MSc Digital Transformation for the Health and Care Professions

BMDS7005

Optimising Healthcare Data Flows and Interoperability: Exploring Challenges and Opportunities in NHS Wales

Dissertation submitted in partial fulfilment of the award of
Master of Science in Digital Transformation for the Health and Care Professions

Jay Hier-Jones
2111607
6th May 2025

Acknowledgements

This is dedicated to my mother who I lost and my son who I gained during the undertaking of this qualification.

* * *

I would like to extend my heartfelt thanks to:

My husband and my dad for their love, encouragement and the time they have spent with Elwyn, allowing me to concentrate on completing this work. I am also deeply grateful for the unwavering support from my incredible sister and friend, Liz Jones.

My Supervisor, Dr. Ben Duxbury for his invaluable guidance, thoughtful feedback, and the time he has dedicated to reviewing my work. His constructive insights have been instrumental in keeping me on the right path.

My manager and team for listening to my ideas, supporting me through the stresses, and giving me the time to invest in this project.

My original Group 4 who I abandoned for maternity leave – you supplied much needed moral support and comic relief.

PTHB Library Services, notably Jane Parry, for her advice on literature searches and the **many** book loans.

This journey would not have been the same without you.

Abstract

This dissertation investigates the challenges and opportunities associated with healthcare data flows and interoperability within NHS Wales, with a particular focus on Powys Teaching Health Board (PTHB). Through a mixed-methods approach, combining qualitative insights from healthcare professionals' survey responses with quantitative analysis of documentary sources, the research highlights the barriers posed by fragmented data systems, inconsistent data entry practices, and reliance on manual processes. The findings touch upon the potential of interoperable systems to enhance clinical decision-making, reduce delays in care delivery, and optimise resource allocation. By integrating technological advancements with real-world needs, this study recommends investment in modern systems and IT infrastructure, adoption of interoperability standards, and a cultural shift toward digital literacy, increased collaboration and data-sharing to achieve seamless data integration and improved patient outcomes. The research contributes to the ongoing digital transformation efforts within NHS Wales, aligning with broader NHS and UK Government strategies to create a more connected, patient-focused healthcare system.

Declaration

This work has not previously been accepted in substance for any degree and is not being concurrently submitted in candidature for any degree.

I confirm this essay is my own original work and does not include text created by a generative artificial intelligence tool.

Signed……Jay C. Hier-Jones………………………………………

Date……………01.05.2025…………………………………

Table of Contents

Chapter 1: Introduction 9

- Purpose of Research Activity 9
- Background and context 9
  - Big Data and Modern Architectures 10
  - Interoperability and Standards 11
  - Global Innovations in Digital Healthcare 12
  - Digital Transformation and Interoperability within the NHS 13
  - Digital Maturity within NHS Wales and PTHB 14
- Aim and Objectives of Research Activity 15
- Summary 15

Chapter 2: Literature Review 16

- Introduction 16
- Literature Search Strategy 16
- Literature Review Findings 18
  - Technical Factors 19
  - Organisational Factors 23
  - Macroenvironmental Factors 26
- Summary 29

Chapter 3: Methodology and Methods 30

- Introduction 30
- Research Approach 30
- Research Design 31
  - Access and Sampling 33
- Data Collection Methods 35
  - Online Questionnaire 35
  - Documentary sources 37
- Data Handing and Analysis 37
  - Qualitative Analysis 38
  - Quantitative Analysis 39
- Research Integrity and Rigour 40
- Summary 41

Chapter 4: Results 43

- Introduction 43
- Questionnaire Participant Background 43

Research Objective 1. How do healthcare professionals perceive the interoperability of the systems they use, and what impact does this have on their work? 45
  Perceptions of Interoperability.....45
  Consequences of Siloed Systems.....50
Research Objective 2. In what ways do data quality and accessibility influence decision-making and patient outcomes? 56
Research Objective 3. What factors are perceived as barriers or enablers to improving data flows and interoperability? 62
Conclusion.....66
Chapter 5: Discussion, Conclusions and Implications 68
Introduction.....68
Meeting of aim and objective.....68
Significance of the study: 69
Strengths and limitations.....69
Discussion of Findings 70
Implications for Policy, Practice, Education and Research 73
  Policy Recommendations 74
  Practice Recommendations.....75
  Education Recommendations 75
  Research Recommendations.....75
Concluding Remarks 76
References.....77
Appendices.....100
Appendix 1. HIMSS Electronic Medical Records Adoption Model (EMRAM) Model Stages (HIMSS, 2022).....100
Appendix 2. Literature Review Appraisal 103
Appendix 3. Classification of informatics study types by broad study foci, the stakeholders most concerned and the method(s) most commonly employed.....112
Appendix 4. Count of full time equivalent (FTE) NHS Wales staff by staff group and organisation for June 2024. (Stats Wales, 2024).....113
Appendix 5. Research Data Collection Mediums Strengths and Weaknesses.....114
Appendix 6. UWTSD Ethical Approval Form 115
Appendix 7. PTHB Research Innovation and Improvement Ethical Approval.....132
Appendix 8. Questionnaire; Participant Consent Information 133
Appendix 9. Python (Pyspark) Code and Full Results of Likert-Scale Question Pair Correlation.....140
Appendix 10. SQL Code for PTHB Data Analysis and Engineering Team Azure DevOps 144

Appendix 11. SQL Code for PTHB Activity Duplicated Between WPAS and WCCIS 144
Appendix 12. SQL Code for PTHB Multiple EHR System Demographics 145
Appendix 13. Python (Pyspark) Code for PTHB Multiple EHR System Demographic Correlation 146
Appendix 14. Welsh Health Board Allocations Summary: Baseline discretionary capital funding 2025-26 147
Appendix 15. MSc Project Plan Gantt Chart 148

Table of Figures

Figure 1. PRISMA 2020 flow diagram for new systematic reviews which included searches of databases and registers only 17
Figure 2. Graph highlighting the volume of results per search terminology within PubMed per article publish year 18
Figure 3. Research Aim and Objectives 30
Figure 4. Research Design, Data Collection Methods and Their interactions 32
Figure 5. Example Code-to-Theme Process 39
Figure 6. Questionnaire Responses Relating to Data Use and Experience Per Role Type and Organisation 44
Figure 7. Word Cloud for Participant Service Departments 45
Figure 8. Data sources used by questionnaire participants by role type 47
Figure 9. MS Forms Chart Displaying Percentage of Likert Scale Responses per Question 48
Figure 10. Selected data standards/protocols used within questionnaire participants organisation by role type 50
Figure 11. Graphical comparison of Activity per WCCIS and WPAS 53
Figure 12. Selected aspects affected by data flow problems within questionnaire participants role/organisation by role type 55
Figure 13. Comparison of Mismatching record counts per System and Field 59
Figure 14. Correlation Matrix Between Systems Based on Mismatch Counts 60
Figure 15. Dendrogram for Hierarchical Clustering of System Based on Mismatch Correlations 60
Figure 16. Selected factors impacting the adoption and use of real-time data flows and interoperable systems by questionnaire participant role type 66
Figure 17. Healthcare Data Interoperability: Current Challenges vs. Ideal Future 76

Table 1. Summary of Common Digital Health System Challenges and Contributing Factors (Aerts and Bodgan-Martin, 2021 12
Table 2. SPICE Framework used for Literature Review Search 16
Table 3. Example Ovid Database Literature Search (MEDLINE(R) ALL) 16
Table 4. Literature Review Themes and Article Frequency 19
Table 5. Target Sample Organisation and Role Type 34
Table 6. Questionnaire Responses Per Organisation and Role Type 43
Table 7. Thematic Analysis: Potential of Interoperable Systems/RealTime Data Flows 45
Table 8. Thematic Analysis: Current State of NHS Interoperability 46
Table 9. Number of Likert Scale Responses per Question 48
Table 10. Likert-Question Pair Correlation 49
Table 11. Thematic Analysis: Consequences of Siloed Systems 50

Table 12. Summary Results of PTHB Data and Analysis Team AzureDevOps Active Manual Processing Jobs 51
Table 13. Activity Duplicated Between WPAS and WCCIS per Year and Specialty........................52
Table 14. Summary of Patient Records Per PTHB EHR System 58
Table 15. Summary of Potential Duplicates Per PTHB EHR System 58
Table 16. Number of Unique Patient Records per Field with System Matches and Mismatches.......58
Table 17. Thematic Analysis: Barriers to Interoperability 62
Table 18. Thematic Analysis: Enablers to Interoperability 64
Table 19. Study Recommendations 73

List of abbreviations and acronyms

ADO	**Azure DevOps**
AI	**Artificial Intelligence**
API	**Application Programming Interface**
CASP	**Critical Skills Appraisal Program**
CDSS	**Clinical Decision Support Systems**
COREQ	**Consolidated Criteria for Reporting Qualitative research**
DGH	**District General Hospital**
DHCW	**Digital Health Care Wales**
DIP	**Data Intelligence Platform**
DPIA	**Data Protection Impact Assessment**
EHI	**Electronic Health Information**
EHR/EPR/EMR	**Electronic *Health/Patient/Medical* Record**
EMRAM	**Electronic Medical Record Adoption Model**
EPMI	**Electronic Patient Master Index**
FDP	**Federated Data Platform**
FEDIP	**Federation for Informatics Professionals**
FHIR	**Fast Healthcare Interoperability Resources**
HES	**Health Episode Statistic**
HIMSS	**Healthcare Information Management System Society**
HIT	**Health Information Technology**
HL7	**Health Level Seven International**
IG	**Information Governance**
IT	**Information Technology**
LHS	**Learning Health System**
ML	**Machine Learning**
MMAT	**Mixed Method Appraisal Tool**
NAO	**National Audit Office**
NCCID	**National COVID-19 Chest Imaging Database**
NDAP	**National Data Analytics Platform**
NDR	**National Data Resource**
NHS	**National Health Service**
NLP	**Natural Language Processing**
NPfIT	**National Programme for IT**
PTHB	**Powys Teaching Health Board**

PRISMA	**Preferred Reporting Items for Systematic reviews and Meta-Analyse**
RCT	**Randomized Control Trial**
SCR	**Shared Care Record**
SDE	**Secure Data Environments**
SQL	**Structured Query Language**
TRE	**Trusted Research Environment**
UWTSD	**University of Wales Trinity Saint David**
WG	**Welsh Government**

Data Role Identifiers

A-DN	Administrative/Non-patient facing – District Nursing
A-MH	Administrative/Non-patient facing – Mental Health
A-PC	Administrative/Non-patient facing – Primary Care
A-SN	Administrative/Non-patient facing – Specialist Nursing
A-T	Administrative/Non-patient facing – Therapies
A-O	Administrative/Non-patient facing – Other/Unknown
C-M	Clinician/Patient-facing – Medical
C-SN	Clinician/Patient-facing – Specialist Nursing
C-T	Clinician/Patient-facing – Therapies
CI	Clinical Informatics
DT	Digital/Technical
SMC-C	Senior Manager Clinical – Community
SMC-T	Senior Manager Clinical – Therapies
SMC-PH	Senior Manager Clinical – Public Health
SMC-O	Senior Manager Clinical – Other/Unknown
SMC-WC	Senior Manager Clinical – Womens and Childrens
SMA-F	Senior Manager Administrative – Finance
SMA-PC	Senior Manager Administrative – Planned Care
SMA-O	Senior Manager Administrative – Other/Unknown

Chapter 1: Introduction

This chapter outlines the research purpose and context, covering modern data architectures, interoperability[1], and digital healthcare innovations globally and within the National Health Service (NHS). The chapter concludes with the specific research aim and objectives.

Purpose of Research Activity

Existing research indicates that effective data sharing and interoperability can help optimise decision-making processes and improve patient outcomes (Choun and Petre, 2022; Perlman and Elsner, 2019; Welsh Government, 2023). The Welsh Government's (WG) *Digital and Data Strategy for Health and Social Care* (2023) supports this, calling for standardised data infrastructure and cross-sector collaboration to facilitate seamless sharing of patient information.

This study explores data flows and interoperability within NHS Wales, with a particular focus on Powys Teaching Health Board (PTHB). Noting the complementary nature of process (how something is achieved) and outcome (the effectiveness of what has been achieved), Robson (2024) describes evaluation as a tool for both assessment and improvement, contending that measuring outcomes alone is insufficient. Accordingly, this study seeks to understand how effectively data is shared and used, and potential factors influencing interoperability. The study aspires to provide evidence to support data flow optimisation, particularly within rural healthcare contexts like that of PTHB.

Background and context

There is growing recognition of the vital role data plays in healthcare (World Economic Forum, 2024), particularly real-time data[2], which can improve care delivery by providing clinicians access to the right information at the right time (Savitz *et al.,* 2024; National Academy of Medicine*,* 2018; Hammond *et al.,* 2020). Although efforts to digitise and implement electronic patient records (EPRs) began over 30 years ago in the United States, widespread adoption was slow until the 2009 *American Recovery and Reinvestment Act* (Honavar, 2020). In the United Kingdom (UK), a significant push for digitisation and record standardisation occurred in 2002 with the launch of the National Programme for Information Technology (NPfIT)[3],

---

[1] Interoperability within healthcare is defined by the Healthcare Information Management System Society (HIMSS) as 'the ability of different information systems, devices, and applications to access, exchange, integrate, and cooperatively use data in a coordinated manner, within and across organizational, regional, and national boundaries.' (HIMSS, 2024a, para.1)

[2] Real-time data can be defined as 'information that is made available for use as soon as it is generated.' (Qlik, n.d.)

[3] NPfIT was described as the 'world's largest civil information technology (IT) programme' (Justinia, 2017, p.2), but dismantled in 2011. It was heavily criticised largely due to overambitious scale, resistance against top-down decision-making, inadequate engagement and poor change management approaches (Justinia, 2017).

laying the foundation for modern NHS EPRs (Justinia, 2017).

Large-scale efforts to digitally integrate health and social care are driven by the objectives of the Quintuple Aim, an increasingly prominent framework for guiding digital transformation and healthcare policy (Elizondo, 2024). Evolving from the Institute for Healthcare Improvement's Triple Aim (improving population health, patient experience, and cost reduction), the Quintuple Aim adds two additional priorities: addressing workforce burnout and promoting health equity (Nundy *et al.,* 2022; Mate, 2022). As Health Information Technology (HIT) and Electronic Health Record (EHR) systems have evolved, so has the volume, complexity and reliance on healthcare data, necessitating modern data architectures to manage and interpret vast datasets effectively (Choun and Petre, 2022; Idowu *et al.,* 2023; Ristevski and Chen, 2018).

Big Data and Modern Architectures

Industry sources often cite that 90% of the world's data was generated in the last two years (Choun and Petre, 2022; Phiri, 2022; Hackenberger, 2019). This figure appears to have originated from IMB's 2011 *Global CMO Stu*dy, yet its recurrent reference underscores the recent exponential rise in data volumes. The increase in volume, velocity and variety of data are key characteristics used to describe Big Data[4], known as the '3Vs' (Shahid and Sheikh, 2021, p.578). Other words associated with Big Data include variability, veracity, value (Ristevski and Chen, 2018), as well as vulnerability, volatility and visualisation (Idowu *et al.,* 2023; Banerjee *et al.,* 2018). To manage and process Big Data effectively, innovative architectures like data lakes[5] and lakehouses[6] have emerged (Idowu *et al.,* 2023; Ambrust *et al.,* 2021).

However, growing volumes of data also demand efficient, interoperable systems capable of delivering real-time, patient-centred insights (Welsh Government, 2023). Mistry *et al.'s* (2022) research for the King's Fund highlights the importance of interoperability and the joining of people, services and organisations, noting that missing information creates challenges for staff, frustrates patients and puts them at risk. The reason for the current 'fragmented' data landscape (Federation for Informatics Professionals, 2024, p.4) and disjointed care

---

[4] De Mauro *et al*. (2016, p.127) define Big Data as 'the information asset characterised by such a high volume, velocity and variety to require specific technology and analytical methods for its transformation into value'.

[5] Data lakes are large pooled, low-cost, unstructured databases for storing Big Data (Idowu et al., 2023).

[6] Lakehouses are directly accessible data management systems which combine the flexibility, cost-efficiency, and scalability of a data lake with the benefits of a traditional data warehouse (Ambrust *et al.,* 2021; Oreščanin and Hlupic, 2021). Lakehouses are expected to replace data warehouse architectures in the future due to their unified open format for storage, processing, and analytics, their simplified data architectures, addressing challenges such as data reliability, total cost of ownership and data lock-in (Armbrust *et al*., 2021).

experience is attributed to legacy[7] health systems and siloed culture (Mistry *et al.,* 2022). Isolated systems provide a limited view of data, but interoperability and integration[8] can help bridge these silos (Patel, 2019).

Interoperability and Standards

Efforts to address interoperability challenges include the development of global health data standards. In 1989, Health Level Seven International (HL7) developed a framework for exchanging, integrating and retrieving electronic health information (HIMSS, 2024). Interoperability has three paradigms, ranked by increasing complexity and desirability: transport, structured and semantic (Braunstein, 2022). Semantic interoperability, where data is standardised and understandable to both sender and receiver, is often referred to as 'the holy grail' of healthcare interoperability (Silsand, 2017, p.142). Achieving this requires the use of standardised terminologies[9] and interoperability standards, which provide a common language, structure and set of system rules (Braunstein, 2022; Choun and Petre, 2022; Gliklich and Leavy, 2019). However, due to the multi-level and multi-sectoral complexity within the NHS, achieving semantic interoperability goes beyond data exchange and coding, requiring collaboration with various stakeholders, from application design to knowledge sharing (Braunstein, 2022; Das and Hussey, 2022).

'Fast Healthcare Interoperability Resources' (FHIR), developed by HL7, is the emerging standard for electronically exchanging healthcare information, designed to facilitate seamless patient data sharing across different systems (Health Level 7 International, 2023). Braunstein (2022) describes FHIR as being widely accepted as the best available approach for achieving interoperability, providing a universal way to represent[10] and access[11] data. In 2023, HL7 FHIR was formally mandated as the foundational interoperability standard for all IT and digital systems across NHS Wales (Welsh Government, 2023a). The Healthcare Information Management System Society (HIMSS) (2024) suggests we are entering a new era of healthcare communication, with the next phase of national interoperability progressing in line

---

[7] Legacy systems can be defined as IT systems which are built on outdated technologies but are still relied upon for daily operations. (Gartner, n.d.)

[8] Data integration is the process of combining data from multiple disparate sources to create single, unified view. (Microsoft Azure, n.d.)

[9] Examples include SNOMED CT (a comprehensive clinical terminology), ICD-10 (for coding diseases and health conditions), DICOM (Digital Imaging and Communications in Medicine, for medical imaging), and LOINC (Logical Observation Identifiers Names and Codes, for laboratory tests and clinical observations) (Braunstein, 2022)

[10] FHIR is typically represented using JavaScript Object Notation (JSON), a text format that is easy to understand and used for storing and transporting data (W3Schools, n.d.)

[11] FHIR data can be accessed via an Application Programming Interface (API), which is a set of definitions and protocols for building and integrating software and helps to mediate and transfer. It uses REST (Representational State Transfer) architecture, which is a set of architectural constraints and guidelines which can be fast to implement, lightweight, and scalable. (RedHat, 2020)

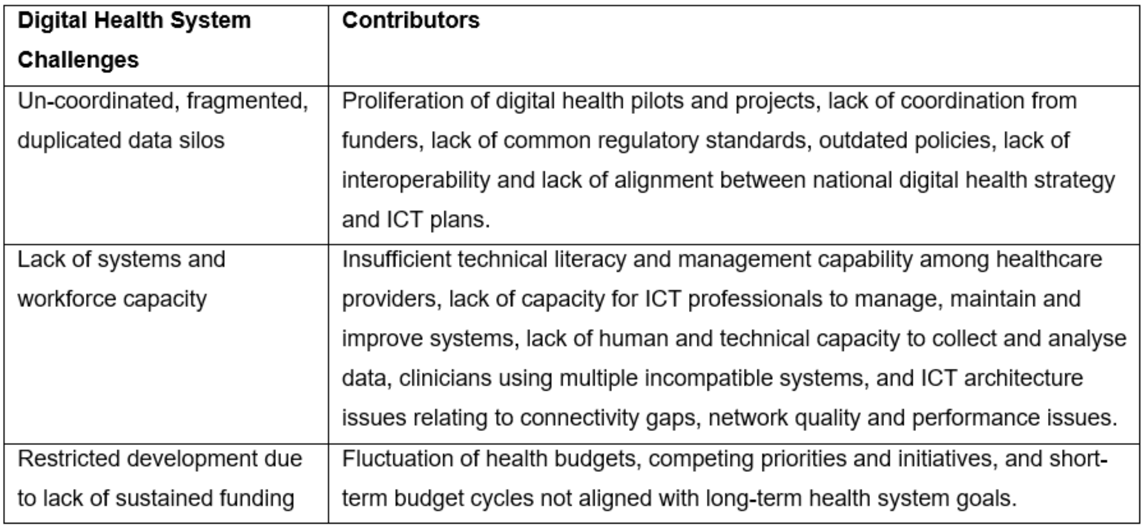
with the collective ambition[12] of the Trusted Exchange Framework and Common Agreement.

Global Innovations in Digital Healthcare

Globally, there is a growing emphasis on prioritising patient-centred care, with increased efforts to involve patients in the co-design of healthcare services to better meet their needs (World Health Organization, 2015). Empowering patient participation in healthcare decision-making, including access to their own data, could bridge the gap between patients, science, and technology, enhancing both individual care and the broader healthcare system (Choun and Petre, 2022). Simultaneously, digital transformation through data-driven technologies such as artificial intelligence (AI) and machine learning (ML), has the potential to modernise healthcare services and improve patient outcomes (Siegel, 2024; Welsh Government, 2023).

The COVID-19 pandemic accelerated the adoption of tools like remote monitoring and telehealth, strengthening their potential to address global healthcare challenges, such as rising costs, workforce shortages, increasing chronic conditions, and inequitable outcomes (Choun and Petre, 2022; Sheikh *et al.,* 2021). The World Economic Forum's 2024 *Transforming Healthcare Report* suggests that digital health innovation and data interoperability can help tackle these challenges. However, as the National Audit Office (NAO) (2020) points out, realising these benefits requires a strong digital infrastructure and integrated data systems. Aerts and Bogdan-Martin (2021) document similar digital health system challenges, summarised in Table 1, and note that the successful development and management of digital health systems hinge on the ‘building blocks’ (p.3) of national strategy, regulatory frameworks, infrastructure, interoperability, partnerships, and financing, all supported by visionary leadership, intersectoral collaboration, and robust ICT.

Table 1. Summary of Common Digital Health System Challenges and Contributing Factors (Aerts and Bodgan-Martin, 2021



Digital Health System Challenges	**Contributors**
Un-coordinated, fragmented, duplicated data silos	Proliferation of digital health pilots and projects, lack of coordination from funders, lack of common regulatory standards, outdated policies, lack of interoperability and lack of alignment between national digital health strategy and ICT plans.
Lack of systems and workforce capacity	Insufficient technical literacy and management capability among healthcare providers, lack of capacity for ICT professionals to manage, maintain and improve systems, lack of human and technical capacity to collect and analyse data, clinicians using multiple incompatible systems, and ICT architecture issues relating to connectivity gaps, network quality and performance issues.
Restricted development due to lack of sustained funding	Fluctuation of health budgets, competing priorities and initiatives, and short-term budget cycles not aligned with long-term health system goals.

---

[12] The Trusted Exchange Framework and Common Agreement (TEFCA) collective ambition is ‘to achieve an interconnected health system where data informs care at the point of service and contributes to a larger ecosystem of health intelligence that benefits all (HIMSS, 2024, p.4)

Successful global case studies of digital transformation in healthcare include India's ReMiND project, which is reported to have significantly improved maternal and newborn health, Chile's AccuHealth, which uses AI for chronic condition management (Perlman and Elsner, 2019), and Estonia's X-Road[13] data exchange layer (Choun and Petre, 2022), which has been widely adopted, including within South American and Asian countries (Plantera, 2024, 2024a). France's Health Data Hub created in 2019, also facilitates secure access and integration of multiple data sources, aiming to improve healthcare delivery (Health Data Hub, 2021). However, challenges persist, with technical limitations and regulations, designed in a pre-internet era, posing barriers to safe and effective information exchange (Braunstein, 2022).

