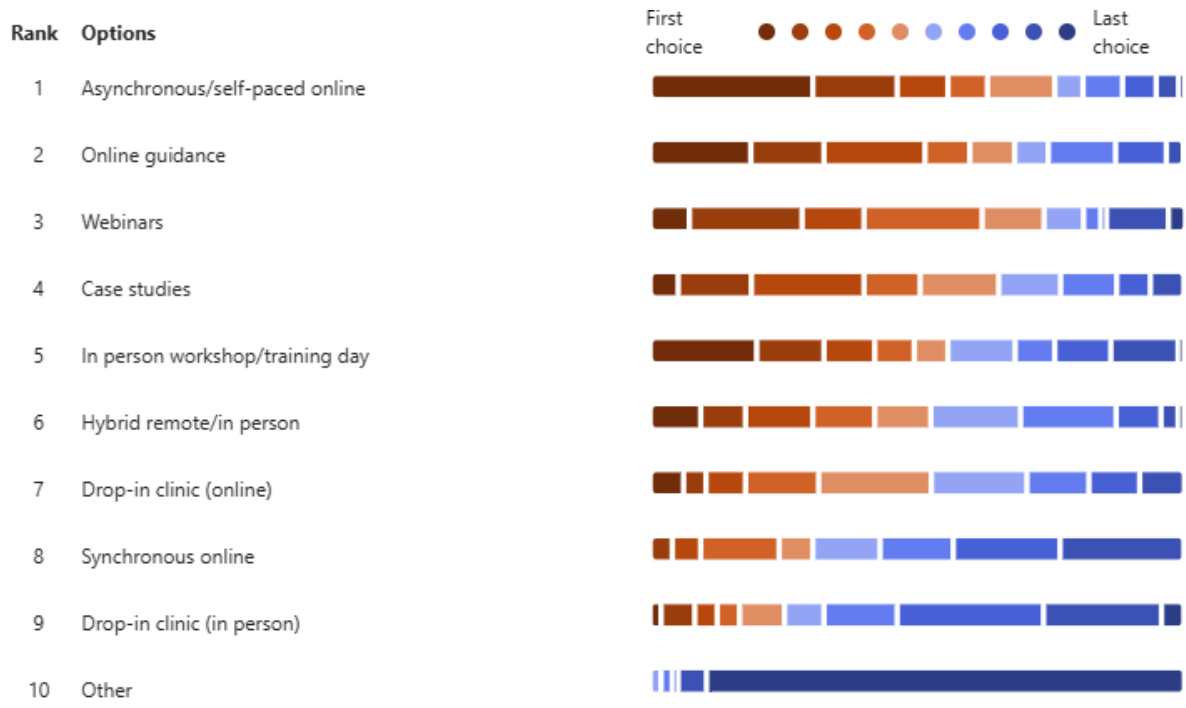
Figure 4: Mean knowledge by confidence accessing/using HSD by job category



4.6. Preferred learning formats

Ninety-five (86.4%) respondents ranked their preferred learning formats. Asynchronous/self-paced online was most preferred and in-person drop-in clinics were least preferred (fig, 5).

Figure 5: Rankings of preferred learning formats



4.7. Mentorship and support

Ninety-nine (91.8%) respondents indicated that they would consider accessing a mentor and 90 (82.6%) would consider joining a support group. Some respondents felt that access to mentorship or support would help enable learning from others.

"Being able to discuss with someone more experienced with HSD would be useful." (Respondent 29, statistician)

"Happy to join a group to share knowledge and experiences." (Respondent 33, data management)

However, despite this openness, there was some hesitancy around capacity, ease of access and uncertainty around how much this level of support would be needed.

"Dependent on the availability and ease of contact" (Respondent 58, trial management)

"I am unsure how much I would use a mentor" (Respondent 59, statistician)

Some respondents were unsure whether a support group would be the "*right format"* (Respondent 105, statistician) for receiving help and support.

The 9.2% of respondents that would not access a mentor felt that they had sufficient expertise, were able to access support already, or that HSD use was not relevant to their role.