Digital Transformation and Interoperability within the NHS

Specific challenges in data sharing and real-time data access exist within the UK, with NHS IT systems being termed 'outdated and inefficient' (NAO, 2020, p.6). The 2020 NAO report highlighted the ever-present need for digital transformation given previous unsuccessful attempts. More recently, an independent investigation revealed the NHS' critical state, including poor waiting list performance and health inequalities (Darzi, 2024), prompting the Prime Minister to state that the NHS must 'reform or die' (Starmer, 2024). The Department of Health and Social Care's '*Data Saves Lives'* Strategy (2022a) outlines a commitment for improving health and social care interoperability, underpinned by FAIR (findability, accessibility, interoperability and reusability) principles. This strategy promotes open, practical data standards and supports the implementation of Secure Data Environments (SDEs), centralised platforms designed to securely link diverse data sources for research and analysis (Department of Health and Social Care, 2022b). NHS England's Federated Data Platform (FDP) also aims to improve data quality by bringing siloed data together (Ghafur *et al.,* 2023). Initiatives for integrated systems and shared care records (SCR) within NHS England promote seamless data exchange between healthcare staff, patients, and carers, supporting timely, safe decision-making (Department of Health and Social Care, 2022).

Klein (2013) highlights a rhetoric-reality gap within NHS policies and programmes, pointing to a disconnect between their stated objectives and actual achievements. Multiple initiatives have failed to meet the original requirements (Federation for Informatics Professionals, 2024), and this pattern of unmet goals specific to data interoperability and digital integration has persisted since 1999 (Wilson and Davies, 2021), illustrating the ongoing challenge to achieve the intended outcomes. Darzi (2024) also highlighted limited progress for integrated systems, stating that 'including "integrated care" in the title of organisations does not make it thus' (p.78). Within the wider digital context, Banerjee *et al.* (2018) quote Wachter, noting that

---

[13] The X-Road open-source software enables systems to securely share and access data directly through a standardised interface, eliminating the need for a central database (Choun and Petre, 2022).

there is ‘more promise than reality’ (p.162) and more hype than evidence, with barriers related to ethical AI implementation, infrastructure, workforce, training, and culture also hindering successful implementation of SCRs (Elizondo, 2024).

Digital Maturity within NHS Wales and PTHB

In today's healthcare environment, optimising operational efficiency is critical for delivering high-quality patient care, yet healthcare systems often lag in digital innovation (World Health Forum, 2024); arguably, this is particularly pronounced in rural areas like Powys. Reliance on outdated, siloed IT systems that do not interconnect is seen as unsustainable in an increasingly digital world (National Audit Office, 2020). The National Data Resource (NDR) Programme emerged from WG’s long-term vision in *A Healthier Wales*: *Our Plan for Health and Social Care* (2018), which called for a seamless, person-centred health and care system supported by integrated digital technologies. Although the NDR aimed to enable large-scale information sharing by 2020, the NDR Data Analytics Platform (NDAP) has only recently become available, with organisational use still limited (DHCW, 2024).

The HIMSS Electronic Medical Record Adoption Model (EMRAM) provides an eight-stage maturity model, offering a roadmap for digital transformation and evaluating care institutions worldwide (HIMSS, 2022; Appendix 1). Within NHS England, very few trusts have achieved the highest level (Digital Health Intelligence, 2023), while in Wales, most health boards are at stage 1. Notably, both Cwm Taf Morgannwg and PTHB are at Stage 0 (John, 2024; HIMSS, 2023). John (2024) notes the main gap identified across Wales is the lack of a solution which provides a unified view of patients records. Nonetheless, PTHB has recently advanced its data infrastructure by adopting lakehouse technology and implementing a central data repository via the Databricks Data Intelligence Platform (DIP) (Databricks, 2024). This could be considered a significant step towards leveraging advanced analytics and future-proofing digital infrastructure.

PTHB faces significant operational challenges, including financial pressures, which in November 2024, under the WG’s Joint Escalation and Intervention Arrangements, led to an increase in escalation status from ‘enhanced monitoring’ to ‘targeted intervention’ for finance, strategy and planning (PTHB, 2024). WG’s baseline capital health board allocations for 2025/26 (Appendix 14) reveal that PTHB receives less funding compared to other Welsh Local HBs[14]. As the primary healthcare provider for Wales’ largest county, PTHB is responsible for delivering NHS services across a wide geographic but sparsely populated area (ONS, 2024). Without a district general hospital (DGH) patients face significant travel times to access acute and emergency care, and consequently, a significant portion of Powys'

---

[14] This allocation is reportedly based on several considerations, including the size of the local population, age profiles, and levels of ill health and poverty (Lloyd-Jones, 2024).

patient care is commissioned to other health boards within Wales and England (Powys County Council, 2014). The differing policies in these regions also create challenges in ensuring that patients receive equitable and high-quality care across providers (Powys County Council, 2014).

Aim and Objectives of Research Activity

This project aimed to answer the following research question; What factors influence healthcare data interoperability and the utilisation of real-time data flows, and how effectively do they support healthcare delivery within the NHS?

To achieve this aim, the project sought to address the following sub-questions, presented as project objectives:

- *How do healthcare professionals perceive the interoperability of the systems they use, and what impact does this have on their work?*
  Rationale: To explore user experiences of system interoperability, assess its impact on healthcare delivery, and gain insight into existing data mechanisms
- *In what ways do data quality and accessibility influence decision-making and patient outcomes?*
  Rationale: To assess the importance of data reliability and availability in supporting informed clinical decisions and patient care.
- *What factors are perceived as barriers or enablers to improving data flows and interoperability?*

  Rationale: To identify key influences on data management practices and areas for potential enhancement.

Summary

This thesis explores the challenges and opportunities associated with healthcare data flows and interoperability within the NHS. The chapters progress logically from the study's scope and background, introduced in this chapter, to a review of existing literature, followed by the methodology and findings, before concluding with the discussion, final reflections and recommendations.

Chapter 2: Literature Review

Introduction

The previous chapter outlined the research aim, objectives, and context, highlighting key concepts such as Big Data, interoperability, and digital transformation within the NHS. This chapter reviews relevant literature on NHS data interoperability and real-time data exchange. It begins with the literature search strategy, discusses common themes, and concludes with a summary of the findings.

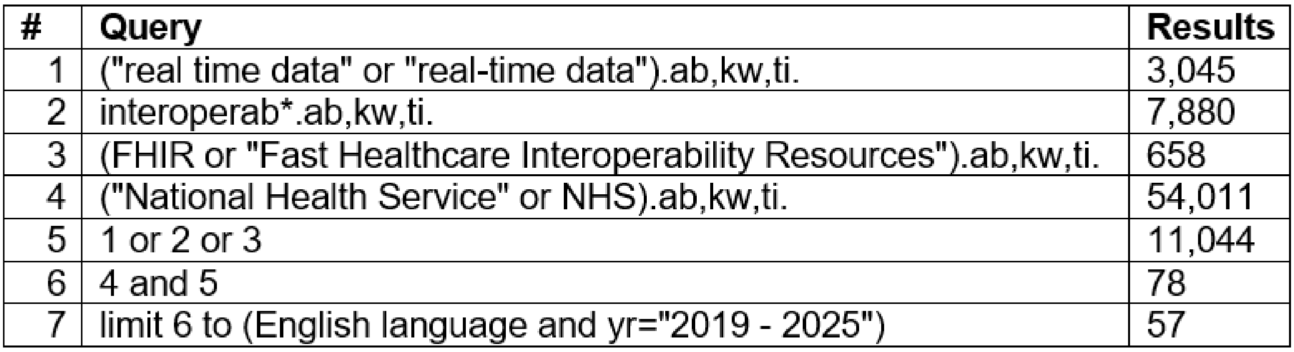
Literature Search Strategy

The literature review was conducted systematically, examining full-text articles published between 2019 and 2024. Searchers were performed using the Ovid Interface for Medline, EMBASE, and Health Management Information Consortium databases, alongside manual searches within IEEE Xplore, PubMed, ScienceDirect and BMJ Journals. Guided by the SPICE Framework (Booth, 2004), test searches were used to refine the search strategy detailed in Table 2. The search was restricted to NHS settings, given the high volume of results for 'healthcare'. Other search words included 'interoperability', 'FHIR' and 'real-time data'. Table 3 depicts an example OVID search.

Table 2. SPICE Framework used for Literature Review Search

Setting	NHS healthcare system.
Perspective	NHS stakeholders, including healthcare professionals, policymakers, and patients.
Intervention	Implementation of healthcare interoperability frameworks, interoperable systems and real-time data exchange mechanisms.
Comparison	Situations with/without interoperable systems or real-time data flows.
Evaluation	Identification of contributors and barriers to interoperability and data exchange.

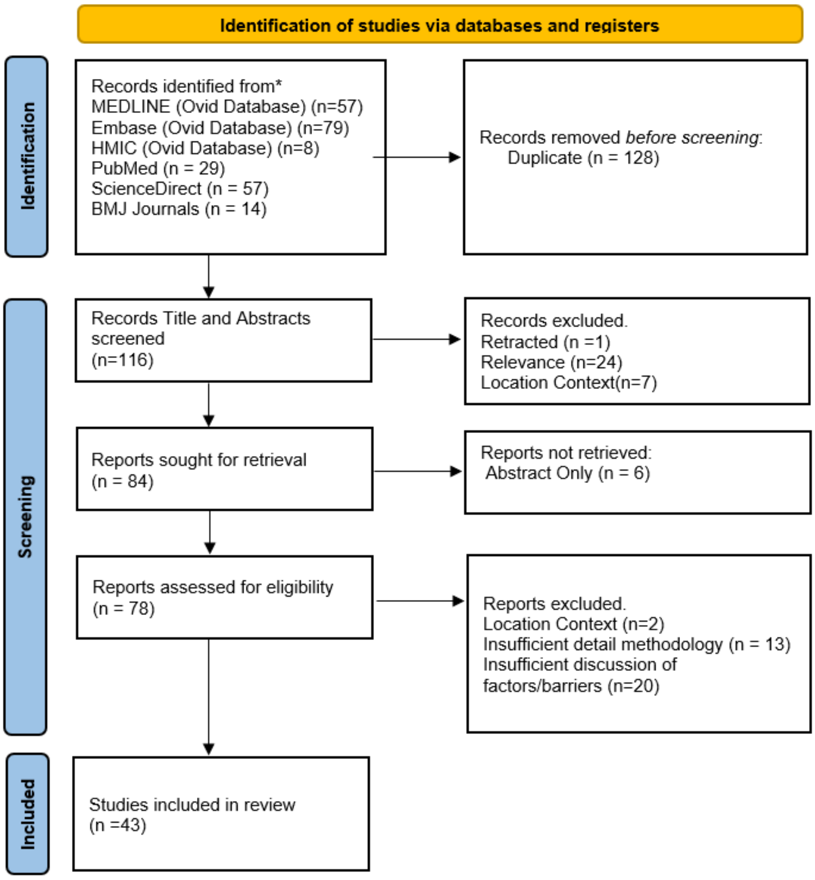
Table 3. Example Ovid Database Literature Search (MEDLINE(R) ALL)



#	Query	Results
1	("real time data" or "real-time data").ab,kw,ti.	3,045
2	interoperab*.ab,kw,ti.	7,880
3	(FHIR or "Fast Healthcare Interoperability Resources").ab,kw,ti.	658
4	("National Health Service" or NHS).ab,kw,ti.	54,011
5	1 or 2 or 3	11,044
6	4 and 5	78
7	limit 6 to (English language and yr="2019 - 2025")	57

Additional literature was identified through manual journal searches and the 'snowball technique' of citation tracking (Booth, 2004). Mendeley Reference Manager was used to organise, detect and remove duplicate articles. Following PRISMA guidelines (Page *et al.,* 2021), only peer-reviewed, English-language original research studies were included, encompassing both empirical and non-empirical literature, such as research studies and expert opinion articles. While the hierarchy of evidence typically applies to clinical research (Aveyard, 2023), a balanced approach incorporating opinion pieces and evidence-based studies was used to support evidence-based practice, which Moule (2020) describes as the integration of the best available research evidence with clinical expertise and patient preferences. Limitations such as database selection, language, and keyword constraints may have excluded relevant studies using synonymous terms or published in niche sources (Aveyard, 2023).

A summary of reviewed literature is provided in Appendix 2, and the PRISMA Diagram in Figure 1 outlines the search process. Titles and abstracts were screened for relevance before full-text review. Studies outside the NHS context, lacking reference to data flows or interoperability, lacking methodological detail or focused primarily on clinical outcomes were excluded. Quality was assessed using an adapted GRADE framework (Guyatt *et al.,* 2008), the Mixed Methods Appraisal Tool (MMAT) (Hong *et al.,* 2018), and the CRAAP Test (Meriam Library of CSU, Chico, 2010) helped assess the currency, relevance, authority, accuracy, and purpose of the literature. The next section presents the literature review findings.



From: Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021;372:n71. doi: 10.1136/bmj.n71

Figure 1. PRISMA 2020 flow diagram for new systematic reviews which included searches of databases and registers only

Literature Review Findings

The literature displayed commonality across academic and practice-based discourse, emphasising the importance of interoperability and real-time data exchange. Figure 2 (created by the author) illustrates an upward trend in PubMed publications on healthcare interoperability since 2015, coinciding with the emergence of FHIR-related articles. However, UK-specific research remains relatively sparse, especially within the context of NHS Wales. Warren *et al.* (2019) also noted the lack of empirical studies on interoperability within the NHS. This gap is particularly evident when compared to the United States, where FHIR has been more widely adopted and analysed (Ayaz *et al.,* 2021). Li *et al*. (2023) also observed that research on EHR interoperability has traditionally been US-centric from a provider perspective.

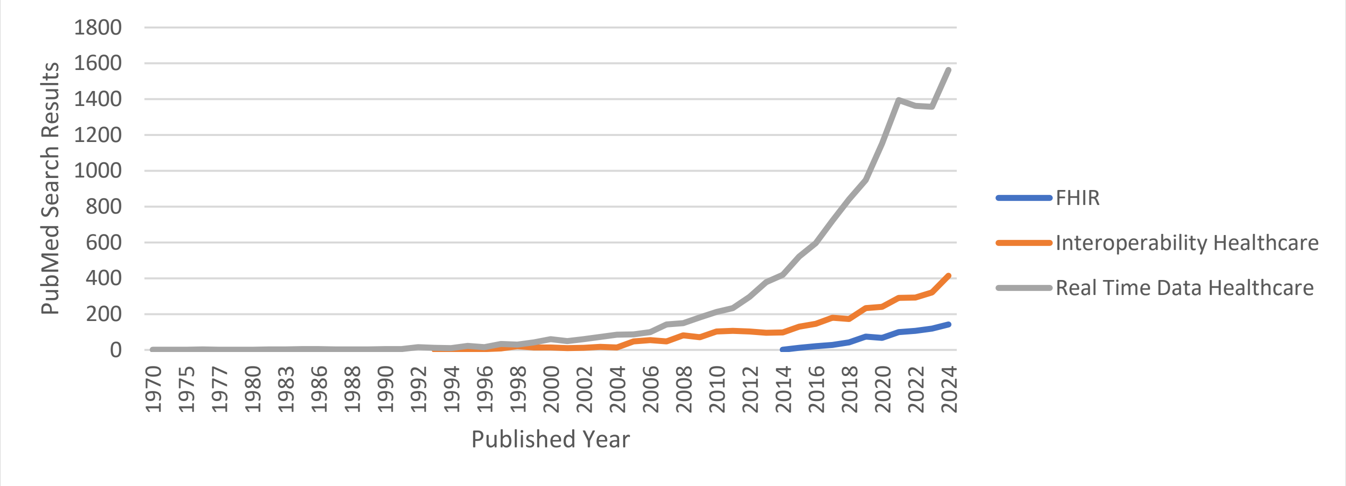


Figure 2. Graph highlighting the volume of results per search terminology within PubMed per article publish year

Most included studies were qualitative, and arguably, lacked quantitative data substantiating issues such as duplication of work, manual processes and poor data quality. Despite cited potential of healthcare interoperability (Braunstein, 2022), questions remain about the technical readiness of NHS systems and the timeline for widespread NHS interoperable data-sharing.

Three overarching themes and eight sub-themes were identified (Table 4, p.19), drawing parallels with the Technology, People, Organisational, and Macroenvironmental (TPOM) framework[15] (Cresswell *et al.,* 2020). The themes, which are generally complementary and intrinsically linked, offer insight into the practical challenges of implementing healthcare data flows, reflecting the contributors identified by Aerts and Bogdan-Martin (2021), presented in Table 1 (p.12).

---

[15] The TPOM framework was developed as a formative evaluation for HIT implementation and digital transformation efforts. (Cresswell *et al.,* 2020).

Table 4. Literature Review Themes and Article Frequency

	Technical Factors			People and Organisational Factors			Macroenvironmental Factors	
	System Architecture and Infrastructure	Data Standards and Protocols	Data Quality, Consistency and Presentation	Organisational Culture, Staff Engagement and Acceptance	Workforce Expertise, Training and Digital Literacy	Capacity and Financial Challenges	Governance, Privacy and Ethical Considerations	Nationally Aligned Policy and Directives
Theme Occurrence (Count of Literature Articles)	**24**	**12**	**10**	**14**	**12**	**11**	**17**	**12**

The following sections explore each theme, beginning with Technical Factors.

Technical Factors

The theme of technical factors comprises three subthemes: system architecture and infrastructure, data standards and protocols, and data quality and consistency. Symons *et al.* (2019) comment that the prospect of universally connecting multiple organisations, systems, and exchange protocols is daunting, with technical and socio-technical barriers presenting the most significant challenges.

System Architecture and Infrastructure

System architecture involves the structural design of IT systems (hardware, software, databases, and networks), while infrastructure includes the supporting physical and virtual components (Dovgyi and Kopiika, 2022). Healthcare IT environments are often complex, characterised by multiple legacy systems that are difficult to update or integrate with modern technologies (NAO, 2020; Warren *et al.,* 2019), and reflected within Morris *et al.'s* (2023) reflective account of integrating remote monitoring in Child and Adolescent Mental Health Services. Adaba and Kebebew (2018) similarly found legacy systems lacked functionality, leading to manual processes, staff frustration, data errors, and increased costs. Using two cycles of action research, guided by socio-technical systems theory, they found implementing real-time data recording in an NHS theatre department improved communication, coordination, and collaboration, reducing administrative workload.

Warren *et al.* (2019) using retrospective observational analysis of Hospital Episode Statistics and organisational data from NHS England acute trusts, found EHR distribution across more than twenty-one platforms, with some using multiple systems and in-house solutions. However, it is unclear what the interoperable capabilities of these systems are or whether other factors contributed to data fragmentation. The diversity of systems often necessitated manual workarounds[16], a consequence also seen in qualitative literature (Li *et al*., 2023;

---

[16] The term 'manual workaround' refers to alternative, ad hoc processes that are implemented to compensate for limitations or inefficiencies in existing systems. In this context, these workarounds typically involve human effort, such as manually inputting or transferring data.

McVey *et al.,* 2021; Sharma *et al.*, 2022). Patient transfers between trusts were common but having found minimal coordination between EHRs, Warren *et al.* concluded that fragmentated patient records impact the provision of quality, cost-effective, safe care. This aligns with other literature which suggests integrated care systems support seamless patient transfers and continuity of care (Litchfield *et al*., 2022; Wain, 2021). Zhang *et al.* (2023) also note the variation in systems presents technical data exchange barriers, particularly between primary and secondary care, having used retrospective longitudinal analysis over a seven-year period to evaluate data-sharing implementation across English Trusts. Using national aggregate secondary care datasets and self-reported Clinical Digital Maturity Index Assessments, their covariate analysis associated data-sharing with improved patient-reported experiences and reduced Accident and Emergency target breaches.

Arguably, a qualitative dimension could compliment Warren *et al*. (2019) and Zhang *et al.'s* (2023) findings by revealing perceptions of interoperability. For instance, Wain (2021) explored integrated health and care systems using an interpretivist design and qualitative interviews with community nurses (n=5) in NHS England. Using Colaizzi's descriptive phenomenology analysis, non-interoperable IT systems were found to hinder integrated care. Naqvi *et al*. (2019) reported similar challenges in integrating primary and social care, following semi-structured interviews with GPs (n=18) and Practice Managers (n=7) in London. Relatedly, Litchfield *et al.* (2022) through interviews with a snowball sample of primary care managers (n=13) and clinicians (n=15), found interoperability and data-sharing issues as key barriers to timely and accurate communication between primary and secondary care. Conversely, Sullivan *et al*. (2023) offer a more optimistic technical perspective, suggesting the national GP Connect Programme is an established mechanism for sharing GP data between organisations.

Li *et al*. (2023) identified system usability concerns, multiple logins, duplication of effort, and limited access to external records as persistent barriers to interoperability. Their semi-structured interviews (n=15) targeted the 'under-represented' (p.2) role of clinical information officers in NHS England through convenience sampling, with interview coding and thematic analysis conducted collaboratively by multiple researchers. Similarly, Sharma *et al.* (2022), through qualitative phone interviews (n=23) with transplant coordinators in all UK renal centres found that non-interoperable systems necessitated manual workarounds, wasting resources and delaying care. From staff interviews (n=54) and observations (102hrs) across five NHS hospitals McVey *et al.* (2021) also found significant effort was required to retrieve data from unintegrated systems.

Research supports the idea that improved system integration enhances clinical outcomes, for example Camacho *et al.* (2024), drawing on expert input and published data, modelled how interoperable records could reduce prescribing errors, costs, and patient harm. These findings align with other research suggesting that interoperable infrastructure helps reduce workflow inefficiencies caused by legacy EHR systems and poor usability (Banerjee *et al.,* 2018; Li *et al.,* 2023; Sheikh *et al.,* 2021). Herlitz *et al.* (2023) found data fragmentation and infrastructure limitations hindered technology adoption in their mixed-method evaluation of remote monitoring during COVID-19, using surveys (staff=292, patients and carers=1069) and interviews (staff=58, patients and carers=62). However, remote monitoring appeared well-received when supported by dependable, high-quality, and clinically useful systems. Outcomes from three Big Data workshops led by Banerjee *et al.* (2018) provide an evidence-based analysis of the NHS's potential to evolve into a learning health system (LHS), suggesting progress is constrained by inadequate infrastructure and inconsistent EHR availability. Their recommendation for secure, integrated systems and clear data standards (discussed in the next section), aligns with the WG's *Digital and Data Strategy* (2023), which highlights these as foundational to data-driven healthcare improvements.

Data Standards and Protocols

Data standards and protocols refer to the agreed rules and formats for structuring and exchanging healthcare data consistently across systems, and are considered fundamental to achieving interoperability, yet studies have found a lack of common standards in practice (Braunstein, 2022; Li *et al.,* 2023). Reflecting on their experiences of establishing the National COVID-19 Chest Imaging Database, the opinion piece by Cushnan *et al.* (2021) advocates for national datasets and adherence to data standards, suggesting that inconsistent regulations and the absence of unified datasets delay national data collection. The health policy paper by Sheikh *et al.* (2021) also notes that contracts often fail to specify interoperability standards, so even when technological infrastructure exists, systems struggle to integrate smoothly.

To promote common approaches for data exchange national initiatives have introduced the NHS Interoperability Framework, Interoperability Toolkit (ITK) (NHS Digital, 2023), and FHIR UK Core (NHS Digital, 2024). However, Trivedi *et al.'s* (2023) implementor report, one of the few studies to mention FHIR within the NHS context, notes slow adoption of FHIR standards by EHR vendors. Their proof-of-concept indicates that FHIR-enabled software could unify data across systems, supporting accessibility and decision-making. Relatedly, Smith *et al.* (2020), through the National Institute for Health Research Health Informatics Collaborative, integrated hepatitis data across five centres. Despite a structured governance framework and clinician-led dataset refinement, data inconsistencies across sites, including field definitions and formats limited comparability. This practice-based literature suggests that even well-governed integration efforts can be hampered by underlying data inconsistencies.

Standardisation may also aid clinical decision-making, for instance, Xiao *et al.* (2022) developed a protype Clinical Decision Support System (CDSS), conducted a comparative case study evaluation, and concluded that encoding clinical arguments using standard formats supports interoperability. Relatedly, Kapadi *et al.* (2024), through thematic analysis of semi-structured interviews with a purposive sample of professional stakeholders (n=23) from five UK cancer centres, found the need for data standardisation, investment, and educational support. The National Institute for Health and Care Excellence (NICE) also supports that structured, coded guidelines could have 'significant methodological implications' (p.3), improving the clarity and consistency of clinical recommendations. These findings are drawn from NICE's workshops with CDSS developers and vendors, as well and the understanding that structured data and knowledge enable a continually LHS (Mitchell, 2020).