"I feel I have sufficient experience to not require mentorship" (Respondent 18, clinical)

"I could access relevant support from within my unit if needed" (Respondent 69, statistician)

"This is probably more relevant for people involved in the use of HSD (which is not my role)" (Respondent 83, trial management)

5. Discussion

Findings from this study contribute to the literature by capturing information about the nature and distribution of HSD experience, knowledge and confidence among the UK clinical trials community. To our knowledge, this is the first UK-wide study to quantify these attributes across the clinical trials workforce and the first to interpret findings through a recognised skills acquisition framework. As such, the findings provide an evidence base to support national capability-building activities relating to data-enabled trials.

A key finding from this study was that experience and knowledge are unevenly distributed across the community. Although respondents reported a wide range of experience, average knowledge levels were aligned with the advanced beginner stage of skill acquisition (23). A small number of respondents reported having expertise across all eight domains, suggesting that expertise exists but is concentrated within a limited group. This finding is consistent with McKay et al (27) and Sydes (28), who identified that both knowledge gaps and pockets of expertise can be found across the clinical trials community.

This concentration of knowledge within small sections of the community may reflect limited opportunities available to participate in HSD-enabled trials. According to Dreyfus and Dreyfus (1980) (23), expertise develops through repeated engagement with tasks and practical applications of knowledge. Similarly, theories of situated learning (acquirement of knowledge through social interaction and active participation in a real-world context) and communities of practice (groups of people who share a passion, craft or problem who interact regularly to learn) emphasise that expertise is acquired and shared through participation alongside more experienced practitioners (16,29). Where opportunities to run HSD-enabled trials has been limited, expertise is likely to have accumulated within a relatively small number of individuals, research groups and CTUs (3,28). Consequently, efforts to build capability across the trials community should focus on a combination of education, training and increasing opportunities for practical involvement, mentorship and knowledge exchange.

However, improving individual capability alone is unlikely to be sufficient. Many of the barriers identified in previous research, including fragmented access systems, governance complexity, interoperability challenges and concerns regarding data quality, operate at organisational and system levels and may continue to limit HSD adoption even where knowledge and confidence are improved.

The average positioning of respondents in the advanced beginner stage indicates an ability to recognise recurring aspects of practice but a continued need for support in applying knowledge in more complex situations. Dreyfus' model suggests that progression beyond this stage relies on practical experience and exposure to authentic tasks (30). However, the presence of respondents across all stages of skill acquisition suggests that resources are also required that support skill acquisition in novices, such as glossaries of terms and guidance documents expanding basic concepts and ideas, to proficient/experts, where leadership opportunities and contribution to the development of guidance may be more appropriate (23,31). As a result, capability development is likely to require a combination of written guidance, incorporating mentorship, case-studies, communities of practices and opportunities to participate in HSD-enabled trials.

Respondents were generally receptive to mentorship schemes and peer support networks, though concerns around capacity were a potential barrier. This barrier is particularly relevant in academic trials, where time and resource constraints are well recognised (32). Therefore, to be sustainable, mentorship and support initiatives would need to provide clear value to those seeking support and those providing it.

Notable differences in knowledge and confidence were observed across professional groups. Those in trial management roles were least confident, despite reporting being at least as knowledgeable as those in other roles. This may be related to challenges surrounding professional identity, workload burden and feelings of being under-valued (33,34), rather than actual or perceived knowledge gaps. In an effort to address these concerns, the UK Trial Managers' Network (UKTMN) competency framework was developed to encourage career development and training opportunities across Trial managers, taking into account all levels of experience (35). These differences may also reflect

variation in roles and responsibilities within clinical trials. For example, data managers and statisticians routinely engage with data collection, quality assurance, management, and analysis processes, whereas trial managers are more likely to be involved in study delivery, governance, and data access procedures. As a result, exposure to HSD-related activities is likely to vary, contributing to differences in knowledge and confidence, suggesting a need for role specific-approaches to training and development (36,37).

Confidence was also variable and not always aligned with experience or knowledge. This suggests that confidence may reflect factors beyond technical knowledge alone. For example, confidence may be influenced by previous experiences of obtaining and using HSD. Around a third of respondents with more than 10 years' experience reported low confidence, while approximately one-third of those reporting high levels of confidence did not identify themselves as experts in any aspect of HSD. Given the well-documented challenges associated with HSD, including governance requirements linkage limitations, timeliness, and missing data (3,38,39), lower confidence among more experienced users may reflect a greater awareness of complexity rather than lack of expertise. Equally, higher confidence among some less experienced respondents may reflect successful engagement with a limited number of HSD activities, such as participant identification or feasibility assessment, rather than broad expertise across all stages of HSD use. Creating opportunities for users to share their knowledge and lessons learned through publications, mentorship and communities of practice may therefore help build capability and confidence among the wider trials community (29).

Strengths and limitations

A strength of this study was its wide reach across the UK clinical trials community. The survey received a high number of responses from trialists across the UK, job roles and experience, so these data likely represent the diversity of UK trialists. Another strength was the application of Dreyfus' model, which provided a structured developmental framework that allowed the capture of nuanced perceptions of knowledge, confidence and experience, in a way that had not been explored previously, to support the development of targeted resources in appropriate formats.

A limitation of this study is the possibility of selection bias. Respondents self-selected and those who responded may have stronger feelings about HSD use than those who chose not to participate. It is also of note that members of the TDfT team did not participate in the study. This is a group with considerable HSD knowledge and experience that were not taken into consideration when exploring the distribution of expertise across the UK clinical trials community. Additionally, while the Dreyfus model effectively maps foundational skill growth, its application is restricted by the assumption of a static domain with stable rules. Infrastructure around HSD is ever changing. Consequently, this framework may not fully account for how trialists may need to adapt where established intuition may become obsolete due to regular changes and updates. Finally, whilst the aggregation of knowledge to the mean score across the domains is a pragmatic approach to analysing this data, it is acknowledged that domains are likely to be unequal in scope, complexity and relevance to certain trial staff which may require additional exploration and consideration during resource development.

Recommendations

HDR UK's Transforming Data for Trials programme was established partly to equip the clinical trials community with the skills needed to use HSD more effectively (19). The programme's emphasis on developing resources, training, case studies and networks suggests an acknowledgement that expertise needs to be deliberately developed and disseminated across the broader trials community. While training and resource development are important, efforts to increase HSD adoption should sit alongside broader initiatives to reduce structural barriers to access and use. Capability-building interventions are therefore likely to be most effective when implemented in parallel with improvements in infrastructure, governance processes and access pathways. To support the effective development and dissemination of resources and support, based on findings from this study, we make the following recommendations:

1. Resources currently under development should be tailored to meet the different learning needs of the trials community, based on their stage of skill acquisition and learning preferences.

2. Identify the experts from within the trials community and find ways to support their involvement in leadership and development of materials and guidelines.
3. Collaborate with the trials community to support the development of a mentoring scheme and/or support group to enable community and peer support for accessing/using HSD.

Additional work is needed to further examine the views and experiences of trialists to gain a more in-depth understanding of how to tailor resource development, not just to knowledge but also to anticipated use of HSD, professional role, organisational context and perceived need for HSD-related skills. Lower knowledge levels do not necessarily indicate a training need, as required expertise will depend on how individuals use, or intend to use, HSD within clinical trials.

6. Conclusion

Data-enabled clinical trials have the potential to improve efficiency of trial conduct for trialists and reduce data collection burden for participants. Based on the study findings, resources should be tailored to the varying learning needs and skill levels of the trials community. However, increasing adoption of HSD-enabled trials will require more than training alone. Alongside capability-building initiatives, continued efforts are needed to address organisational and system-level barriers relating to data access, governance, interoperability and data quality. Experts within the community should be supported to lead the development of guidance and materials, alongside collaborative efforts to establish mentoring and peer-support networks that facilitate the uptake and use of HSD. Further work is needed to explore the impact of other factors that may affect knowledge and to measure the impact of learning resources and support on the uptake and implementation of HSD in clinical trials.