Nevertheless, practical issues exist, as Mitchell (2020) describes the difficulties agreeing on coding standards and responsibilities for their application. More promisingly, Sullivan *et al.* (2023) suggest SNOMED CT has unified coding across hospitals and practices supporting seamless data exchange and continuity of care, though they caution that standards alone are insufficient because the meaning behind the data must be universally consistent and understandable. Thimbleby's (2022) critique echoes this, contending that digital healthcare regulation requires urgent attention because current healthcare standards are inconsistent and lack the rigour needed, concluding interoperability problems may originate from diverse and incorrect interpretations of 'vague standards' (p.1). The next section discusses data quality challenges.

Data Quality, Consistency and Presentation

High-quality data is argued as essential for ensuring accuracy, reliability, and consistency across datasets, supporting seamless system integration (Dawoodbhoy *et al*., 2021; Camacho *et al.,* 2024; Xiao *et al.,* 2022). Data quality issues have been found to stem from inconsistent clinical practices and non-interoperable EPRs leading to missing data, with variable formats and free-text fields complicating analysis and overall robustness (Smith *et al.,* 2020; Li *et al.,* 2023; Kapadi *et al.,* 2024). McVey *et al*. (2021) found that staff often distrust the quality of data within siloed EHRs, relying on paper forms to validate the data, but supported partial automation of the most time-consuming aspects. Relatedly, Dawoodbhoy *et al.* (2021), combining a narrative literature review with twenty semi-structured interviews across NHS mental health units, examined the potential for AI to improve patient flow and resource allocation. While CDSS' were seen to offer value in streamlining administrative tasks and identifying emerging patterns, concerns were raised about the current quality of NHS data limiting AI effectiveness.

Several studies suggest that data volume and complexity can be overwhelming, and that data should be structured and manageable to support accuracy (Beasant *et al.,* 2023; Chu *et al.,* 2022). When implementing digital technologies for early detection of dementia-causing diseases, Wilson *et al.* (2023) found, through semi-structured interviews with primary and secondary care professionals (n=18), that easy-to-interpret data enhanced trust and understanding. Similarly, Jager *et al.* (2023) in a realist evaluation, interviewed a purposeful sample of primary care commissioners (n=23), analysed public care commissioning meetings and found that visually appealing data from credible sources improved engagement.

The broader issue of data quality and consistency may be tied to organisational factors that shape how data is collected, interpreted, and acted upon. Khanbhai *et al.* (2021), using purposeful sampling and semi-structured interviews with fifteen participants (9 lead nurses and junior clinical staff, 6 non-clinical patient experience team members), reported that data inconsistencies and large amounts of free-text hindered analysis and reduced the perceived value of real-time data. Participants reported that data presentation was often difficult to interpret, limiting data-driven action. The authors conclude that organisational and micro-level factors should be tackled together prioritising data utilisation, making visualisations accessible, straightforward, and engaging. Likewise, Banerjee *et al.* (2018) suggest that while infrastructure and analytical tools are essential, they are often misunderstood and underdeveloped, requiring investment and clinical engagement for high-quality data input. These organisational factors will be discussed next.

Organisational Factors

Organisational subthemes relate to culture, digital literacy, training and workforce capacity. Investment in integration and data processes likely depends on these elements, underpinned by staff engagement, and leadership (Dawoodbhoy *et al.,* 2021; Kapadi *et al.,* 2024; Li *et al.,* 2023; Pope *et al.,* 2024*).*

Organisational Culture, Staff Engagement and Acceptance

Organisational culture, defined as the shared values, beliefs, and practices shaping workplace behaviour, is widely recognised as a facilitator of digital transformation (Chalmers and Brannan, 2023). Camacho *et al.* (2024) emphasise the necessity of staff buy-in, as systems may underperform if the workforce does not value them. Staff acceptance is considered crucial for embedding interoperable health systems within the organisational culture, as posited by Catlow *et al* (2024), who in a large NHS endoscopy randomized controlled trial (RCT), found that active engagement with real-time feedback systems improved clinical outcomes. Behavioural resistance documented within Wood and Proudlove’s (2022) case study of an NHS IVF clinic, linked data delays to perceptions of it being a low priority. Using process waste mapping to identify non-value-adding activities, Plan–Do–Study–Act (PDSA) cycles, and regular staff engagement sessions, they reported improved workflow, collaboration, and data quality.

Behavioural change frameworks can support cultural alignment during digital transitions, for example, the COM-B model (Michie *et al.,* 2011) proposes that capability, opportunity, and motivation are prerequisites for new behaviours, while McKinsey and Co's Influence Model (Basford and Schaninger, 2016) emphasises leadership's role in modelling behaviours, incentivising change and communication—the 'golden thread' of change management (Dwyer *et al.* 2019). However, Naqvi *et al.* (2019), identified poor communication, hostile culture, and weak cross-sector relationships reduce accountability and collaboration. Their qualitative interview analysis details inefficiencies in meetings, and overworked staff as barriers to integration. Sharma *et al*. (2022) also found variation in data management and digital maturity across UK transplant centres, many of which lacked departmental digital leadership.

Khanbhai *et al.* (2021) highlight risks of disengagement, finding that staff at a London NHS Trust undervalued and underused patient feedback data due to its format and presentation, leading to missed improvement opportunities. Similarly, McGowan *et al.* (2024), through qualitative interviews with twenty NHS England commissioners, found scepticism toward remote monitoring technologies, largely due to limited evidence of benefit, and perceptions of imposed implementation, rather than clinically led. One participant noted success was more likely when linked to a clearly defined problem, although organisational views on what constitutes a genuine need varied.

Structured tools are thought to support staff engagement, such as the customised NoMAD tool[17] (Finch *et al.,* 2018) and Lewin's Change model (1958). In a Birmingham case study of EHR change during the COVID-19 pandemic, Pankhurst *et al.* (2021), found that staff engagement can help maintain digital system buy-in. Co-design has also been proposed as strategy to foster engagement and reduce resistance (Bird *et al.,* 2021). Pope *et al*. (2024) agrees that clinical engagement is fundamental to healthcare innovation but notes the need for protected time and in-house expertise, an issue explored in the next section.

Workforce Expertise, Training and Digital Literacy

Successful implementation of interoperable health systems is suggested to rely on a workforce with the skills to use them effectively, with several studies emphasising that staff understanding of interoperability, alongside targeted training and improved digital health literacy, could strengthen integration efforts, increase confidence, and overall benefit (Li *et*

---

[17] The NoMAD Tool, based on Normalisation Process Theory, provides a framework for evaluating system acceptance across four domains: coherence (understanding the intervention), cognitive participation (engagement), collective action (practical implementation), and reflexive monitoring (ongoing assessment and adaptation) (Finch *et al.,* 2018)

al., 2023; Sullivan *et al.,* 2023; Herlitz *et al*., 2023). Adaba and Kebebew (2018) also suggest that training can create a culture which values data.

Alhmoud *et al.'s* (2022) mixed-methods evaluation on real-time auditing at Barts Health NHS Trust identified a lack of staff engagement due to system usability challenges and the need for clinicians to learn digital tools. Interviews with three nurses, two PDSA cycles and retrospective descriptive EHR data analysis revealed advanced data-driven methods improved the management of critically ill patients, but participants required clear guidance, and concise data displays. These findings reflect broader issues around data interpretation and usability, suggesting that ongoing training is needed to support effective data use (Davidson *et al*., 2022; Jager *et al*., 2023; Khanbhai *et al.,* 2021; Wain, 2021). The commentary article by Goldacre *et al.* (2020), argues that little formal analytics training has left clinicians ill-equipped to interrogate data, resulting in poor decision making. However, training initiatives such as the NHS Digital Academy's (2023) digital leadership programme, and NHS England's Making Data Count training (Riley *et al.,* 2021) aim to promote data-informed decisions. Chen and Banerjee (2020) also advocate for the rapid upskilling of workforce digital capabilities, noting that knowledge of health informatics, data science, and digital technology could future-proof the medical profession.

Authors also highlight gaps in technical expertise. For example, Pope *et al.* (2024), reflecting on learnings from a partnership formed in 2021 between Great Ormond Street Hospital and a pharmaceutical company, suggest these roles are an important, often-overlooked dimension, with NHS ICT teams lacking the cloud skills and resources needed to support modern infrastructure. Thimbleby's (2022) opinion piece raises related concerns about the lack of regulatory oversight and accountability in healthcare software development, suggesting developers should be held to the same standards and accountability as other professions. The Federation for Informatics Professionals (2022) promotes the professionalisation of health informatics, having established competency frameworks. Furthermore, Camacho *et al*. (2024) and HIMSS (2024b) emphasise the value of hybrid clinical informatics roles, which bridge clinical and technical domains to advance healthcare technologies. Arguably, this requires the organisational capacity and financial resources to hire and train staff, as discussed next.

Capacity and Financial Challenges

Financial investment can support the workforce, digital infrastructure, and governance required for effective digital transformation, yet many organisations face budgetary constraints (Camacho *et al.,* 2024; Li *et al.,* 2023; Naqvi *et al*., 2019). Zhang *et al.* (2023) suggest that interoperability benefits depend on foundational investment in workforce resource and capacity. Pope *et al.* (2024) raise concerns regarding the cost of digital implementation, especially as NHS trust budgets operate under capital expenditure models. However, Rastogi's (2022) discussion article suggests that SCRs between primary and secondary care lead to new value-based care models, generating cost-savings which Camacho *et al.* (2024) suggest can help offset implementation costs. Nevertheless, Cushnan *et al.* (2021) stress that sustained funding is required for maintaining data infrastructure and supporting long-term interoperability goals. Naqvi *et al.'s* (2019) qualitative study concluded that limited funding, lack of pooled budgets and misaligned incentives between health and care sectors exacerbates barriers to interoperability and integrated patient care.

Pankhurst *et al.'s* (2021) reflection on the rapid system development and EHR-lab integration for real-time COVID-19 case identification in Birmingham, highlighted risks of data burden, disengagement, and alert fatigue. Similarly, through deductive and inductive thematic analysis of six semi-structured interviews with clinical ICU ward staff in two UK hospitals, Davidson *et al.* (2022) found concerns about digital literacy and poorly timed alerts causing stress and demoralisation. However, participants valued real-time, flexible, and accessible dashboards for reducing clinical errors, easing cognitive load, and improving continuity of care during patient handovers.

Researchers indicate that staff capacity is strained by manual workarounds and multi-system duplication (Cushnan *et al.,* 2021; Khanbhai *et al.,* 2021; Li *et al.,* 2023.; Pope *et al.,* 2024; Sharma *et al.,* 2022; Wood and Proudlove, 2022). In the Hull and East Yorkshire region, Crowther *et al.* (2022) held exploratory meetings (staff=7, patients=3), a staff feedback event and a codesign event which revealed preference for real-time data for staff to compare, support and learn from. However, concerns from healthcare professionals were that data transparency could highlight poor practices, demotivating individuals. Nurses were also concerned about the repetitive nature of data recording, increasing staff fatigue and reducing data quality. Similarly, Wilson *et al.* (2023) highlighted staff concerns that early detection systems could increase workload, and overburden the NHS, requiring dedicated resources, implementation plans, and protection from unintended consequences. These financial and workforce challenges are potentially shaped by the broader healthcare environment which is discussed next.

Macroenvironmental Factors

Macroenvironmental subthemes include national policy, governmental directives, and data ethics. Effective data systems are said to require coordinated efforts across these domains to support sustainable digital transformation and interoperability (Schmitt *et al.* 2023). National policy alignment and clear ethical guidance are seen as both barriers and potential enablers for digital improvements (Aerts and Bodgan-Martin, 2021).

Governance, Privacy and Ethical Factors

Ethical and governance concerns are considered critical in efforts to enhance interoperability and real-time data use, particularly given historical controversies surrounding data privacy and the sensitivity of personal data (Elizondo, 2024). Elizondo's five-year multi-method longitudinal study, examined the governance complexities of implementing SCRs within the Southwest region of England, using public documents, governance meeting observations (12hrs), and interviews (n=50) with selected NHS and social care staff involved in developing the interoperability initiative. The research revealed difficulties in aligning stakeholders, establishing data-sharing agreements, and ensuring GDPR compliance. These challenges were echoed by Litchfield *et al.* (2022), who found that GDPR added significant complexity to data use and integrated care efforts, with one participant describing it as a 'massive barrier' (p.782). This tension between ensuring robust data security and achieving effective interoperability for SCRs was further highlighted by Schmitt *et al.* (2023) and Sharma *et al.* (2022).

Additionally, Elizondo (2024) emphasises that the absence of clear governance frameworks can cause implementation delays, fostering distrust and confusion within healthcare organisations. The research draws attention to the importance of continuous dialogue, workshops, and narrative framing to facilitate social learning and idea diffusion. The findings indicate that opportunities for open discussion helped address concerns, aligned diverse perspectives and reportedly strengthened ethical norms, including transparency in decision-making. Similarly, Cushnan *et al.* (2021), based on their experience with the NCCID, recommend standardising processes, reducing participation barriers, and leveraging secure Trusted Research Environments. They suggest that data governance and standardisation contributes to ethical data use, data security, and transparent public engagement.

Further challenges in ensuring that interoperable systems and real-time EHR data are used in fair, secure, and ethical ways, were identified by Alhmoud *et al.* (2022), who noted conflicts between public interest, patient safety, and optimising HIT as barriers to realising the full potential of real-time data. Beasant *et al.* (2023) warn that HIT could widen health inequalities, particularly lower-income individuals with limited digital access. Relatedly, Wilson *et al.* (2023) thematically identified staff concerns about digital exclusion and health disparities in the early detection of dementia, and Herlitz *et al.'s* (2023) study found remote monitoring favoured younger, higher-educated patients. These findings suggest that ethical frameworks should balance innovation with patient protection, ensuring equitable access to digital health tools.

Dawoodbhoy *et al.* (2021) suggest AI presents further ethical concerns, having identified issues related to confidentiality, regulatory gaps, patient autonomy, and transparency in AI models from interviews with mental health professionals and AI experts. Participants stressed the need for clearer governance and oversight to mitigate potential misuse and ethical concerns. The ethical implications of data transfer and reuse are also explored by Golinelli *et al.* (2018) whose research and interpretation of European legal framework advocates that the NHS should adopt more appropriate, innovative data policies to better support Big Data opportunities while prioritising patient interests.

Vezyridis and Timmons (2019), using Actor-Network Theory and semi-structured interviews (11 primary EHR data researchers, 7 GPs, and 9 citizens), identified participants unease with the 'care.data' programme [18] particularly due to concerns over insufficient consent mechanisms, heavy-handedness and limited patient choice. Despite past scepticism, public attitudes to data sharing appear to be shifting, as Jones *et al.* (2022) found growing support for using NHS data in research and public health efforts to improve healthcare quality following the COVID-19 pandemic. Their co-designed anonymous online survey received 29,275 responses, and statistical analysis of Likert-scale data revealed concerns over data privacy and transparency remained, especially regarding how data would be used and by whom. Willingness to share data was closely tied to trust in how healthcare organisations would handle the data. Zhang *et al.* (2023a) recommend refining opt-out mechanisms to give patients greater control, reflecting Vezyridis and Timmons (2019) suggestion for collaborative governance frameworks where experts provide informed opinions of potential data users and usage, but individuals retain final control. The final subtheme to be discussed relates to national policy alignment.

Nationally Aligned Policy and Directives

Arguably, effective healthcare data interoperability cannot be achieved in isolation; it requires a unified approach, supported by national and organisational policies. Elizondo (2024) concluded that insufficient policy alignment, alongside disparities in digital maturity and data quality across regional bodies, complicates efforts to standardise healthcare data management, with overlapping administrative agendas exacerbating the issue. Both Morris *et al.* (2023) and Sullivan *et al.* (2023) suggest clear national guidelines for development and sufficient investment in cloud infrastructure are essential for achieving interoperability.

The absence of aligned policy frameworks within the UK is thought to have undermined the potential for widespread interoperability, with the decentralised nature of system procurement resulting in a 'patchwork landscape' (Zhang *et al.,* 2023a, p.1). Using public resources and Freedom of Information Requests, Zhang *et al.* (2023a) systematically mapped electronic

---

[18] The care.data programme was a UK government initiative announced in 2013, intended to bring health and social care data from different settings together, but was paused following significant resistance from both the public and clinicians whose concerns were that its use was unclear, and therefore resulted in a large number of opt-outs. (Vezyridis and Timmons, 2019)

data flows across England and identified non-transparent multi-stage processes. The researchers found that many processes failed to adhere to best practices for secure access and produced duplicate data assets that diminished overall value. Similarly, Elizondo (2024) calls for future policies to address the current inconsistent, fragile, ad-hoc integration methods.

Regulatory fragmentation potentially compounds interoperability challenges, as Wain, (2021) found the existence of different regulators and governance frameworks between health and social care sectors increases the potential for system errors. Cushnan *et al.* (2021) warn that without nationally aligned regulations, the disjointed data landscape will persist, a view echoed by Li *et al.* (2023) and Sheikh *et al.* (2021). Kapadi *et al.* (2024) suggest method clarity, educational support, and national guidance could facilitate implementation. The lack of centralised IT strategies or standardised procurement practices may also hinder interoperability in decentralised healthcare environments (Sharma *et al.,* 2022; Litchfield *et al.,* 2022). Zhang *et al.* (2020) suggest that the responsibility for setting national digital policy and standards lies with NHSX, a body established to lead the digital transformation of NHS technology and data (Department of Health and Social Care, 2019). However, Sheikh *et al.'s* (2021) health policy paper recommends that both government-led and bottom-up innovation are needed to address front-line challenges. Incentivising organisations to integrate and provide high-quality, structured data may also support digital transformation goals (Schmitt *et al.,* 2023; Sheikh *et al.,* 2021).

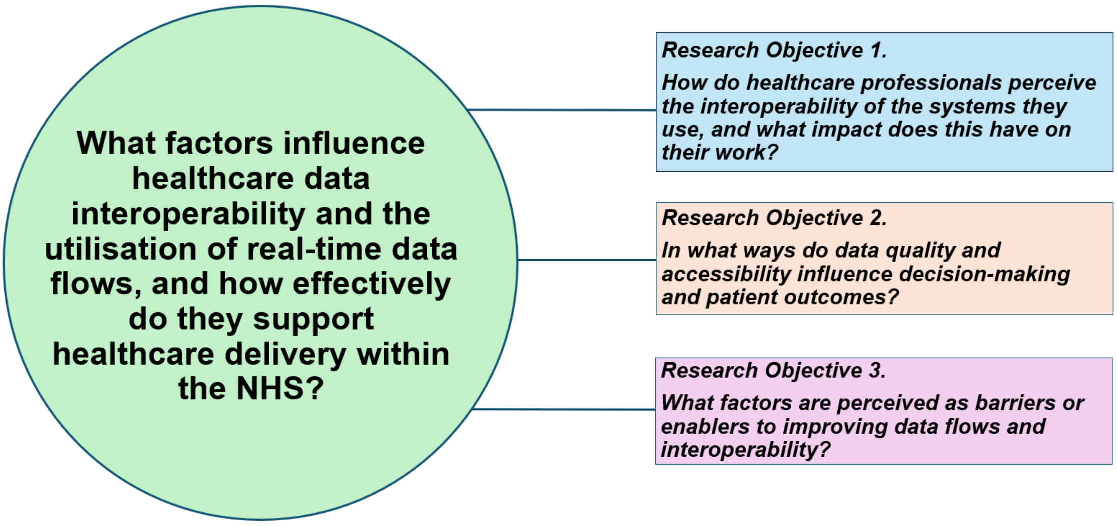
Summary

The literature suggests that while digital transformation and data interoperability present potential opportunities for improving healthcare delivery and patient outcomes, realising these benefits within the NHS is likely to require sustained effort. Progress appears to be impeded by inter-related technical, organisational, and macroenvironmental factors. Advancement likely depends on the implementation of interoperable systems, but also on workforce engagement, and a collaborative culture that values data quality. Developing digital literacy through ongoing training, infrastructure investment, and cross-sector collaboration surface as potential enablers to overcoming siloed working and improving data practices. Building trust, both in digital technology and amongst healthcare professionals and patients may be another enabler, supported by national policies which align with organisational priorities and reinforce common data standards. Nevertheless, further research is needed to identify effective strategies and best practices for enabling ethical and secure data sharing within the NHS. However, the limited volume of UK-specific literature on interoperability may constrain a full understanding of interoperability barriers, indicating a need for further research.

Chapter 3: Methodology and Methods

Introduction

The previous chapter reviewed existing literature, exploring challenges surrounding data flows and interoperability within healthcare. This chapter outlines the methodological approach and methods used to investigate the research aim and objectives introduced in Chapter 1 and depicted in Figure 3.



What factors influence healthcare data interoperability and the utilisation of real-time data flows, and how effectively do they support healthcare delivery within the NHS?

Research Objective 1.
How do healthcare professionals perceive the interoperability of the systems they use, and what impact does this have on their work?

Research Objective 2.
In what ways do data quality and accessibility influence decision-making and patient outcomes?

Research Objective 3.
What factors are perceived as barriers or enablers to improving data flows and interoperability?

Figure 3. Research Aim and Objectives

The philosophical approach and research design are discussed first, providing rationale for selecting a mixed methods approach within the pragmatist paradigm. The concurrent data collection methods are then described, including procedures for data collection, handling and analysis. The chapter concludes by discussing research rigour.

Research Approach

Philosophical paradigms provide distinct approaches for conducting and interpreting research (Clark *et al.,* 2021). Examples include positivism, which focuses on structured, measurable data; interpretivism, which focuses on subjective experiences; and realism which emphasises the existence of an independent reality beyond perception (Clark *et al.,* 2021). The strengths and limitations of each paradigm are often seen as mutually restrictive, influenced by their underlying assumptions, values, and methods. For example, positivism may overlook the nuances of subjective experience, while interpretivism can struggle with generalisability. Pragmatism offers a flexible alternative, preferring actionable insights and problem-solving over rigid adherence to any single philosophical framework (Dudovskiy, 2020; Robson, 2024). This middle ground rejects the dualisms of other paradigms, enabling researchers to adopt the methods best suited to addressing the research question or problem (Johnson and Onwuegbuzie, 2004).

Qualitative and quantitative research represent distinct methodological traditions with differing epistemological orientations. Qualitative methods are typically inductive and theory-generating, resonating with interpretivism, while quantitative methods are deductive and theory-testing, consistent with positivism (Clark *et al.,* 2021). There is support for integrating these approaches in 'mixed methods' designs, which collect and analyse both data types (Creswell and Creswell, 2023). Bowling (2023) organises mixed methods into four types; triangulated, embedded, explanatory and exploratory. Triangulation, although not exclusive to mixed methods, can enhance confidence in results, balance research biases and improve the depth and validity of findings by cross-referencing multiple data sources (Clark *et al.,* 2021; Noble and Heale, 2019). Mixed method triangulation has been referred to as 'the ultimate triangulation' (Friedman *et al*., 2022, p.404), and involves obtaining 'different but complementary data on the same topic' (Bowling, 2023, p.458). This may be 'as simple as comparing and contrasting qualitative and quantitative responses in a questionnaire' (Scott, 2016, p.105).

Robson (2024) notes that although data triangulation strengthens research rigour, different methods may complicate direct comparisons. Mixed methods research also faces challenges related to the epistemological tensions[19] over whether opposing paradigms can be meaningfully integrated (Dawadi, Shrestha and Giri, 2021). Mixed method studies tend to be more time consuming and complicated than mono-method[20] studies (Scott, 2016). Despite this, mixed methods research is believed to leverage the strengths of each method while minimising their individual limitations, offering a more comprehensive understanding of complex issues (Doorenbos, 2014; Johnson and Onwuegbuzie, 2004; Scott, 2016). This is particularly valuable in healthcare and informatics, where technical, clinical, and human factors intersect (Scott, 2016). The suitability of mixed methods is therefore considered in relation to the research context[21], with the *contingency theory* proposed by Johnson and Onwuegbuzie (2004) acknowledging that quantitative, qualitative, and mixed research 'are all superior under different circumstances' (p.22).