List of abbreviations

CAP	Cluster Randomised Trial of PSA testing for Prostate Cancer
CORMORANT	Consensus on Methodological Opportunities for Routine data Trials
CTU	Clinical Trials Unit
DUCkS	Data Utility Comparison Study
HDR	Health Data Research
HSD	Health Systems Data
NIHR	National Institute for Health and care Research
RCT	Randomised Controlled Trial
RECORD	REporting of studies Conducted Using Observational Routinely-collected Health Data
RECOVERY	Randomised Evaluation of COVID-19 ThERapY
TDfT	Transforming Data for Trials
TMRP	Trial Research Methodology Partnership
UKTMN	UK Trial Managers Network

Declarations

Ethics approval and consent to participate

This study was conducted in accordance with the Declaration of Helsinki and was approved by the UCL REC [project ID: 1697]. Participants provided their written informed consent to take part.

Consent for publication

Not applicable

Availability of data and materials

The datasets used and/or analysed during the current study are available from the corresponding author upon reasonable request.

Competing interests

JT, SBL, MB, FLW, RT and MLM declare that they have no competing interests. ISM declares Institutional research grants from HDR UK, British Heart Foundation, AstraZeneca, Wellcome Trust, Tayside Institute for Cardiovascular Research (TICR), RTI Health Solutions, University of Bordeaux and NovoNordisk; honoraria from Amgen, Amarin, Novartis and NovoNordisk; support from Novartis,

Menarini and AstraZeneca to attend meetings; participant of data monitoring committee for WAVE, AFFECT and CARE-UK; vice-chair of Scottish Heart and Arterial disease Risk Prevention (SHARP), and member of British and Irish Hypertension Society (BIHS) collaborative research working group.

Funding

This project was funded by the HDR UK Transforming Data for Trials programme.

The Centre for Trials Research receives infrastructure funding from Health and Care Research Wales

Funders did not have any role in the conceptualisation, design, data collection, analysis, decision to publish or preparation of the manuscript.

Authors' contributions

JT was involved in the conceptualisation, design, delivery, analysis, and was the lead author of the manuscript. SBL supported the design and analysis of the work and contributed to the writing of the manuscript. MB supported the design of the survey and commented on a draft manuscript. FLW supported the design of the survey and contributed to the draft manuscript. RT supported the design of the survey and commented on a draft manuscript. MR supported the design of the survey and commented on a draft manuscript. ISM supported the design of the survey and contributed to the draft manuscript. MLM was involved in the conceptualisations of the survey, the study design, analysis and writing of the manuscript. All authors read and approved the final manuscript.

Acknowledgements

We wish to thank survey respondents for their time and contributions to this work. We would also like to acknowledge members of the HDR UK Community Insight Group (CIG) who circulated and participated in the survey, the Institute of Clinical Trials and Methodology at UCL, UKTMN, TMRP and UK registered CTU Network for their support in circulating the survey to their members.

References

1. Sydes MR, Barbachano Y, Bowman L, Denwood T, Farmer A, Garfield-Birkbeck S, et al. Realising the full potential of data-enabled trials in the UK: A call for action. BMJ Open. 2021 Jun 16;11(6). doi:10.1136/bmjopen-2020-043906 PubMed PMID: 34135032.

2. Mc Cord KA, Al-Shahi Salman R, Treweek S, Gardner H, Strech D, Whiteley W, et al. Routinely collected data for randomized trials: Promises, barriers, and implications. Trials. BioMed Central Ltd.; 2018. doi:10.1186/s13063-017-2394-5 PubMed PMID: 29325575.

3. Macnair A, Love SB, Murray ML, Gilbert DC, Parmar MKB, Denwood T, et al. Accessing routinely collected health data to improve clinical trials: recent experience of access. Trials. 2021 Dec 1;22(1). doi:10.1186/s13063-021-05295-5 PubMed PMID: 33971933.

4. Thorn JC, Turner EL, Hounsome L, Walsh E, Down L, Verne J, et al. Validating the use of hospital episode statistics data and comparison of costing methodologies for economic evaluation: An end-of-life case study from the cluster randomised triAl of PSA testing for prostate cancer (CAP). BMJ Open. 2016;6(4). doi:10.1136/bmjopen-2016-011063 PubMed PMID: 27130167.

5. Anderson BR, Gotlieb EG, Hill K, McHugh KE, Scheurer MA, Mery CM, et al. Registry-based trials: A potential model for cost savings? Cardiol Young. 2020 Jun 1;30(6):807–17. doi:10.1017/S1047951120001018 PubMed PMID: 32605679.

6. Toader AM, Gamble CL, Dodd S, Williamson PR. The use of healthcare systems data for RCTs. Trials. 2024 Dec 1;25(1). doi:10.1186/s13063-023-07846-4 PubMed PMID: 38287383.

7. Williams ADN, Davies G, Farrin AJ, Mafham M, Robling M, Sydes MR, et al. A DELPHI study priority setting the remaining challenges for the use of routinely collected data in trials: COMORANT-UK. Trials. 2023 Dec 1;24(1). doi:10.1186/s13063-023-07251-x PubMed PMID: 36997954.

8. Toader AM, Gamble CL, Dodd S, Xie CX, Gillies K. Researchers' perspectives of accessing healthcare systems data for use in randomised controlled trials. J Clin Epidemiol. 2026 Jul 1;195. doi:10.1016/j.jclinepi.2026.112268 PubMed PMID: 41950974.

9. Sudlow CLM. Uniting the UK's Health Data: A Huge Opportunity for Society: Executive Summary [Internet]. 2024 Nov. Available from: https://www.gov.uk/government/publications/independent-investigation-of-the-nhs-in-england

10. Lugg-Widger F V., Angel L, Cannings-John R, Hood K, Hughes K, Moody G, et al. Challenges in accessing routinely collected data from multiple providers in the UK for primary studies: Managing the morass. In: International Journal of Population Data Science. Swansea University; 2018. doi:10.23889/ijpds.v3i3.432 PubMed PMID: 34095522.

11. Murray ML, Mafham M, Hartley S, Carpenter J, Landray M, Love SB, et al. Use of NHS Digital datasets as trial data in the UK: a position paper Heather Pinches 3. 2022.

12. O'Shaughnessy J. Commercial Clinical Trials in the UK: The Lord O'Shaughnessy review [Final report]. 2023 May.

13. Sydes MR, Murray ML, Ahmed S, Apostolidou S, Bliss JM, Bloomfield C, et al. Getting our ducks in a row: The need for data utility comparisons of healthcare systems data for clinical trials. Contemp Clin Trials. 2024 Jun 1;141. doi:10.1016/j.cct.2024.107514 PubMed PMID: 38537901.

14. Mc Cord KA, Ewald H, Ladanie A, Briel M, Speich B, Bucher HC, et al. Current use and costs of electronic health records for clinical trial research: a descriptive study. CMAJ Open. 2019 Jan 1;7(1):E23–32. doi:10.9778/cmajo.20180096 PubMed PMID: 30718353.

15. Lensen S, Macnair A, Love SB, Yorke-Edwards V, Noor NM, Martyn M, et al. Access to routinely collected health data for clinical trials - Review of successful data requests to UK registries. Trials. 2020 May 12;21(1). doi:10.1186/s13063-020-04329-8 PubMed PMID: 32398093.

16. Lave J, Wenger E. Situated Learning: Legitimate peripheral participation. Cambridge University Press; 1991.

17. Waithira N, Mutinda B, Shah K, Kestelyn E, Bull S, Boggs L, et al. Addressing the gap in health data management skills: an online self-guided course for researchers and health professionals. BMC Med Educ. 2024 Dec 1;24(1). doi:10.1186/s12909-024-06405-y PubMed PMID: 39614233.