Research Design

The study adopts the pragmatist paradigm, as both qualitative and quantitative data help

---

[19] Debates and conflict in the 1980s between researchers termed the 'paradigm wars' centred on the perceived sense of incompatibility between interpretive and causal paradigms. (Clark *et al.,* 2021)

[20] There is a distinction between mono-method studies, which use a single research method, and multi-method studies where multiple types of qualitative or quantitative data are collected. In contrast, mixed methods studies follow a distinct design that explicitly integrates both qualitative and quantitative approaches from the study's conception throughout its execution. (Creswell and Creswell, 2023; Scott, 2016)

[21] For example, randomized controlled trials (RCTs) are often considered the gold-standard for studying causal relationships, especially regarding on efficacy and effectiveness (Hariton and Locascio, 2018). Yet Scott (2016) comments not all research questions can be answered solely by a RCT design.

address the research aim and objectives. Pragmatism is considered instrumental when investigating organisational processes, demonstrating real-world applicability, and producing results that can be used to evaluate and inform healthcare practices and future development (Kelly and Cordeiro, 2020). Pragmatism is often the philosophical underpinning of mixed methods research (Biesta, 2010); together they form the basis of this study's design.

This study’s mixed method approach[22] facilitates the exploration of real-world perceptions, incorporating both quantitative and qualitative elements within an online questionnaire, alongside quantitative documentary evidence from existing datasets. Figure 4 illustrates how this combination explores both measurable outcomes and contextual, human factors, offering a richer understanding than either method alone (Scott, 2016). Pragmatically, a cross-sectional survey was deemed the most appropriate design, considering the constraints of time, word limit, and the lone researcher’s limited research experience. The applied multi-strategy design, likened to evaluation research (Robson, 2024), helps address literature gaps, such as the lack of quantitative evidence on data quality issues, and the inclusion of both technical and clinical perspectives. Design was also influenced by the informatics study types and appropriate methodological approaches[23] suggested by Friedman *et al.* (2022) illustrated in Appendix 3.

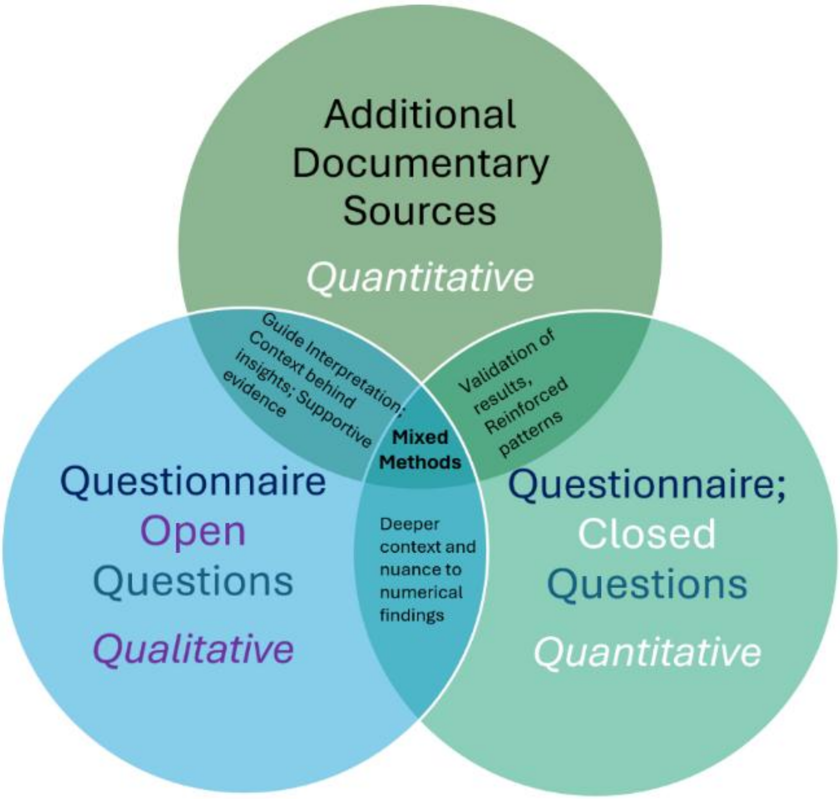


Figure 4. Research Design, Data Collection Methods and Their interactions

---

[22] Mixed methods research can incorporate various designs, including case studies, experimental (e.g. RCTs), cross-sectional (e.g. surveys and observations), longitudinal, and comparative designs (Clark *et al.*, 2021).

[23] For example, ‘identifying the resource need’ aligns with qualitative methods, ‘resource performance’ typically involves quantitative approaches, and ‘evaluating resource effectiveness’ requires both. (Friedman *et al.* 2022)

Mixed methods studies are typically classified by two factors: priority (the weighting of each element) and sequence (the order of data collection) (*Clark et al.,* 2021). Explanatory and exploratory sequential designs allow findings from one phase to inform the next, while convergent designs, also known as concurrent triangulation, collect data simultaneously offsetting the weaknesses of a single approach (Creswell and Creswell, 2023). Using Morse's (1991) notation system[24], this study is represented as **QUAL + QUAN,** signifying equal weighting of qualitative and quantitative data collected and analysed concurrently in a convergent design (Clark *et al.,* 2021). This approach supports mixed methods integration principles, aiming to create a rich combination of mutually dependent and complementary findings (Halcomb, Massey and Gunowa, 2023; Scott, 2016). The next section discusses the study's approach to access and sampling.

Access and Sampling

Access to survey participants and documentary data was facilitated through the author's professional role within PTHB. Prior to data collection, ethical approval was obtained from both the University of Wales Trinity Saint David (UWTSD) (Appendix 6), and the Research and Development Hub at PTHB (Appendix 7)[25]. Beauchamp and Childress (2013) outline four foundational principles of biomedical ethics: beneficence (acting in participants' best interest), non-maleficence (avoiding harm), autonomy (respecting individual rights), and justice (ensuring fairness and equality). Similarly, Atenas *et al.* (2023) emphasise informed consent and ethical research practices that consider individual interests, justice, and collective well-being. Recognising that ethics extend beyond legal compliance (Creswell, 2013), this study was grounded in these principles to maintain participant trust and adhere to regulatory and ethical standards.

Sampling methods are broadly categorised into probability (random, systematic) and non-probability (convenience, purposive) methods (Martínez-Mesa *et al.* 2016). Probability sampling is typically associated with quantitative research to enhance generalisability, while non-probability sampling is commonly used in qualitative research to gain in-depth insights from a specific group (Patton, 2015). Fair and equitable participant selection upheld the principle of justice (Beauchamp and Childress, 2013). Given the mixed methods approach, seeking both breadth and depth, professionals working within NHS Wales and Welsh Local Authorities were purposively sampled. As Johnson *et al.* (2020) note, qualitative research samples should be 'defined purposively to include the most appropriate participants in the

---

[24] Notation provides a symbolic representation of research procedures, clarifying the weight of each component through capitalisation, where a plus sign (+) denotes concurrent data collection and an arrow (→) represents sequential phases (Creswell and Creswell, 2023).

[25] Since this study is classified as a service evaluation rather than health and care research, and involves minimal risk, NHS Research Ethics Committee approval and a Data Protection Impact Assessment (DPIA) were not required.

most appropriate context for answering the research question' (p.141). Target participants (n=280) were individuals with prior engagement with the PTHB Data Engineering and Analysis (DEA) Team, identified via the 'requested by' field in the team's Azure DevOps (ADO) System[26]. This non-random selection aimed to deepen the understanding of the research area by involving individuals with potential interest and familiarity with healthcare data and interoperability (Gill, 2020), thereby increasing questionnaire relevance.

Participants, contacted verbally, over MS Teams and by email, were given an information and consent statement outlining the study objectives, data use, and withdrawal rights (Appendix 8). This supports Beauchamp and Childress' (2013) principle of autonomy by respecting participants' right to make informed decisions regarding their involvement, addressing Okorie *et al.'s* (2024) view that informed consent and privacy are foundational to ethical data handling. Table 5 shows the target sample breakdown by organisation and role. Most participants were from PTHB (76%), including clinical (n=103), administrative (n=85) and digital/technical (n=49) roles. The sample also included 72 individuals from other health and care organisations, and individuals in clinical informatics or senior management roles. The diversity of roles supports a rounded approach to addressing the research aim. Participants were invited to share the questionnaire with colleagues. This technique of snowballing sampling, a form of non-probability sampling, is thought to be useful when there is difficulty identifying members of the population (Martínez-Mesa *et al.,* 2016; Robson, 2024). This was intended to help reach staff who had an interest in healthcare data from organisations who had not worked directly with the PTHB DEA Team.

Table 5. Target Sample Organisation and Role Type

Organisation	Administrative	Clinical	Clinical Informatics/ Digital	Digital/ Technical	Senior Manager (Non-Patient Facing)	Senior Manager (Clinical)	Grand Total
Aneurin Bevan University Health Board			2	1			**3**
Betsi Cadwaladr University Health Board	1	1	2	3			**7**
Cardiff and Vale University Health Board		2	2	3			**7**
Cwm Taf Morgannwg University Health Board		1	2	1	1	2	**7**
Digital Health and Care Wales	1			7			**8**
Health and Care Research Wales	2				1		**3**
Health Education and Improvement Wales		2					**2**
Hywel Dda University Health Board	1		2	1			**4**
NHS Executive	5	1		6			**12**
Powys County Council				10			**10**
Powys Teaching Health Board	85	103	3	15	4	13	**223**
Swansea Bay University Health Board	1		2				**3**
Velindre NHS Trust			3	1			**4**
Welsh Ambulance Services NHS Trust			1	1			**2**
Grand Total	**96**	**110**	**19**	**49**	**6**	**15**	**295**

---

[26] Azure DevOps is a Microsoft Product used within PTHB to support work planning and record tasks requested and carried out by the Data Engineering and Analytics Team.

Targeted sampling strategies are also thought to significantly improve response rates[27] (Wu *et al*., 2022), helping to mitigate low engagement typically associated with online surveys (Clark *et al*., 2021). Martínez-Mesa *et al.* (2014) highlight that appropriate sample size calculation is important for reducing error, bias and enhancing the reliability of research findings. However, acceptable response rates can vary depending on the sample and collection methods employed, with lower response rates considered less of an issue for non-probability sampling methods (Clark *et al.,* 2021). Although not typically suggested for non-probabilistic samples, Krejcie and Morgan's (1970) method provides a useful benchmark for evaluating representativeness (Memon *et al.,* 2020) and considers 370 responses representative of NHS Wales staff (StatsWales, 2024; Appendix 4). Wu *et al.* (2022) suggest there is little agreement on adequate online survey response rates, but found published research had an average response rate of 44.1%. This study achieved a response rate of 31.9%. The methods for data collection are discussed in next section.

Data Collection Methods

Data collection for both qualitative and quantitative components of the mixed methods design, was guided by the pragmatic principles outlined by Kelly and Cordeiro (2020). The first principle**,** 'emphasis of actionable knowledge' (p.3) was addressed by collecting data that reflected real-world situations, ensuring the findings were both theoretically relevant and practically applicable. As Yin (2014) suggests, understanding how and why certain processes and practices occur can offer insight supporting future practical implementation, such as identifying specific conditions, resources, or constraints that require attention. Kelly and Cordeiro's (2020) second principle, 'recognition of the interconnectedness between experience, knowing and acting' (p.3) guided the identification of patterns in participants' responses, offering a richer understanding of the underlying factors shaping the topic. The following sections detail the data collection methods used in this study, starting with the online questionnaire.

Online Questionnaire

Survey research, defined as 'the collection of information from a sample of individuals through their responses to questions' (Check and Schutt, 2012, p.160), is often used in social research to describe and explore human behaviour (Ponto, 2015). Considering the strengths and weaknesses of various survey administration modes, summarised by Clark *et al.* (2021) (Appendix 5), an online survey was chosen for its practicality in reaching a large, diverse sample, enabling the exploration of complex contextual factors alongside outcome measurement. Curtis and Drennan (2013) describe how contemporary survey approaches

---

[27] The calculation for response rate is the number of usable questionnaires divided by the total sample of unsuitable or uncountable members (Clark *et al.,* 2021).

can incorporate ‘single, or multiple quantitative or qualitative methods, or a rich combination of both’ (p.175). Wasti *et al.* (2022) suggest that integrating qualitative methods into a survey alongside structured questions can help capture nuanced perceptions. Morgan (2017, 2014) agrees that the inclusion of both open and closed questions within a questionnaire, contributes to quantitative and qualitative data, thus making it a mixed methods approach.

An online, cost-free researcher-designed questionnaire, (Appendix 8), was created to gather perspectives on healthcare system interoperability, data quality, and accessibility. This method was time-efficient for both participants and the researcher. Delivered via MS Forms, a platform both familiar and accessible to participants, it allowed for responses at their convenience. Once designed and piloted with six digital and clinical PTHB colleagues, questions were refined to improve clarity, ensuring they were easily interpretable for the target sample, and required no further input from the researcher.

Following Dillman *et al.’s* (2014) survey design recommendations, the layout and wording were tailored for clarity and understanding, for example key terms such as interoperability were defined for participant comprehension (Clark *et al.* 2021). Questions were aligned with the research aim and objectives, whilst reflective of themes from the literature review. The researcher’s professional background, and the System Usability Scale (Brooke, 1995), influenced the question choice. The survey incorporated both open (n=9) and closed (n=18) questions, which included multiple select, single picklist, and Likert scale items, with responses ranging from 1 (strongly agree) to 5 (strongly disagree). This design enabled responses to be quantified and comparable between respondents (Clark *et al.* 2021). To minimise potential negative bias, where individuals with negative experiences may be more inclined to provide feedback, the questionnaire included balanced questions addressing both positive and negative aspects of the topic (Groves and Peytcheva, 2008).

Special care was taken to protect participant anonymity and confidentiality, aligning with ethical guidelines and research integrity standards (Hammer, 2017). The first four questions gathered non-identifiable demographic details about organisation, department, role, and data usage experience to understand what works for different groups in various contexts, notably the context–mechanism–outcome (CMO) configuration (Wong *et al.* 2016).

Documentary sources

The documentary data[28] collection, comprised three distinct elements, offering objective quantitative evidence relating to healthcare data quality and interoperability. Scott (1990) distinguishes documentary sources as either personal or official documents, and evaluates their quality based on authenticity, credibility, representativeness, and meaning. While official documents from private sources may be difficult to obtain, they are valued for their authenticity and relevance (Clark *et al.* 2021).

Quantitative data within PTHB's Databricks DIP was used to complement the survey findings. Structured Query Language (SQL) scripts (Appendices 10-12) were employed to extract aggregated, non-personally identifiable data. Since the Databricks DIP tables are updated daily, a snapshot was saved in the researcher's personal area on January 16, 2024, ensuring consistency for subsequent analysis.

Data was extracted from a table containing details of the manual data processing[29] tasks performed by the PTHB DEA Team within their ADO system. Data was gathered to identify data quality issues, including duplication of patient records, missing NHS numbers, and mismatched demographic details for the same individual across multiple systems. The eight systems included within this base dataset are not exhaustive, but the most common systems with existing data flows into the PTHB Databricks DIP. The study also examined duplication of activity between WPAS and WCCIS, looking for instances where activities were recorded for the same patient, specialty[30], and date.

Data handling and analysis methods are discussed in the next section.

Data Handing and Analysis

The analysis of both qualitative and quantitative data in this study followed an approach based on Onwuegbuzie and Teddlie's (2003) conceptualisation of mixed methods data analysis. This model guides analysis through seven-stages; data reduction, display, and transformation, followed by correlation, consolidation, comparison, and integration of the datasets. Visualising data is often a critical step, aiding the interpretation of results by making trends and outliers more identifiable (Unwin, 2020). Tools such as Microsoft Excel, PowerBI

---

[28] In research, this often relates to materials which have not been produced specifically for research purposes (Clark *et al.* 2021).

[29] Within this context, manual data processing tasks can be considered as activities that require human intervention to collect and ingest data from isolated data sources, such as Excel files, into the PTHB Databricks DIP.

[30] As per the NHS Wales Data Dictionary 'Specialties are divisions of clinical work which may be defined by body systems e.g. dermatology, age e.g. paediatrics, clinical technology e.g. nuclear medicine, clinical function e.g. rheumatology, and group of diseases e.g. oncology or combinations of these factors.' (2022)

and the Matplotlib Python package within Databricks DIP were utilised to generate statistical calculations and graphical visualisations.

Data ethics, as defined by Floridi and Taddeo (2016), involves handling data responsibly throughout its lifecycle, addressing storage duration, access, and secondary use. In this study, data was securely stored on NHS cloud infrastructure and retained only for the study's duration. Responses from MS Forms were accessible solely to the researcher, and if necessary, an academic supervisor. As Creswell (2013) advises, participant anonymity was preserved by presenting composite data, using anonymous responses and unique participant identifiers. Confidentiality measures and thoughtful study design minimised potential harm, supporting non-maleficence (Beauchamp and Childress, 2013). The qualitative and quantitative components were analysed independently before integration, where both components complemented each other.

The following subsections outline individual analysis approaches.

Qualitative Analysis

Thematic analysis[31] was undertaken to identify recurring themes and patterns within the open qualitative survey responses. These were analysed systematically in line with the *Framework Method* which consists of a structured process to organise and identify themes (Gale *et al.* 2013). Familiarisation was obtained by reading participant responses multiple times to gain a deep understanding of the content, followed by the interpretative method of coding (Clark *et al.,* 2021), whereby a code was assigned to a word or short phrase summarising topics (Saldaña, 2009). Initial codes described the researcher's interpretation of each passage, which were then grouped into categories and refined to identify themes, as demonstrated in Figure 5. After repeated analysis, following the first twelve responses no new themes emerged. This concept known as saturation suggests that further data collection would likely yield similar perspectives, indicating the study has explored the topic in sufficient depth (Ahmed, 2025). The themes were subsequently summarised in a table, accompanied by illustrative quotes linked to unique participant identifiers, prefixed with a data code (p.8), maintaining the connection between participants and their roles. By breaking down the themes in this way common patterns were visualised, allowing for the identification of certain CMO configurations (Gale *et al*., 2013; Wong *et al.,* 2016). This also demonstrates that the themes include perspectives from a range of participants, rather than a limited few.

---

[31] Thematic analysis is defined by Clarke *et al.,* (2021) as a term used in connection with the analysis and extraction of key themes from qualitative data.

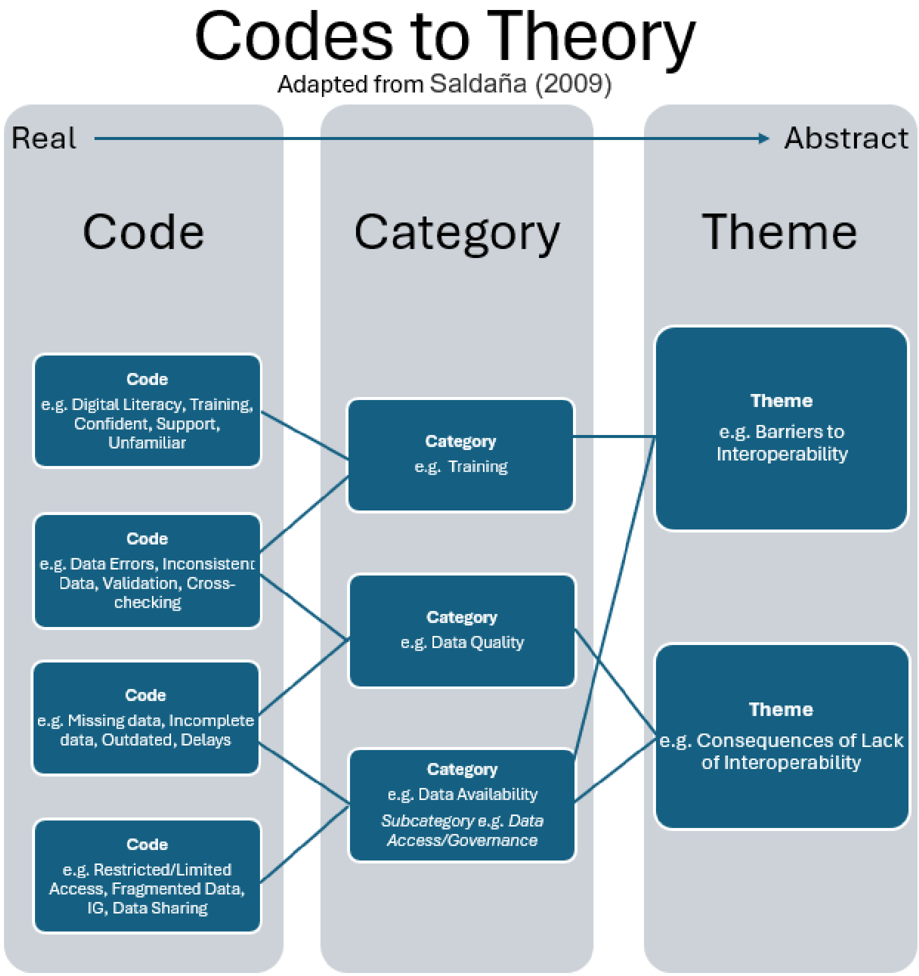


Figure 5. Example Code-to-Theme Process

Quantitative Analysis

Quantitative analysis methods can range from simple descriptive statistics to complex inferential statistics like t-tests and analysis of variance (ANOVA) (Curtis and Drennan, 2013). Appropriate method selection depends on factors such as the research aim, data distribution and type (Mishra *et al.,* 2019). Robson (2024) suggests that straightforward analysis, using descriptive statistics, tables, and visual displays, is often all that is required to draw meaningful insights—a view supported by Cohen (1990) and Gorard (2006) who promote thoughtful, purposeful analysis over unnecessary statistical complexity. In this study, frequencies were calculated and presented in tables for both the closed questionnaire responses and documentary sources. Measures of central tendency (mean, median) and dispersion (standard deviation) were also calculated to understand data distributions. These are standard methods for identifying representative values and understanding variability in a dataset (Mishra *et al.* 2019). One limitation of the mean is its sensitivity to extreme values; therefore, the median (positional average) was also calculated as it is less affected by skewed distributions (Mishra *et al.,* 2019). Group differences were also illustrated using matrices and bar charts within PowerBI.

Correlation, the analysis of relationships between variables, is quantified using the correlation coefficient, a statistical measure that indicates the strength and direction of the relationship between two variables (Clark *et al.,* 2021). Correlation was calculated within this study to identify patterns of demographic data quality errors associated with specific systems or fields (Appendix 13). First, data was normalised using z-scores[32], making the correlations more comparable. The Python `corr()` method was then used to find the Pearson correlation coefficient[33] between fields and systems, and visualised within a correlation matrix. A hierarchical clustering dendrogram was then created to group similar systems based on their mismatch patterns using Ward's method[34].

Appropriate methods were carefully considered for analysing Likert scale data, given its ordinal nature and historical concerns about the appropriateness of parametric tests. These concerns stem from the assumption that the differences between response options may not be equidistant (Sullivan and Artino, 2013). While parametric tests, which assume a normal distribution of results, have been criticised for use with Likert data, Norman (2010) argues that they are robust enough to be used even when assumptions, such as normal distribution, are violated. In contrast, Jamieson (2004) recommends non-parametric tests, such as Spearman's rank correlation to assess the strength of associations. In this study, Likert scale data was first converted to values (1-to-5), with negatively phrased items reverse-scored to ensure consistent directionality (Sexton-Radek and Simmons, 2018). Spearman's rank correlation was then calculated for pairwise comparisons using the `scipy.stats.spearmanr` function[35] (Appendix 9). The MS Forms visualisation was also used to compare the frequency distribution of Likert responses (Sullivan and Artino, 2013).

Research Integrity and Rigour

Rigour relates to the integrity, competence and legitimacy of the research process (Tobin and Begley, 2004). Detailing and justifying chosen methods can enhance rigour (Brown *et al.,* 2015; Tobin and Begley, 2004), although establishing rigour in mixed method research remains debated, partly due to differing philosophical underpinnings and evaluation methods (Harrison *et al.* 2020). For example, the rigour criteria within quantitative research (reliability and validity) are not necessarily appropriate for qualitative research (Clark *et al.,* 2021).

---

[32] The z-score is calculated by subtracting the mean and dividing by the standard deviation, and is a common method used to handle outliers and standardise data. (Codeacademy, n.d.)

[33] The Pearson correlation coefficient is typically used for normally distributed data that have a linear relationship with each other. (Schober *et al*, 2018)

[34] Ward's Clustering Method is a type of agglomerative (hierarchical) clustering algorithm and seeks to minimise overall cluster variance by computing the sum of squares (squared differences between the groups). (Murtagh and Legendre, 2014)

[35] 'SciPy is an open-source scientific computing library for the Python programming language, and a de facto standard for leveraging scientific algorithms'. (Virtanen *et al.,* 2020, p.261)

Believing that no single truth exists, Lincoln and Guba (1985) proposed trustworthiness and authenticity as alternative criteria for qualitative rigour. This aligns with the pragmatic epistemology of this study, within which the trustworthiness criteria of credibility, transferability, dependability, and confirmability are applied.