18. HDR UK. https://www.hdruk.ac.uk/about-us/who-we-are/. 2026. Who We Are.

19. HDR UK. Transforming Data for Trials Transforming access, analysis and understanding of health data for use in clinical trials. [Internet]. 2026 [cited 2026 Aug 14]. Transforming Data for Trials. Available from: https://www.hdruk.ac.uk/research/research-data-infrastructure/transforming-data-for-trials/

20. HDR UK. https://www.hdruk.ac.uk/study-and-train/learn-with-hdr-uk-futures/. 2026. Learn with HDR UK Futures.

21. UCL. https://www.ucl.ac.uk/population-health-sciences/clinical-trials-and-methodology/study/short-courses/using-healthcare-systems-data-clinical-trials-data-utility-comparisons-ducks-workshop. 2026. Using Healthcare Systems Data in Clinical Trials: Data Utility Comparisons - DUCkS Workshop.

22. Microsoft Corporation. Microsoft Forms [Internet]. 2024 [cited 2025 May 28]. Available from: https://forms.office.com/

23. Dreyfus SE, Dreyfus HL. A five-stage model of the mental activities involved in directed skill acquisition. 1980 Feb.

24. Stephan F. https://www.kaizenko.com/the-dreyfus-model-of-skills-acquisition/ [Internet]. 2017 [cited 2026 Sep 1]. The Dreyfus Model of Skills Acquisition. Available from: https://www.kaizenko.com/the-dreyfus-model-of-skills-acquisition/

25. StataCorp. Stata Statistical Software: Release 19. College Station, TX: StataCorp; 2025.

26. Hsieh HF, Shannon SE. Three approaches to qualitative content analysis. Qual Health Res. 2005 Nov;15(9):1277–88. doi:10.1177/1049732305276687 PubMed PMID: 16204405.

27. McKay AJ, Jones AP, Gamble CL, Farmer AJ, Williamson PR. Use of routinely collected data in a UK cohort of publicly funded randomised clinical trials. F1000Res. 2020 May 4;9:323. doi:10.12688/f1000research.23316.1

28. Sydes MR. https://www.hdruk.ac.uk/news/healthcare-data-in-clinical-trials-the-potential-and-the-challenge/ [Internet]. 2021 [cited 2026 Aug 14]. Healthcare data in clinical trials: the potential and the challenge. Available from: https://www.hdruk.ac.uk/news/healthcare-data-in-clinical-trials-the-potential-and-the-challenge/

29. Wenger E. Communities of practice: Learning as a social system. Cambridge: Cambridge University Press; 1998.

30. Rousse BS, Dreyfus E. Revisiting the six stages of skill acquisition . In: Mangiante ES, Peno K, editors. Teaching and learning for adult skill acquisition. Emerald Publishing ltd; 2021. doi:https://doi.org/10.1108/978-1-64802-502-0

31. Field A. Understanding the dreyfus model of skill acquisition to improve ultrasound training for obstetrics and gynaecology trainees. Ultrasound. 2014 Jan 1;22(2):118–22. doi:10.1177/1742271X14521125

32. Hoffmann JM, Bauer A, Grossmann R. Academic vs. industry-sponsored trials: A global survey on differences, similarities, and future improvements. J Glob Health. 2024 Nov 22;14:04204. doi:10.7189/jogh.14.04204 PubMed PMID: 39575754.

33. Mitchell EJ, Arundel C, Babalis D, Connolly S, Constable L, Goodman K, et al. Supporting the professional development of trial managers: how can you help? Trials. BioMed Central Ltd; 2025. doi:10.1186/s13063-025-09199-6 PubMed PMID: 41137154.

34. Mitchell EJ, Goodman K, Wakefield N, Cochran C, Cockayne S, Connolly S, et al. Clinical trial management: a profession in crisis? Trials. BioMed Central Ltd; 2022. doi:10.1186/s13063-022-06315-8 PubMed PMID: 35477835.