To establish credibility, the study aimed to accurately reflect participant perspectives through direct quotes, and by triangulating qualitative and quantitative data offer a realistic view of healthcare data issues. Benge *et al.* (2012) suggest that research validity is enhanced by commensurability between different data types, which was seen in this study's qualitative and quantitative result alignment.

The study's findings were compared with the literature review to assess transferability and applicability within similar settings. By including a range of professional roles and locations, demonstrating Lincoln and Guba's (1985) authenticity criteria of fairness, the study aimed to represent different perspectives, offering insights that could be transferable to other health organisations. The sample selection rationale sought alignment with the study's objectives, however, intentionally selecting participants may have introduced sample bias (Clark *et al.,* 2021), limiting perspectives to those digitally engaged.

To support confirmability, the study demonstrates data transparency by using non-personally identifiable data. Ethical principles, particularly the protection of participants' rights and well-being, maintained the integrity of the research process (Atenas *et al.,* 2023). Verbatim quotations from multiple participants enhances transparency and trustworthiness. Including the data collection and analysis code (Appendices 9 to 13) supports reproducibility, reinforcing the study's reliability and dependability. To further enhance dependability, consistency and relevance were also maintained between research questions, data collection methods, and analysis techniques (Johnson *et al.,* 2020; Robson, 2024).

Reflexivity is also considered an essential aspect of research rigour (Johnson *et al.,* 2020), and while journals are commonly used to document study reflections, this study maintained reflexivity through regular discussions with the research supervisor. This helped identify and address any biases or assumptions brought to the study. Additionally, self-reflection on the researcher's background and design choices helped identify potential influences on the research process.

Summary

In summary, this study employs a pragmatic mixed methods design, integrating both qualitative and quantitative data to address the research aim and objectives. Methodology choice was guided by best practices for robust research design (Creswell and Creswell, 2023; Clark *et al*., 2021), aiming for result credibility, depth, and transferability. By demonstrating the complementarity of the data through triangulation, the study aims to substantiate its findings from multiple perspectives contributing to a cohesive narrative that strengthens the study's rigour. Furthermore, this integration intended to enhance the authenticity of the study, and its potential for meaningful and practical applicability.

Chapter 4: Results

Introduction

This chapter presents the research findings, integrating both qualitative and quantitative data as per the mixed methods design. After summarising participant backgrounds, findings are organised by the three research objectives. Quantitative data from closed survey questions and documentary sources, are combined with thematic qualitative insights and verbatim participant quotes from open survey questions.

Questionnaire Participant Background

The online questionnaire received 95 responses, with 66.3% coming from PTHB Staff. However, this represents only 28.3% of the targeted PTHB sample, indicating a lower-than-expected participation rate. Responses where highest amongst administrative/non-clinical (35.8%) and digital/technical (30.5%) staff. Only 10.5% of participants were clinical staff, the second-lowest response rate of the targeted groups (9.1%). Notably, 80% of targeted clinical senior managers responded. The full breakdown can be seen in Table 6.

Table 6. Questionnaire Responses Per Organisation and Role Type

Organisation	Administrative	Clinical	Clinical Informatics/ Digital	Digital /Technical	Senior Manager (Non-Patient Facing)	Senior Manager (Clinical)	Grand Total	% of Target Sample	% of Response Sample
Aneurin Bevan University Health Board	1						**1**	33.3%	1.05%
Betsi Cadwaladr University Health Board		1					**1**	14.3%	1.05%
Cardiff and Vale University Health Board						1	**1**	14.3%	1.05%
Cwm Taf Morgannwg University Health Board				1			**1**	14.3%	1.05%
Digital Health and Care Wales	2			3			**5**	62.5%	5.26%
Health and Care Research Wales					1		**1**	33.3%	1.05%
Health Education and Improvement Wales								0.0%	0.00%
Hywel Dda University Health Board	1		1	3		1	**6**	150.0%	6.32%
NHS Executive	2			4	1		**7**	58.3%	7.37%
NHS Wales Shared Services Partnership	1						**1**		
Powys County Council	1			3	1		**5**	50.0%	5.26%
Powys Teaching Health Board	26	9		13	6	9	**63**	28.3%	66.32%
Swansea Bay University Health Board						1	**1**	33.3%	1.05%
Velindre NHS Trust				1			**1**	25.0%	1.05%
Welsh Ambulance Services NHS Trust				1			**1**	50.0%	1.05%
Grand Total	**34**	**10**	**1**	**29**	**9**	**12**	**95**	**32.2%**	
% of Target Sample	35.4%	9.1%	5.3%	59.2%	150.0%	80.0%	**32.2%**		
% of Response Sample	35.8%	10.5%	1.1%	30.5%	9.5%	12.6%			

Figure 6 (p.44) shows responses by organisation and role type to a closed question on individual data use. A substantial proportion of respondents selected that they actively seek out and leverage data (31.6%) or are skilled with data and frequently asked to share insights (35.8%). While 26.3% identified they have access to data but face challenges using it effectively, only 5.3% use data occasionally when it is presented in an understandable format, and just 1% stated that data is not critical to their role. This suggests that most participants recognise some level of importance in data use. Data engagement was highest among digital/technical roles, with 21 reporting data skills. Clinical/patient-facing and senior management roles showed mixed data engagement levels, with some encountering challenges in data use.

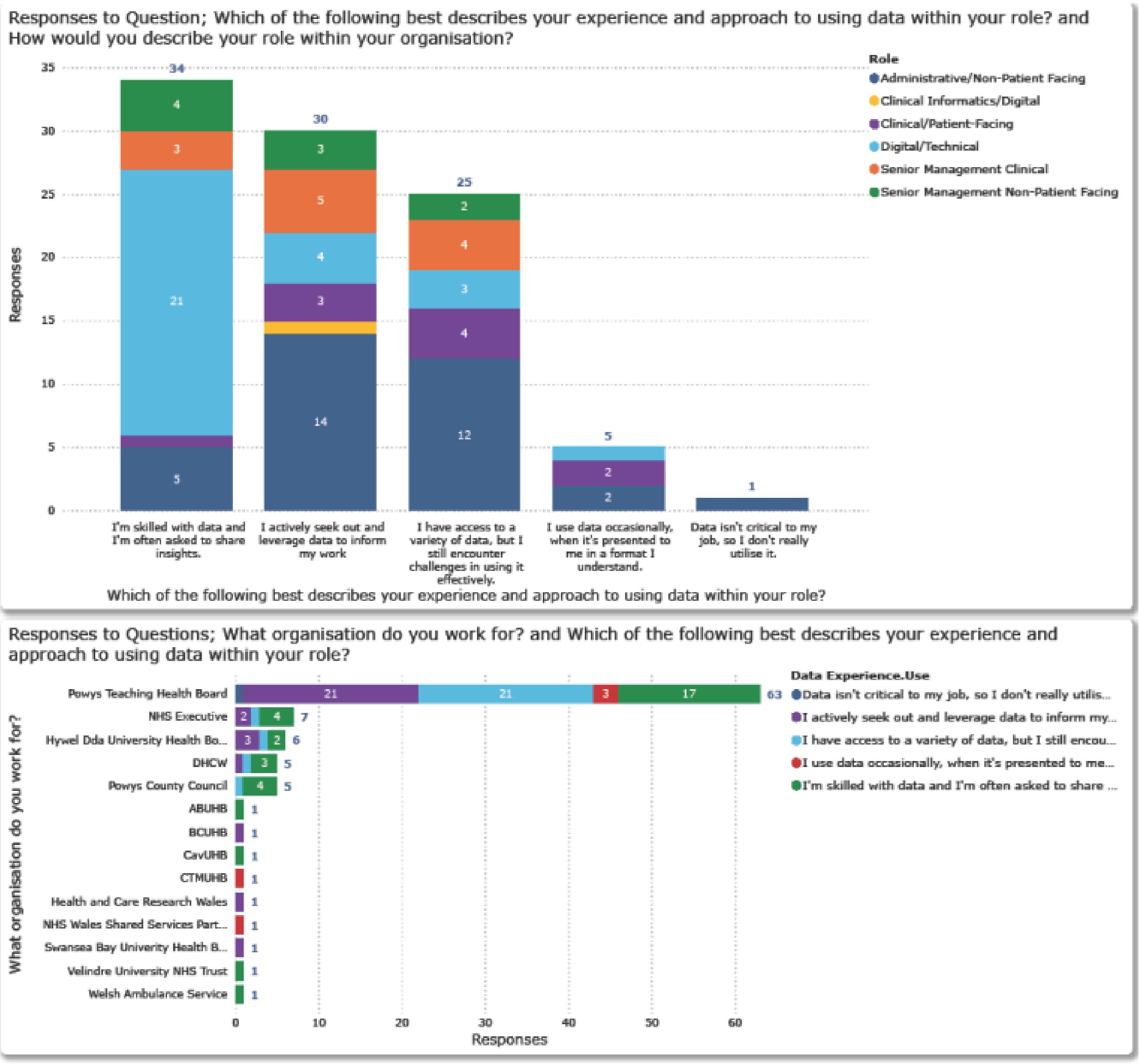


Figure 6. Questionnaire Responses Relating to Data Use and Experience Per Role Type and Organisation

Figure 7 (p.45) displays a word cloud of participants service departments. Frequently mentioned terms appear larger, with 'Digital', 'Intelligence', 'Business', and 'Data' displaying prominently. Generic terms like 'team', 'department', and 'services' were excluded for clarity. The prominence of digital and data-related terms aligns with higher engagement from digital/technical staff, while fewer clinical terms align with the lower clinical response rate. Understanding participants' background including their levels of data engagement provides context for interpreting the findings. The subsequent sections present the results in response to each research objective.

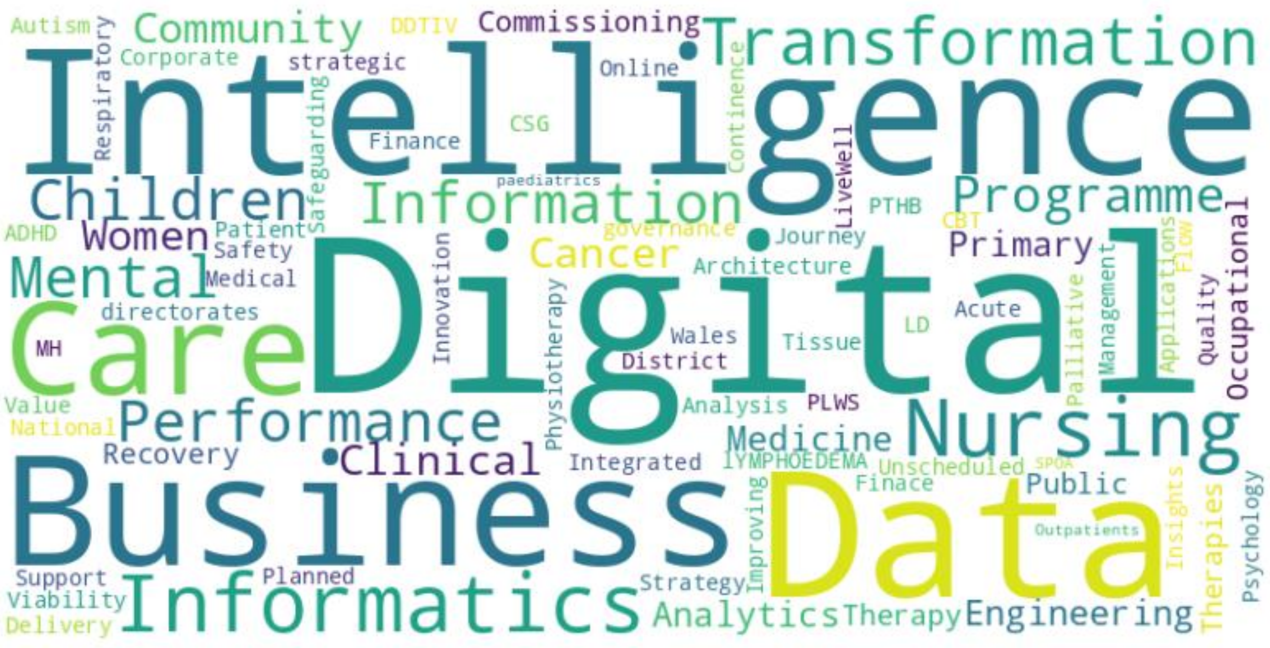


Figure 7. Word Cloud for Participant Service Departments

Research Objective 1. How do healthcare professionals perceive the interoperability of the systems they use, and what impact does this have on their work?

Perceptions of Interoperability

The categories related to the theme of the potential of interoperable systems are demonstrated within Table 7.

Table 7. Thematic Analysis: Potential of Interoperable Systems/RealTime Data Flows

Theme	Category	Representative Quotes
Potential of Interoperable Systems/Real-Time Data Flows *(Respondents acknowledge progress and express hope for future system improvements.)*	**Provide a Unified/Single Source of Truth**	“The positive will be having a unified view of data across multiple systems, automation and the fact that we can allow teams across departments to access shared data and obtain cross-functional insights.” (SMA-O7) "Benefits include being able to get a bigger ‘whole’ picture of a patient journey when able to access elements from different systems." (DT31) “I have come to realise the real positive potential of moving towards interoperable systems and real time data flows. However, there are just so many blockades (mainly political/management level) it feels like an impossible task or at least for the foreseeable.” (DT49)
	Improve clinical-decision making and patient care/outcomes	“Real time, interoperable systems would lead to improved health care by providing clinicians with more complete information about patients and by enabling insights through more complete datasets for analysis.” (DT46) “There is also good access to live data. This allows for informed decision making across the service.” (DT73) “Interoperable systems are essential for safe patient care and will improve the user experience. Data flow and availability will provide efficiencies and release clinicians from the computer to provide more time to care for patients. “ (DT94) “Interoperable systems are key to the success of the NHS for improving patient outcomes.” (SMA-O96)
	Enhance Data Quality and Efficiency	“More work need to be done either with teams across the health board to learn how we can integrate data. Resource needs to be allocated to better systems that work interconnectivity to achieve better data hygiene and efficiency.” (A-T30) “Data flow has improved in the last three years and more data is available digitally but data flow is still clunky and requires validation> The current data flow has improved success and provided more immediate feedback to teams. Efficiency can be improved by a more integrated system across the HB (or even better across Wales!....... or the UK!)” (A-MH50) “The future of healthcare will completely rely on faster data flows which are system agnostic.” (DT60) “Where interoperability works well, this seemingly reduces the work load and risk of transcription error between systems. However, this also allows data outputs to be more operational applicable in presenting a whole picture, as opposed to the feeling of isolated data inputs. Improved interoperability allows the focus to shift from input to output and improved perception of the data processes as a result.” (SMC-C70)

Overall, questionnaire respondents expressed optimism about interoperable systems and real-time data flows, while acknowledging significant ongoing challenges, as DT49 noted:

> I have come to realise the real positive potential of moving towards interoperable systems and real time data flows. However, there are just so many blockades (mainly political/management level) it feels like an impossible task or at least for the foreseeable.

Several participants emphasised that interoperability is essential for safe patient care and improving outcomes. Greater data availability and smoother data flows were seen as key to increasing efficiency and reducing administrative burden.

Despite some recognition of data integration progress, frustrations with current interoperability were common (Table 8).

Table 8. Thematic Analysis: Current State of NHS Interoperability

Theme	Category	Representative Quotes
Current State of NHS Interoperability *(Healthcare professionals perceive system interoperability as poor, leading to fragmented data systems, manual workarounds, duplication of effors and operational inefficiencies)*	**Fragmented and Inconsistent Data**	"The data flows are inconsistent, glitchy and not always accessible due to system errors - both software and hardware." (C-T40) “Systems don’t always talk to each other, especially across Wales. Each health board has its own instance of the system, so data can be inconsistent or incomplete.” (DT52) "Cross-border data flows not taking place.” (A-O85) "Silo working has led to 7 versions of the truth. each HB collecting and displaying data their own way, making it challenging to compare and develop national solutions to meet all niche requirements." (A-PC36) "not all systems tell the same story so this can be challenging to unpick what's really happening within a system." (SMA-PC41)
	Siloed Systems and Disjointed Processes	"Individual systems can work well, lack of communication between systems hinders progress." (SMC-T22) “The biggest gripe I have is about the different systems used for collecting data across health boards.” (DT79) "Data is essential for safe and effective patient care. The benefit of accessing a unified patient record is highlighted when treating patients receiving care in areas not using the same system, leading to frustrations and additional workload. This can lead to delays, omissions or errors.” (DT94) “Different standards, different requirements of the systems and services.” (DT60) "There are challenges sharing with LA colleagues which therefore require work around which result in challenges along with an inability to be timely." (SMC-WC39)

Participants highlighted the need for better collaboration, investment in systems that facilitate seamless data sharing and broader integration, as highlighted by A-T30:

> More work needs to be done either with teams across the health board to learn how we can integrate data. Resource needs to be allocated to better systems that work interconnectivity to achieve better data hygiene and efficiency.

These qualitative findings align with the Likert scale responses (Table 9, Figure 9, p.48). The strongest agreement was with non-weighted/neutral statements, for example 75.8% of respondents strongly agreed that ‘Timely access to data is important to improve patient care’. Similarly, 61.1% strongly agreed that ‘Improved data interoperability would enhance my ability to perform my job’. The impact of data challenges on clinical decision-making is evidenced by 85.3% agreeing or strongly agreeing that ‘delayed or incomplete data negatively affects clinical decision making and direct patient care’.

Both qualitative and quantitative findings highlight poor system interoperability, with recurring frustrations over fragmented data systems. A-PC9 noted, ‘Systems don’t talk to each other. We find it difficult to follow a complete patient pathway’. DT79 echoed this, stating ‘The biggest gripe I have is about the different systems used for collecting data across health boards.’

Relatedly, questionnaire responses regarding data sources used (Figure 8) revealed Excel as the most widely used tool (86.3%), especially by administrative/non-patient facing, digital/technical, and senior management roles. Cloud-reports (n=74) and front-end systems (n=68) are commonly used across all roles, with the latter most common among clinical staff. Despite digital advancements, paper notes (n=33) are still used, especially by administrative and senior management roles.

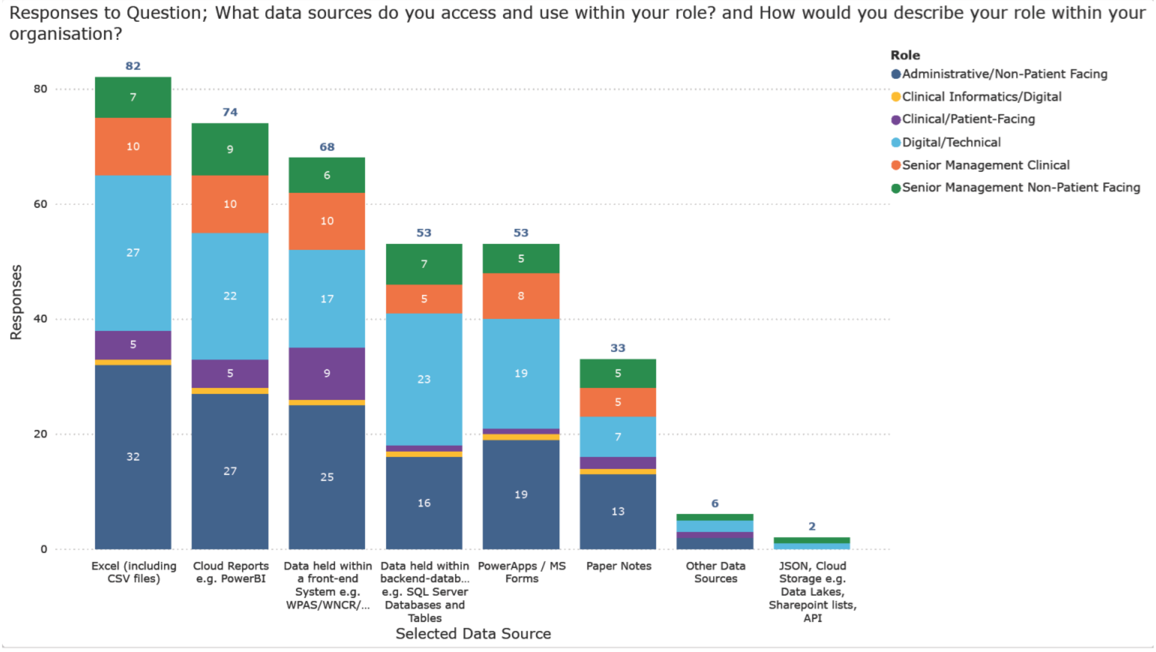

Responses to Question; What data sources do you access and use within your role? and How would you describe your role within your organisation?

Figure 8. Data sources used by questionnaire participants by role type

Poor system interoperability is also evidenced by the questionnaire’s Likert scale responses (Table 9, Figure 9, p.48). Around half of respondents agreed with two negatively worded Likert statements; ‘I often find that data is incomplete or missing when I need it’ (50.5% agreed), and ‘I frequently experience delays in accessing the latest data’ (44.2% agreed or strongly agreed). This is broadly consistent with the respondents who disagreed with the statement that data flows between systems in their organisation are seamless and efficient (45.3% disagreed, 29.5% strongly disagreed). 70.5% disagreed that necessary data is accessible from a central location, providing further indication of fragmented and siloed data systems. General satisfaction with current data systems was moderate, with 47.9% disagreeing that they were satisfied with data systems and availability.

Table 9. Number of Likert Scale Responses per Question

Likert Question	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree	Total
Data flow between systems in my organisation is seamless and efficient	1	5	18	43	28	**95**
Delayed or incomplete data negatively affects clinical decision making and direct patient care	40	41	13	1		**95**
I am satisfied with the current data systems and data available to me	1	15	33	32	13	**94**
I can access all the data I need from a central location	3	17	8	45	22	**95**
I experience challenges when trying to integrate/obtain data from multiple systems	38	38	12	6	1	**95**
I frequently experience delays in accessing the latest data	11	31	22	29	2	**95**
I often find that data is incomplete or missing when I need it	9	48	16	21	1	**95**
I rarely encounter issues with the reliability of data which I access	2	19	20	46	8	**95**
I trust the accuracy of data I use for decision making	6	48	30	11		**95**
Improved data interoperability would enhance my ability to perform my job	58	33	4			**95**
The data I access is up to date	5	34	36	16	3	**94**
Timely access to data is important to improve patient care	72	16	5		2	**95**
Total	**246**	**345**	**217**	**250**	**80**	**1138**

Figure 9. MS Forms Chart Displaying Percentage of Likert Scale Responses per Question

Relationships between Likert scale questions were explored using Spearman's rank correlation (Appendix 9. Python (Pyspark) Code and Full Results of Likert-Scale Question Pair Correlation), with the most statistically significant and strongly correlated results shown in Table 10 (p.49). The strongest negative correlation (-0.515) was observed between the statements 'Improved data interoperability would enhance my ability to perform my job' and 'I experience challenges when trying to integrate/obtain data from multiple systems'. This correlation suggests the greater the challenge in integrating or obtaining data, the stronger the belief that improved interoperability would enhance job performance. The strongest positive correlation (-0.524) was found between ‘I can access all the data I need from a central location’ and ‘I frequently experience delays in accessing the latest data’, indicating that centralised data access could be associated with fewer access delays. Additional positive correlations indicate that data reliability, completeness and timeliness are closely linked. These findings suggest that data integration, access, and quality are central to perceptions of interoperability, supporting the need to address these challenges to improve system effectiveness and job performance.

Table 10. Likert-Question Pair Correlation

Likert Scale Variable Pair	Correlation (r)	p-value
I can access all the data I need from a central location – I frequently experience delays in accessing the latest data	0.524	0.0
Improved data interoperability would enhance my ability to perform my job – I experience challenges when trying to integrate/obtain data from multiple systems	-0.515	0.0
I rarely encounter issues with the reliability of data which I access – I often find that data is incomplete or missing when I need it	0.505	0.0
I am satisfied with the current data systems and data available to me – Data flow between systems in my organisation is seamless and efficient	0.499	0.0
I rarely encounter issues with the reliability of data which I access – I frequently experience delays in accessing the latest data	0.488	0.0
I am satisfied with the current data systems and data available to me – I rarely encounter issues with the reliability of data which I access	0.473	0.0
The data I access is up to date – I frequently experience delays in accessing the latest data	0.472	0.0

Issues with data access and data sharing, especially between Welsh and English service providers, further demonstrate operational challenges affecting interoperability, as highlighted in the quote:

> The biggest challenge is having access to data on our patients on pathways into English service providers. This skews the data available for informing and assessing quality improvement initiatives. (C-M81)

This is supported by DT1, who comments ‘Data sharing agreements often aren't in place to allow for timely and accurate reporting between organisations’, and DT2, who adds ‘It is a challenge to get NHS England to share data’.

Despite the recognised need for improved interoperability, Figure 10 (p.50) highlights a lack of familiarity (n=48) with key healthcare data standards and technologies, particularly within administrative/non-patient facing roles. This unfamiliarity with data standards and protocols likely contributes to the challenges observed in data sharing and system interoperability.

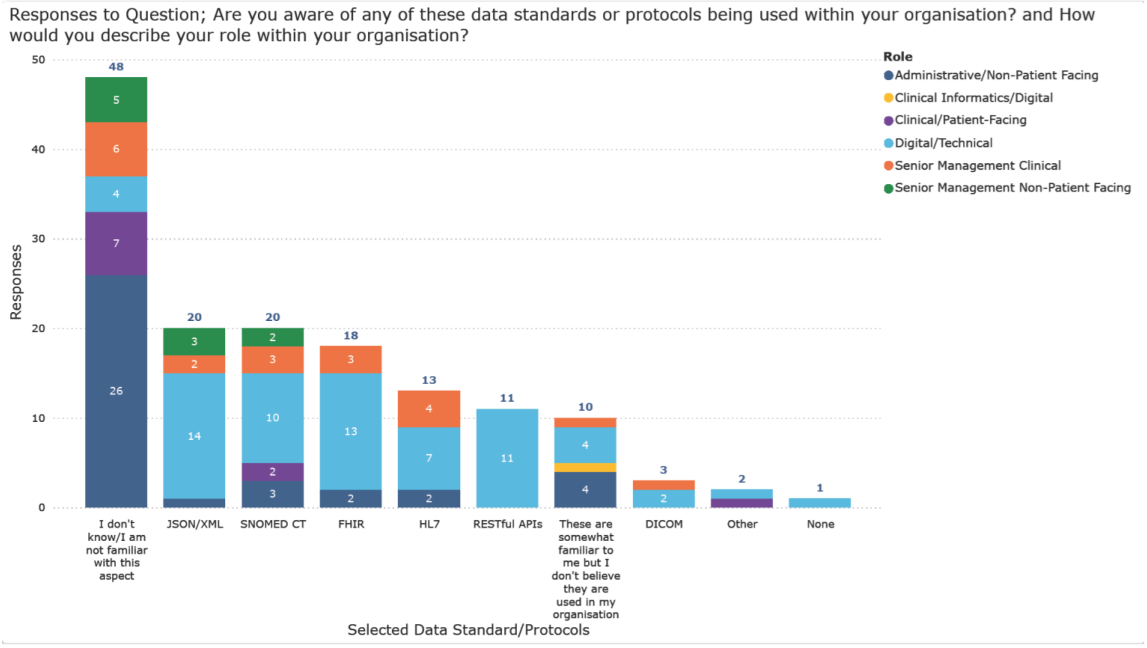


Figure 10. Selected data standards/protocols used within questionnaire participants organisation by role type

The perception of widespread system fragmentation and inadequate data sharing highlights some significant barriers to achieving seamless data flow and subsequently, has both operational and clinical consequences, as discussed in the following section.

Consequences of Siloed Systems

Operational challenges resulting from siloed systems, illustrated in Table 11, include data quality issues, work inefficiencies, reliance on manual workarounds and duplicated efforts to maintain accurate records

Table 11. Thematic Analysis: Consequences of Siloed Systems

Theme	Category	Representative Quotes
Consequences of Siloed Systems *(Healthcare professionals perceive system interoperability as poor, leading to fragmented data systems, manual workarounds, duplication of effors and operational inefficiencies - including impacting on patient care)*	**Data Quality Issues**	"Duplicate databases lends itself to data errors, it is time consuming keeping two databases and various spreadsheets up to date." (A-MH45) "Contradictions in data with multiple systems—BI using live data, other systems are more static—exact figures are misaligned." (A-MH54) "Errors and typos can be created when transferring data between systems." (A-O71) "Impacts on ability to provide accurate data" (A-O37) "If they are set up correctly there are no issues, but occasionally reporting measures are different which gives conflicting data" (SMC-WC39)
	Risk to patient care/service improvements	"Incomplete data access makes patient care dangerous." (C-T40) "The lack of interoperability between clinical systems is resulting in duplication of effort, potential missed information, increased risk to patient safety, dual running paper and digital records." (CI75) "GDPR mean that linking of patient records from multiple data source both internal and external are prevented - i.e. inability to link PREMS to patient records, meaning value of the PREM is diminished as cant be accurately reviewed in conjunction with the patient pathway so aid service improvement." (SMA-F28)
	Inefficient Work Processes inc. Dual Recording	"There is data duplication and dual recording." (DT3) "Lack of interoperability of current systems brings high cost overhead of having to cross-check/update multiple systems." (A-T27) "On a daily basis, when imputing patient data, this needs top be done in some cases on 3 separate systems. This in turn leads ta greater chance of error and severely reduces my efficiency by having to input the same data manually over serval systems" (A-T30) "The lack of interoperability between systems makes it difficult to cross-check data and compare and contrast different aspects" (DT89)

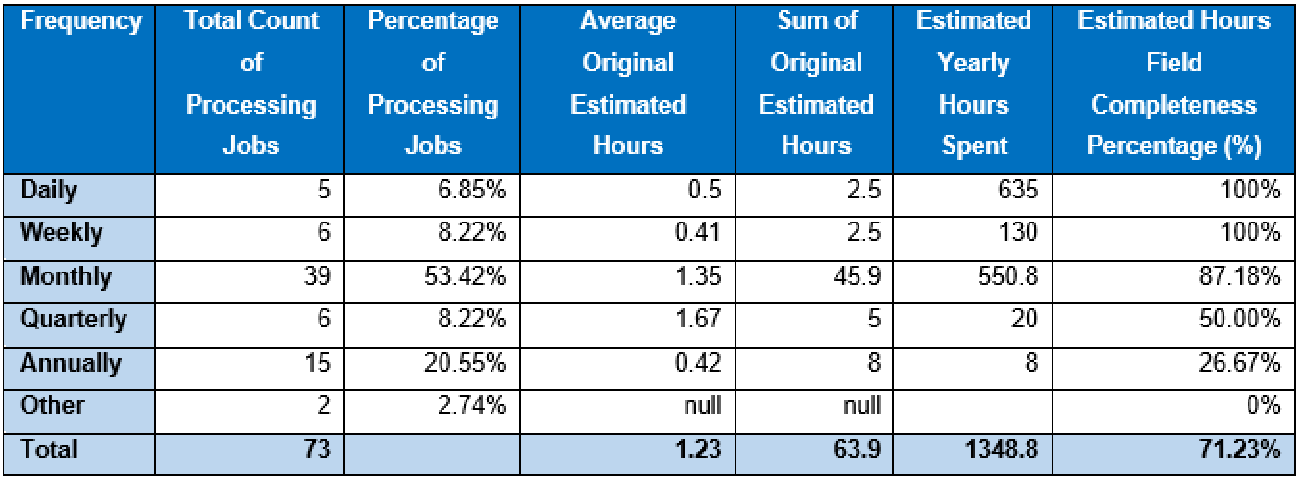
Many participants described these challenges, highlighting that non-integrated systems create additional workload, increase the risk of errors, and ultimately impact patient care. 'Data duplication and dual recording'(DT3) across multiple systems was the most frequently cited issue, and as noted by A-MH45, 'Duplicate databases lends itself to data errors *[sic],* it is time consuming keeping two databases and various spreadsheets up to date'.

The lack of system integration often forces staff to manually transfer and verify data, increasing workloads and operational inefficiency, diverting resources from patient care. As DT10 noted, 'Systems that have not been integrated into the ecosystem have difficulty accessing data, which results in manual transcribing of data'. Duplication of data entry is described as 'a big problem and big time wasting factor' (DT66), also evidenced within the quote:

> Duplication of work is a drain on resources. Also, clinicians use one system and admin use another, so there is a lot of cross-referencing required, and this opens up a margin for error. (A-MH56)

Inefficiencies from non-interoperable data sources are substantiated by the manual processes carried out by the PTHB DEA Team. The SQL query used to identify these specific processes is included within Appendix 10. SQL Code for PTHB Data Analysis and Engineering Team Azure DevOps The findings, summarised in Table 12 reveal that the team performs 73 routine jobs, with 53.4% occurring monthly and 20.6% annually. Based on available data[36], the jobs take on average over an hour to complete, with quarterly and monthly jobs requiring the most time. When the estimated hours are multiplied by the frequency of these tasks, it is estimated that these manual processes account for over 1,340 hours of work annually[37].

Table 12. Summary Results of PTHB Data and Analysis Team AzureDevOps Active Manual Processing Jobs



Frequency	Total Count of Processing Jobs	Percentage of Processing Jobs	Average Original Estimated Hours	Sum of Original Estimated Hours	Estimated Yearly Hours Spent	Estimated Hours Field Completeness Percentage (%)
Daily	5	6.85%	0.5	2.5	635	100%
Weekly	6	8.22%	0.41	2.5	130	100%
Monthly	39	53.42%	1.35	45.9	550.8	87.18%
Quarterly	6	8.22%	1.67	5	20	50.00%
Annually	15	20.55%	0.42	8	8	26.67%
Other	2	2.74%	null	null		0%
Total	73		**1.23**	63.9	**1348.8**	**71.23%**

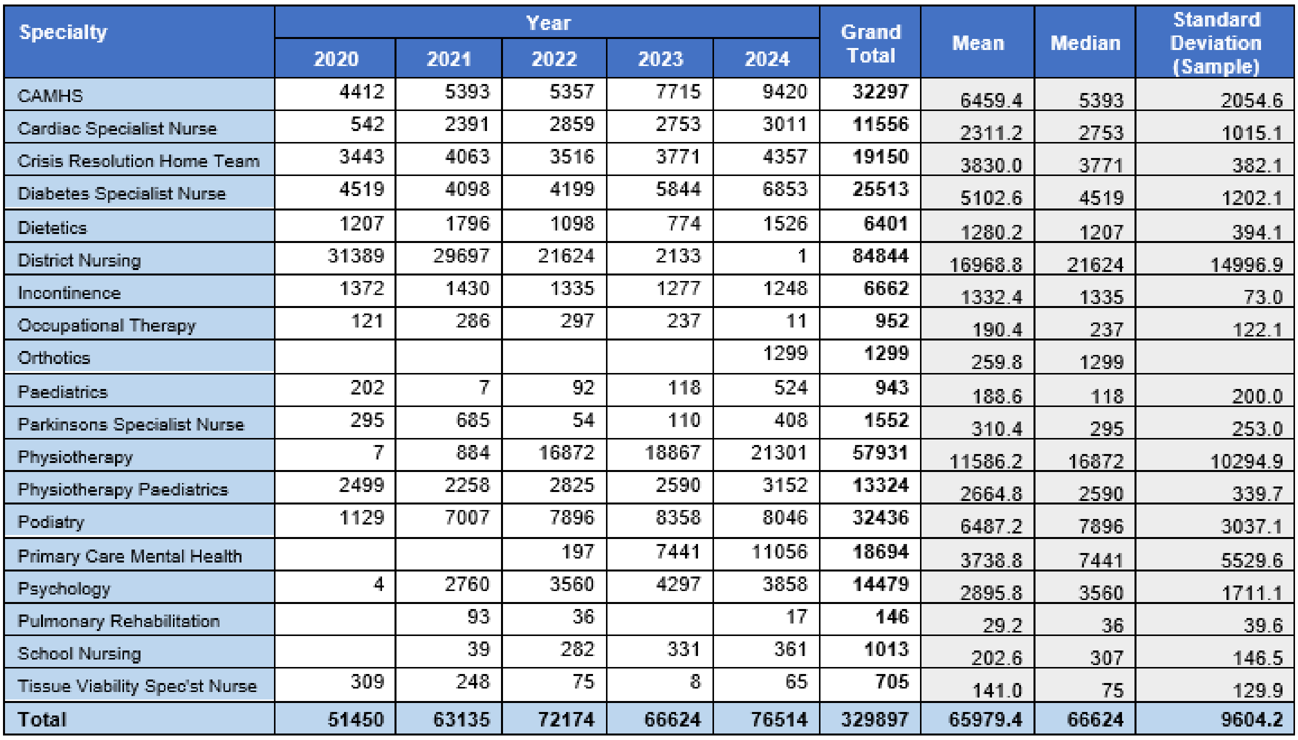
---

[36] The Estimated Hours field exhibits low completeness (<50%) for ad-hoc, annual, and quarterly jobs, indicating a potential data quality issue.

[37] This is calculated based on the on 254 working days within 2024 which excludes weekends and bank holidays.

Documentary analysis comparing activity recorded in WPAS and WCCIS also provides evidence of data duplication. The SQL query used to identify duplicate activity is included within Appendix 11. Initial analysis revealed that some NHS Numbers[38] have multiple activities recorded on the same day, within the same specialty. To prevent inflation and duplication of results when joining the datasets, only one activity per specialty for each patient on a given day was used to identify an equivalent record in the other system. As shown in Table 13, the total number of duplicated activities has increased over time, with a slight decrease in 2023, potentially due to changes in district nursing recording methods. For example, district nursing previously had the highest levels of duplication, peaking at 31,389 in 2020, yet only 1 duplicated activity in 2024. Other specialties such as orthotics (1,299 in 2024), and primary care mental health (197 in 2022, increasing to 11,056 in 2024) show more recent duplication.

Table 13. Activity Duplicated Between WPAS and WCCIS per Year and Specialty



Specialty	Year					Grand Total	Mean	Median	Standard Deviation (Sample)
	2020	2021	2022	2023	2024				
CAMHS	4412	5393	5357	7715	9420	**32297**	6459.4	5393	2054.6
Cardiac Specialist Nurse	542	2391	2859	2753	3011	**11556**	2311.2	2753	1015.1
Crisis Resolution Home Team	3443	4063	3516	3771	4357	**19150**	3830.0	3771	382.1
Diabetes Specialist Nurse	4519	4098	4199	5844	6853	**25513**	5102.6	4519	1202.1
Dietetics	1207	1796	1098	774	1526	**6401**	1280.2	1207	394.1
District Nursing	31389	29697	21624	2133	1	**84844**	16968.8	21624	14996.9
Incontinence	1372	1430	1335	1277	1248	**6662**	1332.4	1335	73.0
Occupational Therapy	121	286	297	237	11	**952**	190.4	237	122.1
Orthotics					1299	**1299**	259.8	1299	
Paediatrics	202	7	92	118	524	**943**	188.6	118	200.0
Parkinsons Specialist Nurse	295	685	54	110	408	**1552**	310.4	295	253.0
Physiotherapy	7	884	16872	18867	21301	**57931**	11586.2	16872	10294.9
Physiotherapy Paediatrics	2499	2258	2825	2590	3152	**13324**	2664.8	2590	339.7
Podiatry	1129	7007	7896	8358	8046	**32436**	6487.2	7896	3037.1
Primary Care Mental Health			197	7441	11056	**18694**	3738.8	7441	5529.6
Psychology	4	2760	3560	4297	3858	**14479**	2895.8	3560	1711.1
Pulmonary Rehabilitation		93	36		17	**146**	29.2	36	39.6
School Nursing		39	282	331	361	**1013**	202.6	307	146.5
Tissue Viability Spec'st Nurse	309	248	75	8	65	**705**	141.0	75	129.9
Total	**51450**	**63135**	**72174**	**66624**	**76514**	**329897**	**65979.4**	**66624**	**9604.2**

To understand the distribution of duplicated activity, measures of central tendency (mean, median) and dispersion (standard deviation) were calculated for each specialty. Specialties like incontinence, pulmonary rehabilitation and Parkinsons specialist nurse, show a median close to the mean, suggesting a relatively normal distribution without extreme fluctuations. These also have the smallest standard deviations, suggesting consistency over time. However, specialties such as district nursing and physiotherapy show wider gaps between the median and mean, indicating greater variability and year-to-year fluctuations.

---

[38] NHS Number is used as the unique identifier for a patients record

Figure 11. Graphical comparison of Activity per WCCIS and WPAS

Figure 11 (p.53) illustrates the percentage of duplicated activity between WPAS and WCCIS, shown relative to the total recorded activities per system and specialty. Certain specialties show a large proportion of WPAS activity is duplicated on WCCIS, such as the crisis home resolution team (95.7%) and physiotherapy paediatrics (83.4%). Equally within WCCIS, 95.5% of podiatry activity recorded also exists on WPAS. However, the figures vary between specialties; for instance, paediatrics shares only around 9% of activity between systems. Over time, the percentage of shared activity on WPAS has increased by nearly 5% annually, reaching 50.4% in 2024, and WCCIS activity has activity has grown by 58% since 2020, with 208,996 activities recorded in 2024.

Beyond inefficiencies of dual recording, the fragmentation of systems affects data quality, consistency, and reliability, as A-DN78 emphasises, ‘Having to use multiple systems leaves room for error when transferring or obtaining data. Also, it can be time-consuming.’ This reflects the operational burden caused by fragmented systems, which contribute to a more labour-intensive and error-prone workflow. While these issues will be explored further in relation to the second research objective, they also have clear consequences extending beyond operational efficiency, directly impacting patient care and clinical decision-making. For instance, A-PC36 remarked, ‘Silo working has led to seven versions of the truth’ highlighting confusion and discrepancies from multiple systems. Inefficiencies caused by such fragmentation can lead to delayed access to critical information, hindering timely decision-making, because as SMC-T13 comments ‘Data is needed to make decisions… Because it's not there, it's very difficult to argue your position’. Similarly, DT62 noted that ‘the data flows present often don’t contribute to effective decision making and very often lead to being the blocker for success’.

Improved data flows and interoperability could contribute to more effective decision-making, as supported by the quote:

> Improved data flows and interoperability would enable my team to provide more timely and accurate information needed by stakeholders to make key business decisions. We would be better placed to join different datasets from separate sources with more confidence. (DT6)

Incomplete or inaccessible records also pose significant risks and ‘makes patient care dangerous’ (C-T40), as CI12 recounted:

> Finding information written in the medical notes which are paper-based can be challenging as it is often illegible and unorganised. I have made a drug error based on the poor handwriting of a Doctor.

This emphasises the need for reliable and accessible data to ensure patient safety.

Participants also stressed the impact on patient experiences, who often face repeated questioning and inconsistencies, for example:

> Patients expect systems to be interoperable and don't understand why they have to repeat information. Boundaries between emergency care, secondary care, primary care, community care and social care often affect the most vulnerable patients as they have multi-faceted need and multiple services involved. (DT35)

Further illustrating the impact of data flow problems, Figure 12 presents the organisational aspects' participants felt were most affected. 'Reporting' is the area of concern most cited, particularly among administrative/non-patient facing and digital/technical roles (23 and 24 instances, respectively). Other notable areas affected include 'strategic planning', 'compliance', and 'auditing'. The limited number of clinician/patient-facing respondents may account for why patient care was not rated higher.

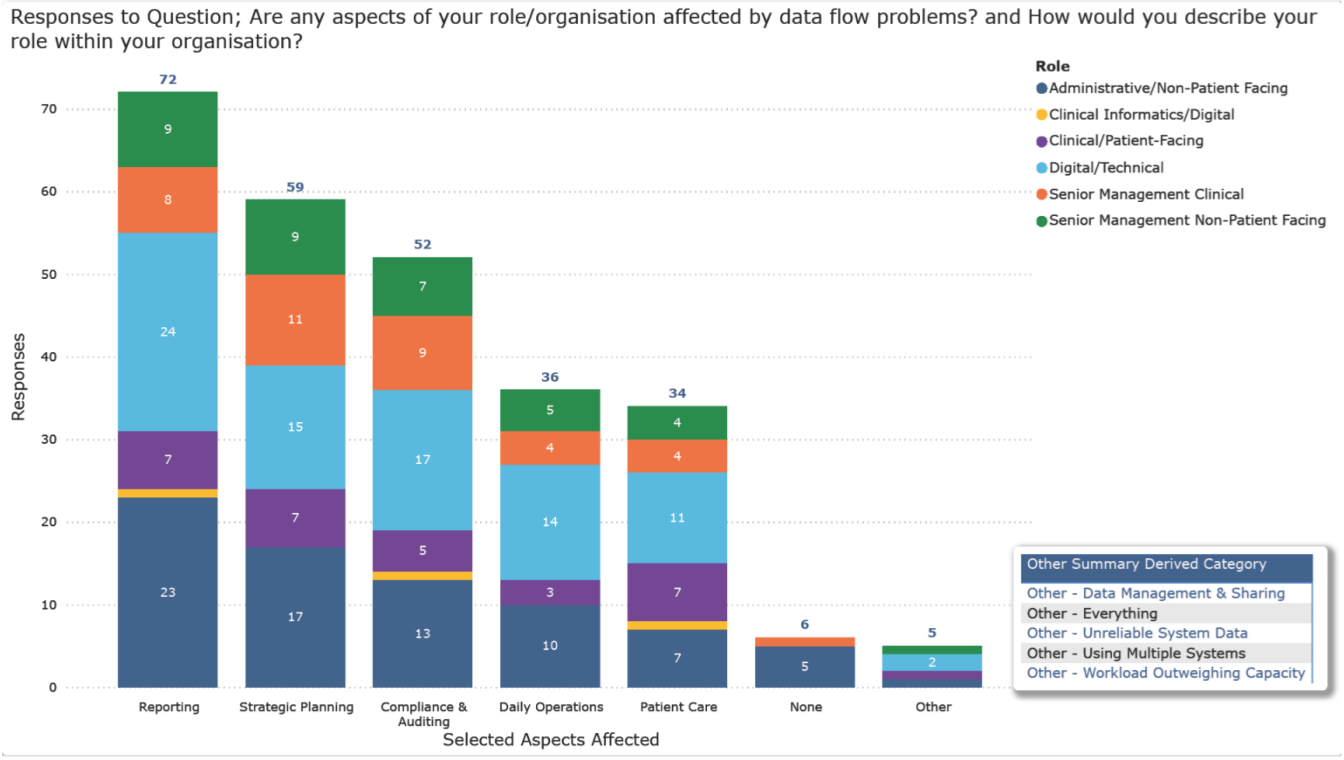


Figure 12. Selected aspects affected by data flow problems within questionnaire participants role/organisation by role type

Addressing issues within these areas requires not only enhanced data flows and interoperability but also accessible, accurate, and timely data—the focus of the second research objective.

Research Objective 2. In what ways do data quality and accessibility influence decision-making and patient outcomes?

Respondents consistently highlighted the importance of timely and accurate data for effective decision-making, optimising resource allocation, and improving patient outcomes:

> It supports patient flow to avoid patient delays and bed blockages…. More timely data is always better than out-of-date data, even if one day is outdated. Things can change in 24hrs in the NHS, especially when reporting patient data. (A-PC9)

Many described how reliable, up-to-date information supports operational efficiency and is ‘vital to delivering services’ (DT60), because ‘Without a full picture, my efforts may not be the most effective or correctly targeted’ (SMA-PC41). SMC-PH17 highlighted that ‘having access to timely and accurate data improves our planning decisions’, while A-OP47 stressed that ‘giving decision-makers access to relevant data helps improve choices made’.

Despite the recognised benefits and requirements for high-quality data, participants recounted challenges arising from incomplete, outdated, or inconsistent information across systems. The Likert responses presented in Table 9 and Figure 9 (p.48) reflect moderate confidence in data reliability with 56.8% of participants disagreeing or strongly disagreeing that they rarely encounter issues with data reliability. Only 6.3% strongly agreed that they trust the accuracy of the data used for decision-making, but 50.5% did agree, indicating moderate trust in available data sources. Only 41.5% agreed that the data they access is up to date, indicating a need for improvements in data timeliness. However, this was the question which received the most neutral responses (36 out of 94).

In complex settings cross-border collaboration and continuity of care are particularly important, as A-SN55 emphasised, ‘it is imperative that data flows are efficient for patient care and accurate continuity’. However, difficulties accessing data including restrictive data sharing agreements, particularly between Welsh and English providers, limit care coordination and accurate reporting, as noted by A-O71:

> Some systems are updated more than others, resulting in contradictory information…Inaccurate information causes delays and risks the wrong allocation of resources.