35. UKTMN. https://www.tmn.ac.uk/resources/73-competency-framework-for-trial-management [Internet]. 2023 [cited 2026 Sep 2]. Competency Framework for Trial Management. Available from: https://www.tmn.ac.uk/resources/73-competency-framework-for-trial-management

36. Yedro S, Tinari E, Napolitano D, Wlderk G, Ribaudi E, Giannone L, et al. DISYNCRO: Perceived roles of clinical study coordinators and data managers: results from a web-based survey of professionals from contract research organizations. Contemp Clin Trials Commun. 2025 Oct 1;47. doi:10.1016/j.conctc.2025.101533

37. Keinänen AL, Lähdesmäki R, Juntunen J, Tuomikoski AM, Kääriäinen M, Mikkonen K. Effectiveness of mentoring education on health care professionals´ mentoring competence: A systematic review. Nurse Education Today. Churchill Livingstone; 2023. doi:10.1016/j.nedt.2023.105709 PubMed PMID: 36638727.

38. Tibble H, Alyami RA, Bush A, Cunningham S, Julious S, Price D, et al. Using routine primary care data in research: (in)efficient case studies and perspectives from the Asthma UK Centre for Applied Research. BMJ Health Care Inform. 2025 Jan 9;32(1). doi:10.1136/bmjhci-2024-101134 PubMed PMID: 39788753.

39. Sathyanarayanan A. The use of routinely collected healthcare records for outcome assessment in clinical trials: a UK perspective. Current Medical Research and Opinion. Taylor and Francis Ltd.; 2024. p. 887–92. doi:10.1080/03007995.2024.2333441 PubMed PMID: 38511976.

582 **Table 3: HSD Expertise by job category**

		Job category				
HSD domains	**Aspects of domains**	**Trial Management**	**Data Management**	**Statistician**	**Clinical**	**Other**
Trial Design and Management	Costing		✓	✓	✓	
	Determining data required for analysis		✓	✓	✓	✓
	Inclusivity				✓	
	Risk Assessment		✓	✓	✓	
	Safety monitoring			✓	✓	✓
Approvals	When to apply	✓			✓	
	differences between providers	✓	✓		✓	
	Differences between devolved nations	✓			✓	
	Nation/regional/local data				✓	
	Information governance		✓	✓	✓	
	Approval duration	✓	✓	✓	✓	
	Amendments	✓	✓		✓	
Informed Consent	Lawful basis for access				✓	
	Role of section 251		✓		✓	
	Importance of transparency	✓	✓	✓	✓	
	Right to withdraw	✓	✓	✓	✓	
	HRA guidance for PIS and GDPR	✓		✓	✓	
	Managing participants' mental capacity				✓	
Site feasibility and selection	Site feasibility assessment				✓	
	Site selection process				✓	
	Identifying populations for recruitment		✓	✓	✓	
Contracts and agreements	Data controllers					
	Data processors					
	Role of your organisation					
	Importance of contacting		✓	✓		
	DPIAs		✓			
	Data sharing agreements		✓	✓		
	Data processing agreements					
	Timelines for data sharing					
Protocol development	RECORD guidelines		✓	✓	✓	
	CONSORT Routine guidelines	✓		✓	✓	
Data management and analysis	Primary care datasets		✓	✓	✓	
	Registries audits and clinical cohorts		✓		✓	
	Using TREs/SDEs		✓	✓	✓	
	Sample size calculations			✓	✓	
	Using algorithms		✓	✓	✓	
	Coding and phenotype libraries		✓	✓	✓	
	Synthetic data					
	Derived data		✓	✓		
	Data utility		✓	✓		
	Dealing with missing data		✓	✓		
Archiving and data sharing	Understanding provider requirements		✓			
	Feeding back to data providers		✓			

	Data ownership	✓				
	De-identification and anonymisation	✓	✓			
	Onward sharing of data	✓	✓			
TOTAL		**9**	**28**	**24**	**29**	**2**

583

584

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Additionalfile1survey.pdf
- Additionalfile2.docx
- Additionalfile3.docx