Similarly, DT3 comments ‘Having missing data shows only half the picture. We are unable to report accurately and give a full account of patient activity’.

Others highlighted the technical challenges of maintaining data integrity across multiple systems and the increased operational workload required to correct data errors:

> To ensure data flows, the integration team has multiple tasks, including mappings of

> data which results in loss of granularity of the data, sometimes missing data and overall low quality. (DT10)
>
> Sometimes the inaccuracies with data cause multiple issues within my role as it requires many changes to reports some of which are public facing, reducing team efficiency. (A-O88)

Data inconsistencies were frequently cited as a barrier to effective healthcare management directly impacting strategic decision-making and patient care. As DT80 highlights 'There are challenges on the timely and accurate availability of data needed to drive strategic decisions which can impact patient care and future service planning'. This is further supported by SMA-F28:

> Incomplete data issues, timely availability and general quality issues always result in lower standard of modelling/analysis with numerous stated caveats around data.

One participant specifically noted, 'Demographic data for the same individual is different between systems' (DT3). SMA-PC41 echoes this, commenting that, 'It's just very time consuming having to cross check information especially as there are often gaps in data and sometimes systems contradict each other'. Such concerns are substantiated by documentary analysis comparing demographic data between PTHB EHRs. The code used to obtain and analyse this data is included within Appendix 12 and 13. Table 14 (p.58) shows that WCCIS has the largest volume of records[39], followed by WPAS. Based on the Modulus 11 algorithm (NHS Wales, 2024), 15.9% of WCCIS records, 9.1% of Auditbase records, and 6.7% of WPAS records contain an invalid NHS Number, often due to the field being blank, impacting data reliability and usability.

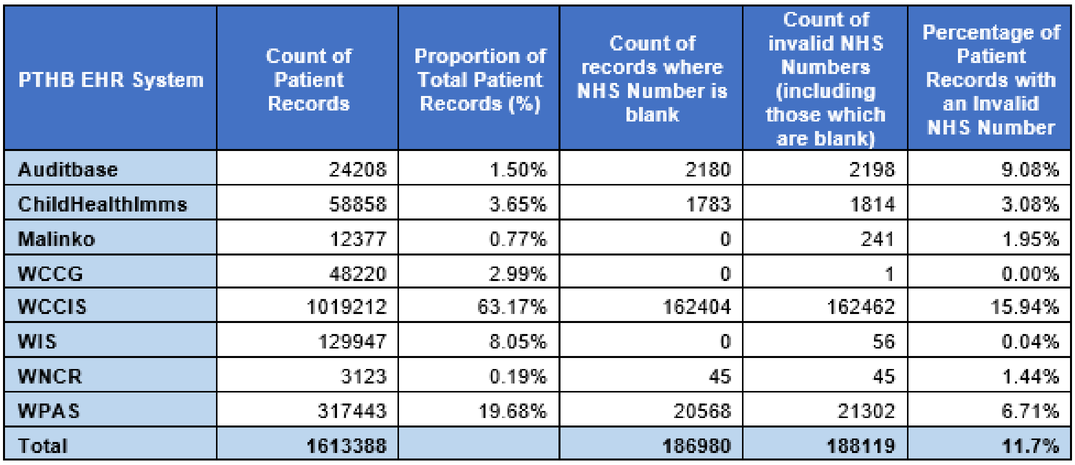
As shown in Table 15 (p.58), WCCIS contained the highest number of duplicate patient records[40], accounting for 84.4% of NHS Number duplicates and 88.50% of 'Fuzzy Logic' duplicates[40]. WPAS ranked second, while Malinko had the fewest. However, NHS Number (0.7%) and 'Fuzzy Logic' (1.1%) duplicates represent a small proportion of total records.

---

[39] This is not surprising given WCCIS serves both health and social care organisations across Wales. (Audit Wales, 2020)

[40] Duplicate patient records are identified based on multiple occurrences of the same NHS Number where present, or by using a 'Fuzzy Logic' match which is based on the concatenation of a patient surname, the first two initials of their forename and date of birth where all fields are populated. For example, DOEJA2000-01-01 would be the 'Fuzzy Logic' search string for someone called Jane Doe born 01-01-2000.

Table 14. Summary of Patient Records Per PTHB EHR System



PTHB EHR System	Count of Patient Records	Proportion of Total Patient Records (%)	Count of records where NHS Number is blank	Count of invalid NHS Numbers (including those which are blank)	Percentage of Patient Records with an Invalid NHS Number
Auditbase	24208	1.50%	2180	2198	9.08%
ChildHealthImms	58858	3.65%	1783	1814	3.08%
Malinko	12377	0.77%	0	241	1.95%
WCCG	48220	2.99%	0	1	0.00%
WCCIS	1019212	63.17%	162404	162462	15.94%
WIS	129947	8.05%	0	56	0.04%
WNCR	3123	0.19%	45	45	1.44%
WPAS	317443	19.68%	20568	21302	6.71%
Total	**1613388**		**186980**	**188119**	**11.7%**

Table 15. Summary of Potential Duplicates Per PTHB EHR System

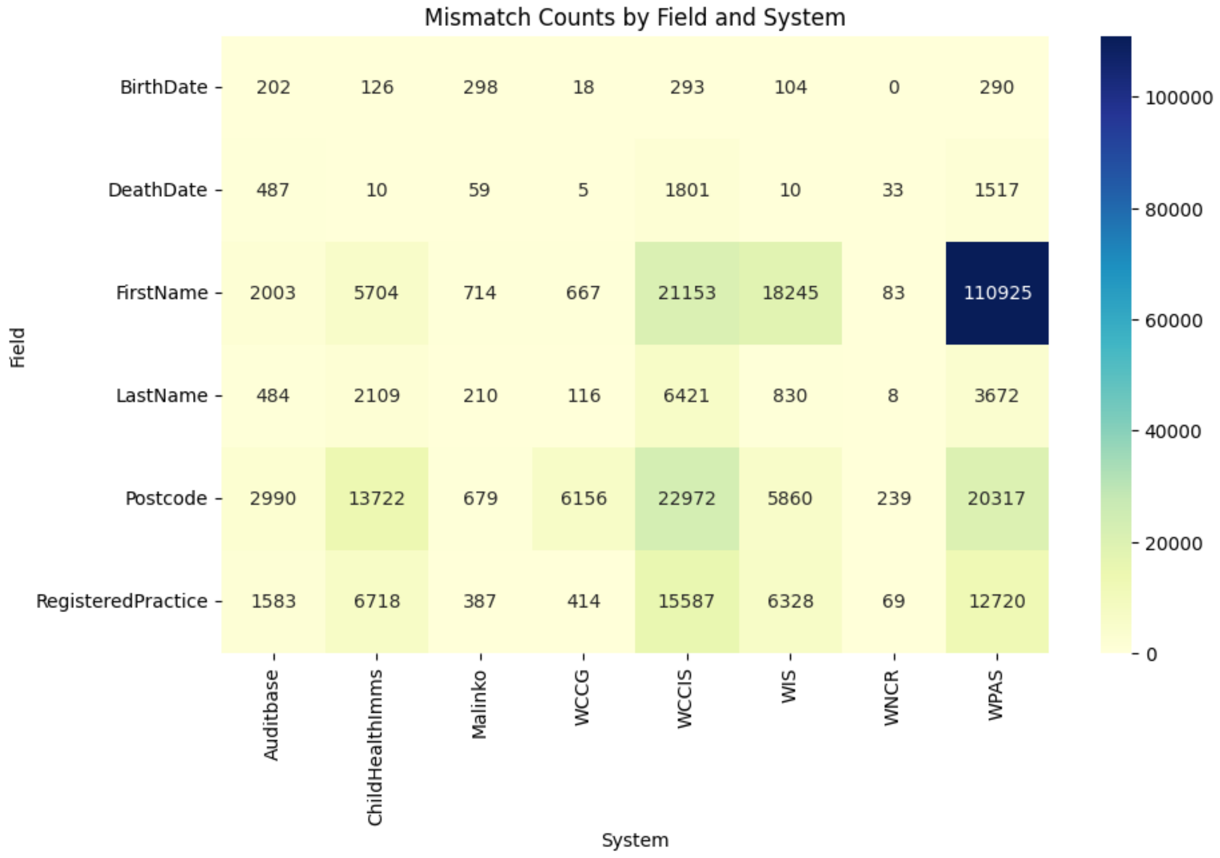
PTHB EHR System	Duplicate Patient Records based on NHS Number	Proportion of Total Duplicate Patient Records Based on NHS Number (%)	Duplicate Patient Records based on Fuzzy Logic Search String	Proportion of Total Duplicate Patient Records Based on Fuzzy Logic Search String (%)
Auditbase	0	0.00%	78	0.43%
ChildHealthImms	1	0.01%	93	0.51%
Malinko	0	0.00%	59	0.33%
WCCG	254	2.44%	174	0.96%
WCCIS	8789	84.42%	16062	88.50%
WIS	127	1.22%	205	1.13%
WNCR	494	4.74%	513	2.83%
WPAS	746	7.17%	966	5.32%
Total	**10411**		**18150**	

A comparative analysis using valid NHS Numbers highlights demographic inconsistencies across systems. Table 16 shows the highest mismatches in ‘first name’ (n=115,846) and ‘postcode’ (n=44,517), reflecting potential errors and variability in data entry. ‘Birth date’ and ‘death date’ exhibited fewer mismatches, indicating a higher level of consistency in these fields. Overall, mismatches across all fields exceed matches, underscoring widespread data misalignment.

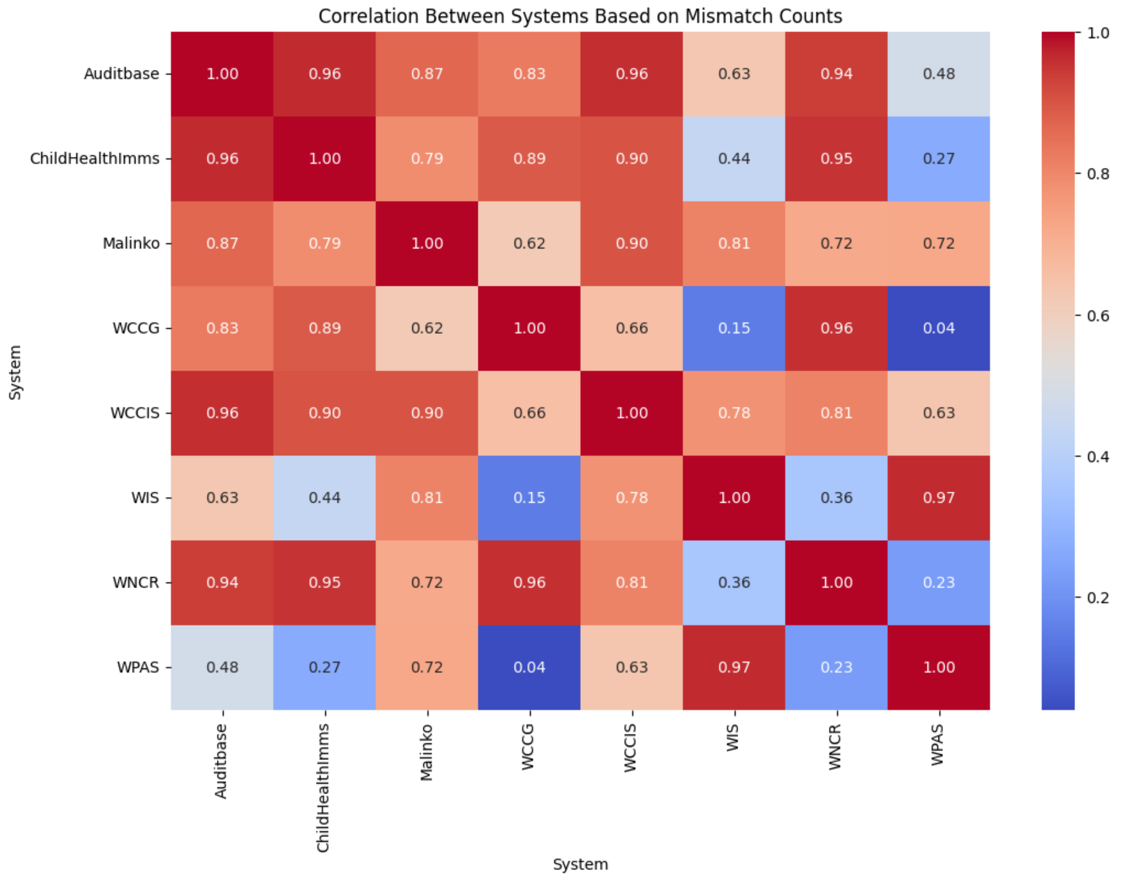
Table 16. Number of Unique Patient Records per Field with System Matches and Mismatches

System Count		First Name	Last Name	Birth Date	Death Date	Registered Practice	Postcode	Total
Mismatch	1	73864	7300	757	551	12595	19141	**114208**
	2	40355	3040	284	1492	12525	22506	**80202**
	3	1590	150	2	129	1952	2705	**6528**
	4	35	5			75	158	**273**
	5	2				1	7	**10**
	Total	**115846**	**10495**	**1043**	**2172**	**27148**	**44517**	**201221**
Match	**1**		61	7		35	87	**404**
	4	39238	3040	141	523	8961	14917	**66820**
	6	25947	1609	155	38	3251	4782	**35782**
	3	8917	877	179	65	2001	2226	**14265**
	2	2327	211	93	27	645	531	**3834**
	7	562	21	5	8	147	103	**846**
	5	106	2			22	2	**132**
	Total	**77311**	**5821**	**580**	**661**	**15062**	**22648**	**122083**

Plotting mismatch counts per field and system pinpoints WPAS as a primary source of First Name mismatches (Figure 13), which could be explained by spelling variations or the inclusion of middle names within this field. Postcode mismatches are especially high within WPAS (n=20,317) and WCCIS (n=22,972), reflecting potential address changes that are not synchronised across systems.



The Correlation Matrix (Figure 14) and Hierarchical Cluster Dendrogram (Figure 15) further illustrate relationships between mismatches across systems (p.60). The pairwise correlation (Figure 14) reveals WPAS and WIS have the strongest correlation (r=0.97), indicating high mismatch alignment, while WCCG and WPAS had the weakest (r=0.4). Clustering via Ward's method (Figure 15) grouped most systems closely, except WPAS and WIS, which formed a distinct cluster, suggesting that they have more unique data discrepancies compared to the other systems. These results point to systemic inconsistencies in demographic data, which can undermine confidence and trust in each system's data, supporting the need for improved validation processes to enhance data accuracy and reliability.



	Auditbase	ChildHealthImms	Malinko	WCCG	WCCIS	WIS	WNCR	WPAS
Auditbase	1.00	0.96	0.87	0.83	0.96	0.63	0.94	0.48
ChildHealthImms	0.96	1.00	0.79	0.89	0.90	0.44	0.95	0.27
Malinko	0.87	0.79	1.00	0.62	0.90	0.81	0.72	0.72
WCCG	0.83	0.89	0.62	1.00	0.66	0.15	0.96	0.04
WCCIS	0.96	0.90	0.90	0.66	1.00	0.78	0.81	0.63
WIS	0.63	0.44	0.81	0.15	0.78	1.00	0.36	0.97
WNCR	0.94	0.95	0.72	0.96	0.81	0.36	1.00	0.23
WPAS	0.48	0.27	0.72	0.04	0.63	0.97	0.23	1.00

Figure 14. Correlation Matrix Between Systems Based on Mismatch Counts

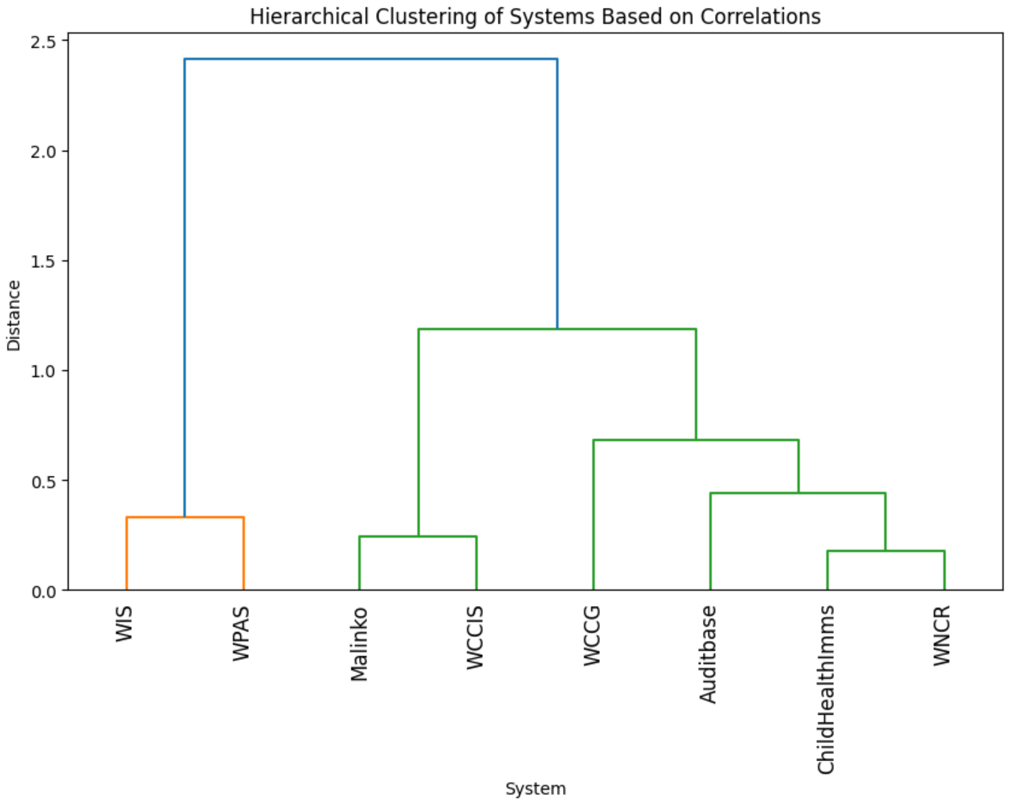


Figure 15. Dendrogram for Hierarchical Clustering of System Based on Mismatch Correlations

Respondents stressed that data reliability depends on consistent, accurate staff input. As SMC-MH26 commented ‘Data input methods are often misinterpreted, and this results in inconsistent data across the various teams pan Powys’, and A-O43 noted, ‘The data is only as reliable as the people entering the data in the first place’. While data-cleansing can improve data quality, efforts should include training and standardising data entry practices from the outset, with respondents commenting that ‘without a decent level of quality, consistency, and adherence to standards at the point of entry, cleansing can only go so far while preserving accuracy’ (DT31), and that ‘data quality starts at data entry.’ (DT49). DT46 reinforces this, commenting that ‘better application systems and better training for admin staff who input data are essential’.

Other suggestions to help address these challenges included the ‘Establishment of data quality groups’ (DT5), the use of ‘automated data quality checks’ (DT73), and investment in advanced technologies like AI, as A-O85 notes, ‘I think AI has a role in supporting with data quality and automation - I'm not sure if we've invested enough in this right now’. These insights reinforce the need for a dual approach: leveraging technology and equipping staff with the training and resources necessary to ensure accuracy at the point of data entry. This naturally leads to the third research objective, which explores other barriers and enablers to improving data flows and interoperability.

Research Objective 3. What factors are perceived as barriers or enablers to improving data flows and interoperability?

The third research objective identified several key barriers to improving data flows and interoperability within healthcare systems from participant perspectives (Table 17).

Table 17. Thematic Analysis: Barriers to Interoperability

Theme	Category	Representative Quotes
Barriers to Interoperability	**Non-interoperable/Siloed Systems** *(Further evidenced by quotes in other themes)*	"Different components of relevant information are within different systems, without a unifying front end." (C-M8) ‘Interoperability, Wi-Fi issues, Downtime of systems’ (DT58) "Systems don’t talk to each other. We find it difficult to follow a complete patient pathway." (A-PC9) "They dont talk to each other and all the data is not in one place or centralised" (SMA-O38)
	Lack of Data Standards/Consistency	"Interoperability of systems, variations in data collection, and poor data formatting are ongoing challenges." (SMA-O15) “Linking disparate systems without common data items is a challenge! For example, systems that use different patient identifiers that have no match between the two.” (DT31) “Our systems do not comply to the same data standards so we are unable to reliably integrate the information from different systems into an overarching picture of the healthcare provision within the health board" (DT46) “Lack of standards throughout the systems and technology limitations of the source systems we use” (DT66)
	Data Sharing/Information Governance Restrictions	"Data sharing agreements often aren't in place to allow for timely and accurate reporting between organisations." (DT1) "It is a challenge to get NHS England to share data." (DT2) "whilst IG is extremely important, particularly in relation to 3rd party orgs trying to gain access to information. It can be an unnecessary barrier to health boards trying to access their own data held by central teams in DHCW...Some of this is down to IG—not all—which, whilst important, shouldn't hinder us from acting in the best interest of our patients as it currently does’" (DT3) "IG policies often delay access to information and can be conflicting between organisations. That noted they are important for ensuring the data is shared appropriately and we legitimately need access." (DT35) "Difficulty accessing cancer pathway data between England and Wales." (SMA-PC84) “lengthy DPA and DPIA processes and sign off from other health boards impacted data access/sharing. difficulties/delays in capturing PTHB data governance within documents to be shared with wider HBs.” (A-MH54) "Security issues block our ability to set up data sharing channels that would improve access to data, particularly for commissioned services and primary care.” (DT46)
	Organisational Culture/Digital Literacy	"Getting people to understand the importance of data and getting them to care about data quality are a big challenge." (DT6) "Staff recording patient information may not always know why it is being collected, and don't see the value of or don't know how to record something accurately… Can be a challenge to understand what the data is telling us, sometimes it’s not clear what is being collected" (SMC-PH17) "We don’t really have the skills within our team. We all use data in medicines management but have not had formal training…We would really benefit from 'data consultant'-type sessions—someone to assess our data needs and show us how to improve/speed up processes." (C-M57) "Too many 'bad habits' exist so when we try and ensure things are done properly it is often met with resistance" (DT66) "struggle to find data never had trainin" (C-SN63) "There is a lack of knowledge re governance/privacy within teams who need access and those granting it. … We are sometimes unsure about what we are allowed to do with the data we do have access to" (DT61) "As a user of data it is challenging to learn what is available, where it is, and how to access it." (SMA-O87) "The biggest issue locally is the lack of real-time data input, which results in a lack of real-time information. This stems from asking staff with minimal data literacy skills and no clinical responsibility to input data into systems that are not user-friendly or easy to access." (DT89)

The primary technical barrier to interoperability stems from the presence of multiple outdated systems that fail to support seamless data flows between disparate systems. DT66 suggests that the ‘Lack of standards throughout the systems and technology limitations of the source systems we use’ contribute to unreliable and inefficient integration efforts. DT31 also comments that ‘Linking disparate systems without common data items is a challenge! For example, systems that use different patient identifiers that have no match between the two’.

The adoption of modern technologies and standardisation of data formats is seen as an enabler of real-time data sharing and seamless interoperability between systems (Table 18, p.64):

> Systems need to talk to each other, data sharing agreements need to be in place and to be shared in real time. Data needs to conform to specific standards so this can happen. (DT3)
>
> Central repositories that are easily accessible by organisations using modern technologies (FHIR API etc) would allow real time data to be surfaced to the clinician in a meaningful way so that the best decision can be made for the individual being treated. (DT62)

Broader macro-environmental barriers related to the governance of data access were noted as obstacles to retrieving necessary data, including complex data-sharing agreements, security constraints, and bureaucratic hurdles imposed by both central authorities and system vendors (Table 17, p.62):

> Various reporting requirements we have been unable to meet because of lack of data, constant ongoing issues and delays trying to gain access to everything needed from DHCW[41]…Barriers put up either centrally or by certain system vendors make it really difficult or impossible to obtain all necessary data needed for various reporting purposes, despite knowing it is captured. (DT3)

Information governance (IG) concerns were seen as both a necessary safeguard and a barrier to improving data access and interoperability. While respondents acknowledged the importance of robust IG policies, many felt that governance processes disproportionately hindered health boards from accessing data, as DT3 noted, 'whilst IG is extremely important, particularly in relation to 3rd party orgs trying to gain access to information it can be an unnecessary barrier to health boards trying to access their own data held by central teams in DHCW'. When discussing broader interoperability challenges, the importance of adhering to IG and respecting data privacy was reinforced by other respondents:

> There is a balance between patient safety and data protection that needs to be considered - governance policies should not hinder patient care unnecessarily. (DT10)

Stronger alignment of governance structures was recognised as an enabler to improving data flows, linking to the perceived lack of alignment between organisations in terms of both data sharing and data collection.

---

[41] Digital Health and Care Wales (DHCW) are a Special Health Authority, tasked with creating the digital solutions needed to improve health and care in Wales. (Welsh Government, 2021)

Table 18. Thematic Analysis: Enablers to Interoperability

Theme	Category	Representative Quotes
Enablers to Interoperability	**Data Standards**	"Systems need to talk to each other, data sharing agreements need to be in place and to be shared in real time. Data needs to conform to specific standards so this can happen." (DT3) "Data standards across health boards, all using the same systems to record information, more checks at the point of submission. We need more digital solutions so that don't have to chase individuals for excel files containing the data we need. " (DT61) "Standardise data entry; Improve training of data entry" (AH-M45)
	Technical Expertise/Digital Staffing	"Interoperability and staffing for digital teams. Ideally, a nominated IT/Digital business partner would help to ease as a nominated point of contact" (CMC-T13) "More investment in the team as I feel they are often swamped with conflicting prioritise - More investment in the data Intelligence and Information teams and more alignment with transformation" (SMA-O38)
	Organisational Change and Investment	"Organisational buy in and investment. This area of work is key and rarely gets the recognition of funding it needs." (DT66) "It will require senior leadership to to take the issues more seriously. We have excellent digital tools available to us now, but the way data is collected and distributed has changed little over the years." (DT86) "Improve communication. Auditing data entry regularly (weekly/monthly/quarterly). Hold people accountable for their data entry and highlight further training requirements" (A-MH45) "More funding and resource specifically geared to primary and community care data (as the managers of 90% of all NHS activity the resources should reflect this). shift the emphases from reactive to proactive this would ensure meaningful data from primary and community care could inform better service delivery and outcomes thereby reducing the reliance on emergency and planned care. " (A-PC36)
	Standard Processes and Data Sharing Agreements Neeeded Across Organisations	"Various reporting requirements we have been unable to meet because of lack of data, constant ongoing issues and delays trying to gain access to everything needed from DHCW…Barriers put up either centrally or by certain system vendors make it really difficult or impossible to obtain all necessary data needed for various reporting purposes, despite knowing it is captured."(DT5) 'There is a balance between patient safety and data protection that needs to be considered—governance policies should not hinder patient care unnecessarily.' (DT10) "Individual commissioning of system contracts means data guardianship is fragmented, making joined-up work between different systems difficult but aligned systems is hit by a quagmire of IG barriers." (A-PC36) "Need for better data sharing agreements between HBs" (SMA-PC84)
	Training	"Training of understanding and implementing interoperability is a big factor in my opinion. With time and more awareness of the positive impact interoperability has, interoperable systems will hopefully be more widely adopted." (DT5) "Education of existing programmes, Blanket adoption of the systems in place, policing of people and departments to ensure compliance with new systems rather than reverting to older less efficient methods." (A-T30) "More training and awareness for staff as to the possible" (A-O32) "Data literacy skills and understanding of operational/clinical staff on the use and impact of data. i.e. if they know how its used, what decisions are made and how its affected if data is incorrect timeliness and accuracy improves."(DT35) "I would love to have the capacity to truly engage and understand the data and how to access." (C-T51)
	Interoperable Systems *(Further evidenced by quotes in other themes)*	"Resource needs to be allocated to better systems that work interconnectivity to achieve better data hygiene and efficiency. " (A-T30) "Systems that talk to each other so data is transferred across the platforms. At the moment the medical records are across multiple systems and all users access (and use) these differently. A clear understanding of what is needed to be complaint when assessing clinical systems that will be used in local services is required." (C-T40) "As nice as having one system that can do everything sounds, I think in the past we have wasted too much time chasing this. It is very unrealistic when you work on a wide programme of work carried out by multiple different organisations that one system can provide everything all partners need, so I think sometimes it's useful to recognise that, understand what each system provides, do your best to create interoperability between them, but not every system needs to be able to do everything for everyone." (DT42)
	User Friendly Reports	"Reporting needs to be easily accessible, intuitive to use, and to be informed by multiple record systems across multiple organisations. Not easily achievable, where systems don't talk to each other. One obvious solution is either single integrated systems across multiple organisations, or interoperability and single reporting systems. Either way, greater investment is then needed in the data and intelligence teams to help construct easily accessible reporting." (SMC-O19) Although our skills have improved over the years, it would be more beneficial to have data presented easily for clinicians to use and not have to spend time crunching data. (C-M57)

Another category that emerged was the lack of training in data handling, reporting, and interpretation. The need for better training and-digital literacy was supported, as highlighted in the quotes:

> The biggest issue locally is the lack of real-time data input, which results in a lack of real-time information. This stems from asking staff with minimal data literacy skills and no clinical responsibility to input data into systems that are not user-friendly or easy to access. (DT89)
>
> Training of understanding and implementing interoperability is a big factor in my opinion. With time and more awareness of the positive impact interoperability has, interoperable systems will hopefully be more widely adopted. (DT5)

Some participants suggested educating service areas on the value of data, for example, SMC-PH17 comments that staff 'may not always know why it is being collected, and don't see the value of or don't know how to record something accurately', noting there are challenges to 'understand what the data is telling us, sometimes it's not clear what is being collected'. Similarly, DT6 remarks that 'getting people to understand the importance of data and getting them to care about data quality are a big challenge'.

Participants also indicate that a change in organisational culture, increased collaboration and learning opportunities could help address existing barriers:

> Invest on interoperable systems and efficient data flows. Break the cultural barrier and the elimination of data silos. Instruct the service areas about the relevance/power of the data, and to become more open (less reluctant on sharing it) -> [*sic*] explain the potential of interoperability/ sharing the data. (SMA-O7)

As DT86 stated 'It will require senior leadership to take the issues more seriously. We have excellent digital tools available to us now, but the way data is collected and distributed has changed little over the years'.

The value of collaboration was also noted, with DT69 emphasising the benefits of cross-Health Board cooperation: 'the willingness of Devs, Data Engineers, and Data Analysts across all Health Boards to share their skills, knowledge, and even products is a huge benefit'. Another participant reflected on improvements in data reliability since working with digital teams:

> Building relationships with the data and business team has been transformative. In the past, we 'lost' clients off our waiting list when we added in a first appointment, but this has been resolved by working with the IT team to develop our coding. Waiting list data was also inaccurate and a 'best guess' for part of our service, but again, collaborative

working has improved this. (A-MH56).

The need for organisational buy in, along with greater funding and investment in digital teams and technology were also raised by participants. DT66 remarked ‘This area of work is key and rarely gets the recognition or funding it needs’. A-PC36 called for ‘More funding and resource specifically geared to primary and community care data’, while SMC-T13 recommended investment in ‘Interoperability and staffing for digital teams’. This was echoed by A-O77 who says that ‘as an organisation we should be investing in workforce who can support services to move over to and improve on data flow optimisation’.

Figure 16 illustrates key factors influencing the adoption of data flows and interoperability. Questionnaire respondents most frequently selected ‘using multiple systems’ (n=65), but ‘work capacity (including time and opportunity)’ (n=63), ‘funding’ (n=63) and ‘resistance to change’ (n=61) were also seen as significant factors. Gaps in digital literacy and system capabilities were related in the selection of ‘digital expertise’ (n=58) and ‘existing technology’ (n=48). Similarly, ‘manual/paper processes’ (n=48) persist as a challenge with reliance on outdated workflows, slowing digital transformation efforts. ‘Nationally aligned directives and policies’ (n=34) and ‘organisational commitment’ (n=37) were less frequently selected but still present as barriers, indicating that alignment with national strategies and internal prioritisation could play a role in adoption success.

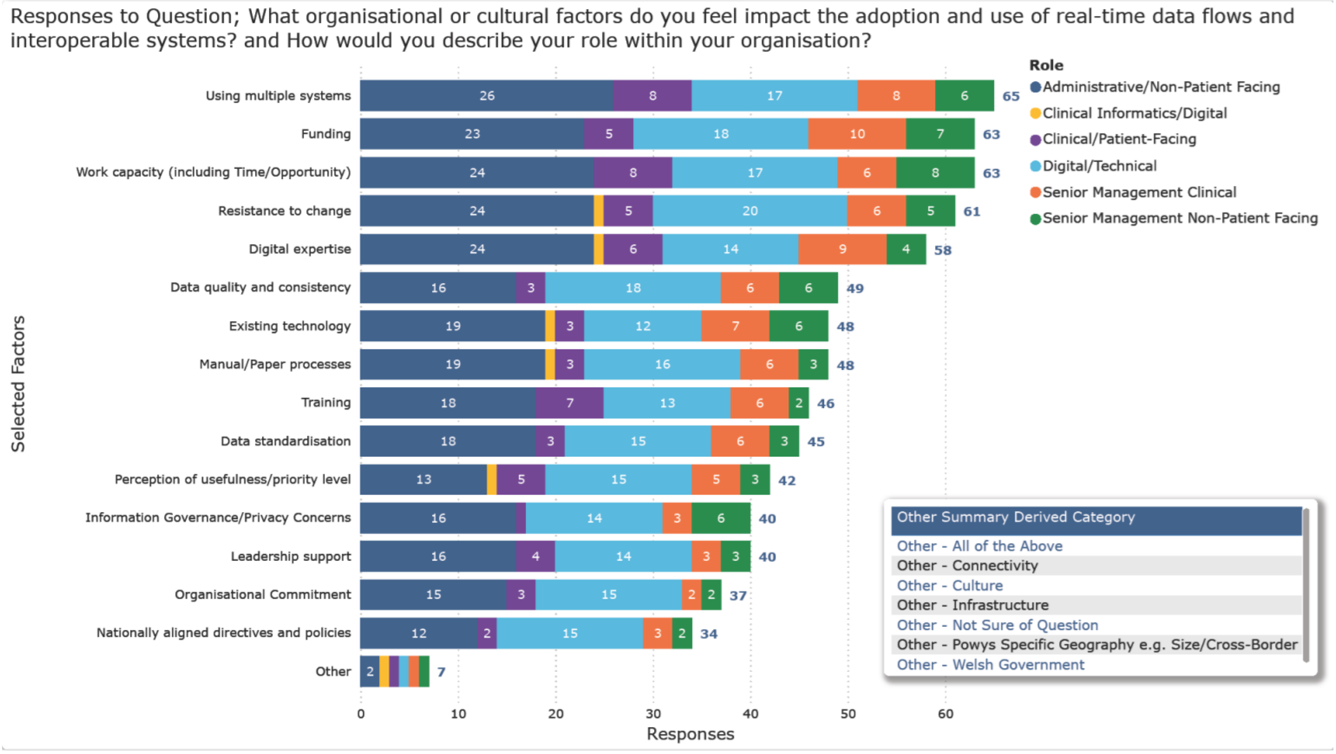


Figure 16. Selected factors impacting the adoption and use of real-time data flows and interoperable systems by questionnaire participant role type

Conclusion

This chapter presented quantitative and qualitative findings from the online questionnaire and documentary sources. Addressing each research objective, the analysis identified key challenges and opportunities surrounding data flows and interoperability. While respondents broadly recognised the potential benefits of improved interoperability, many felt that current limitations undermine confidence in data-driven decision-making, with fragmented data flows resulting in incomplete and unreliable information. Inefficiencies appear to be caused by outdated IT systems, limited interoperability, and limited data sharing. Despite these challenges, interoperability was recognised as essential for improving job performance and patient care, with both technological and cultural advancements facilitating meaningful change. The next chapter will discuss these findings in a broader context, before drawing final conclusions.

Chapter 5: Discussion, Conclusions and Implications

Introduction

Through a mixed-methods approach, the research sought to identify challenges and opportunities associated with healthcare data flows and interoperability, and to understand their perceived impact within healthcare. The findings in the previous chapter indicate the presence of non-interoperable systems, a reliance on manual processes, inconsistent data entry practices, and issues with data quality. This concluding chapter interprets these findings in relation to the main research aim ‘What factors influence healthcare data interoperability and the utilisation of real-time data flows, and how effectively do they support healthcare delivery within the NHS?’. The study concludes with a discussion of its strengths and limitations, findings in relation to existing literature, final reflections and implications.

Meeting of aim and objective

The study aimed to explore how effectively data flows and interoperability support healthcare delivery within the NHS and identify influencing factors and guided by three sub-questions:

- *How do healthcare professionals perceive the interoperability of the systems they use, and what impact does this have on their work?*
  Thematic analysis identified widespread participant frustration with fragmented systems, delays in accessing information, and the burden of duplicate data entry, subsequently impacting daily workflows, collaboration, and ultimately patient care. Quantitative data supported these concerns, evidencing duplicated activity and manual processes.
- *In what ways do data quality and accessibility influence decision-making and patient outcomes?*
  Survey responses indicated that unreliable or missing data hinders strategic planning and clinical decision-making. Quantitative findings confirmed systemic data quality issues across EHRs, compromising the accuracy and completeness of information, particularly patient demographics.
- *What factors are perceived as barriers or enablers to improving data flows and interoperability?*
  Barriers identified from both open and closed survey questions included outdated digital infrastructure, governance restrictions, and inconsistent data entry practices. Enablers included training, collaboration and data sharing, data standards, and investment in digital capabilities. These aligned closely with themes identified in the literature review (p.19), demonstrating a common frustration with multiple systems that do not ‘talk’ to each other.

In relation to the main research aim, although the study could not empirically evaluate the effectiveness of interoperable systems due to limited evidence of implementation, findings revealed positive perceptions of the potential to enhance decision-making, patient safety, and operational efficiency.

Significance of the study:

With a growing global emphasis on real-time data sharing and interoperability, optimising data flows may help improve patient outcomes and operational efficiency. This study offers a realistic, local assessment of the current state, with documentary evidence and insights from participants (including digital and technical professionals, seemingly underrepresented in existing research) providing a nuanced perspective on the challenges and opportunities in achieving interoperability.

For Powys, a rural and expansive region where healthcare is predominantly commissioned (Audit Wales, 2018), streamlining data management and enhancing real-time information exchange could be particularly vital, requiring more cohesive and efficient systems. PTHB's unique interoperability needs have been largely unexplored in previous literature and represent an area worthy of investigation.

Participant frustrations highlight areas requiring improvement that could inform digital strategy within NHS Wales. More broadly, this research supports national transformation efforts, which arguably have been slow and inconsistent, and helps clarify the potential technical and cultural requirements for a more connected, data-driven, and patient-centred health system.

Strengths and limitations

A methodological limitation was the restricted qualitative scope, with qualitative data gathered solely through a questionnaire. The data collection process was constrained by a limited timeframe, no funding or participant incentives. The relatively low clinical response rate may be attributed to various factors including workload pressures or a perception that data use is not central to their role. For example, one clinician suggested rephrasing 'data' as 'information' might have increased participation, while another declined involvement, stating they don't 'use data' within their role. However, it could be argued that data underpins clinical practice, and is the foundation of the *Data, Information, Knowledge, Wisdom (DIKW) Pyramid* (Ackoff, 1989). Consequently, most responses came from individuals in digital and technical roles, who may be more engaged with and receptive to the potential of real-time data flows and interoperable systems.

A leading strength of this study is its mixed-methods design, which provided a deeper understanding of interoperability in practice by integrating qualitative insights from healthcare professionals with quantitative evidence. The research triangulated evidence of data duplication, manual processes, and data quality issues, challenges that the reviewed literature had primarily reported through qualitative data alone. One prominent example is the demographic discrepancies between systems, which creates uncertainty for staff in determining which patient record and associated information is correct. By combining quantitative and qualitative insights from those working directly with healthcare systems the study's findings are problem-driven rather than theory or technology-driven. This convergence of method aims to strengthen the validity, applicability and comprehensiveness of the recommendations, balancing technological advancement with operational need.

Kelly and Cordeiro's (2020) third pragmatic principle, inquiry as an experiential process, could have been better addressed through semi-structured interviews or focus groups, as an interactive approach could have allowed insights to emerge more organically than predefined survey questions. A follow-up qualitative phase could enhance understanding and validate findings, given that the qualitative analysis was conducted by a single researcher without independent verification of consistency or accuracy (Gale *et al.,* 2013; Friedman *et al.,* 2022). Word-limit constraints also limited the extent of discussion that could be included. Despite these limitations, the research provides a qualitative foundation for future studies. The research was conducted in line with a structured project plan (Appendix 15), which supported effective time management, maintained researcher momentum and contributed to the overall organisation and timely delivery of the study.

Discussion of Findings

The findings of this study highlight several aspects influencing data flows and interoperability within NHS Wales, particularly from the perspective of PTHB staff. Participants reported difficulties with data-sharing and system interoperability, which they believe negatively impact service efficiency, co-ordination, and decision-making. Respondents consistently reported fragmented data systems and concerns around accessing accurate, timely data. These issues are consistent with Deloitte's Global Healthcare Outlook (Siegel, 2024), and existing literature (Choun and Petre, 2022; Mistry *et al.,* 2022; Roehrs *et al.,* 2017) which identifies poor interoperability as a critical barrier to effective digital healthcare globally.

The results of this research can be summarised with the following CMO configuration: There are fragmented data flows and multiple non-interoperable systems hindering data-sharing between NHS organisations (*context*). This creates a challenging environment where manual workarounds are frequently required. Staff often engage in dual recording, cross-referencing data from multiple sources, and manually validating data (*mechanism*). These practices, while necessary in the absence of seamless systems, have considerable consequences for operational efficiency leading to staff frustration and capacity challenges. Missing data and data inconsistencies, in turn leads to delays in decision-making, and concerns about the reliability of patient records. (*outcome*).

Relating the findings back to the maturity levels defined within the HIMSS EMRAM (HIMSS, 2022), the lack of seamless data exchange between primary, secondary, and community care services, along with the absence of integrated CDSS, patient portal and remote monitoring data flows, hinder PTHB from advancing their EMRAM Stage. Arguably, the reliance that PTHB and other Welsh health boards have on DHCW to provide the strategic direction and technical solutions may have effectively constrained their ability to pursue independent solutions. As noted in the HIMSS (2023) EMRAM assessment of PTHB, assurance is needed that DHCW has the capacity and capability to deliver in these areas.

Participants identified a lack of data-sharing agreements and missing information as barriers to interoperability, co-ordinated care and effective decision-making, impacting the delivery of patient care, especially when individuals receive treatment at another organisation. While governance frameworks to protect sensitive health information were supported, participants emphasised that these frameworks should facilitate, rather than hinder, data sharing and data-driven improvements. The challenge, therefore, lies in balancing patient privacy with the need for comprehensive data exchange across organisations to support holistic care. Hulsen (2020) highlights growing support for data sharing and suggests the lack of consensus around data ownership and the complexity of privacy laws could be addressed through non-technical means, such as standardised consent processes, or with technical solutions such as FDPs.

In addition to governance and data ownership challenges, geographical and administrative boundaries within health systems have been said to exacerbate data silos (Lau *et al.,* 2024), leading to delays in treatment, misdiagnoses, and an inability to track long-term patient pathways (Birgelen*,* 2024). This aligns with concerns raised by Baird and Wright (2006) which link inequitable health care to rural regions, and challenges connecting Wales' dispersed rural population with urban services (Bauchinger *et al.,* 2021). Within a BBC article, Rees (2025) discusses practical challenges faced by healthcare professionals referring Welsh patients into England, further highlighting how interoperability gaps are exacerbated by differing policies and systems. These issues resonate with the perspectives of staff in this study, who describe how disjointed data flows and poor interoperability contribute to delays, miscommunication, and increase administrative burden through duplicated data entry across systems. Such fragmentation can be especially distressing for patients, and may disproportionately affect those with complex needs, who require coordinated care across services. Previous studies have shown that hospital transfers can result in adverse outcomes and negatively affect patient experiences (Goulding *et al.,* 2015; Mortensen *et al.,* 2019; Uhrenfeldt *et al.,* 2013), which is particularly relevant for PTHB patients where services are frequently commissioned from other organisations.

These challenges are amplified because incomplete data can lead to uncertainty, flawed decision-making and bias (Ayilara e*t al.,* 2019; Helou *et al.,* 2020; Liu *et al.,* 2023). Data quality, as demonstrated by Khong *et al.* (2023) can significantly impact the accuracy, reliability, and timeliness of decisions. These issues are consistent with the findings of this study, where data inconsistencies between systems are reported to hinder the accuracy of patient information, contributing to confusion and poor decision-making. Additionally, poor data quality has been shown to limit the effectiveness of AI and ML applications (Vollmer *et al.,* 2020). As outlined within the literature review, ensuring high-quality, interoperable data is essential for minimising uncertainty, improving decision-making, and ensuring patient safety (Banerjee *et al.,* 2018; Khanbhai *et al.,* 2021).

Despite NHS-wide commitments to interoperability through initiatives such as the NDR, respondents identified the absence of standards, governance, and training gaps as persistent issues, illustrating that the challenges extend beyond technological solutions. This supports the broader critique that technological advancements alone are insufficient, cultural, regulatory, and workforce factors must also be addressed (Camacho *et al.,* 2024). The importance of a digitally literate workforce is widely supported in the literature and considered a contributing factor to achieving better understanding and application of data-driven insights (Chen and Banerjee, 2020; Davidson *et al.,* 2022; Goldacre *et al.,* 2020; Khanbhai *et al.,* 2021; Jager *et al.,* 2023; Wain, 2021). Participants in this study recognised a need for formal training to improve data handling skills and suggested that a better understanding of data's significance would enhance the quality of data entry. Poor digital literacy has long been identified as a barrier to the successful implementation of HIT, and employing a skilled workforce to develop, implement, and evaluate HIT applications is considered equally critical (Hersh and Wright, 2008). This implies that digital training and the recruitment of technical staff could bridge the gap between healthcare needs and digital solutions, an approach also recommended by study participants.

Digital leadership and collaborative communication across healthcare sectors are thought to be essential in driving change (Feely *et al.,* 2023, Schwarz *et al.,* 2020). By communicating the benefits of interoperability and reinforcing change through role-modelling, leaders may be able to drive staff engagement. Without such involvement, interoperability initiatives may fail to address the practical needs of healthcare professionals, leading to resistance or suboptimal use of digital systems. Research has previously shown that user-centred design and frontline engagement boost motivation and reduce resistance to change (Kernebeck *et al.,* 2022).

Prior to the Triple Aim framework, improving care experience and reducing costs were often seen as opposing goals, but it is now recognised that these aims can reinforce one another (Nundy *et al.,* 2022). Arguably, achieving improvements first requires financial investment in interoperable systems and training, which necessitates leadership recognition and adjustments to NHS budget allocation. Positively, this is reflected in the 2025/26 draft NHS Wales budget, which has doubled funding for digital projects and prioritised the establishment of a national digital architecture for consistent and efficient system-wide integration (Welsh Parliament, 2025, p.25). However, as Cresswell *et al.* (2021) highlight, funding allocation often favours well-resourced hospitals, leaving less digitally mature institutions struggling with leadership, capacity building, and funding constraints. This is particularly relevant for smaller health boards such as PTHB, which receives comparatively lower levels of Welsh Government funding (Appendix 14). As a result, such organisations may lack financial resources and face greater challenges in independently investing in the infrastructure or digital expertise needed for interoperable systems. This potential imbalance in digital governance could create a power disparity, where well-funded institutions may advance leaving other organisations behind. It also raises ethical concerns, as the lack of equitable access to digital tools could exacerbate healthcare inequalities. Ensuring inclusive digital transformation strategies could be critical to mitigating these risks and support a more integrated and efficient healthcare system. Ideally, robust regulatory frameworks and ethical guidelines protecting patient data, should also enable seamless data-sharing practices that do not disadvantage particular patient populations.

The final section surmises the implications of this study drawing together recommendations including areas for future research.

Implications for Policy, Practice, Education and Research

Table 19. Study Recommendations

Domain	Key Recommendations
Policy	**System Architecture & Infrastructure**
	- Modernise legacy systems and adopt scalable, future-proof architectures
	- Implement a structured, clinician-informed system procurement/migration plan
	- Adopt interoperable data standards (e.g. FHIR)
	Investment & Resource Allocation
	- Allocate funding to expand digital team capacity
	- Prioritise investment in interoperability initiatives, addressing current manual processes e.g. dual recording
	Organisational Collaboration
	- Establish clear interoperabile reqirements and data-sharing policies
	- Create NHS-wide collaboration frameworks
	- Implement standardised health & social care data-sharing agreements
Practice	- Integrate digital literacy training with cultural transformation efforts
	- Mitigate resistance through leadership engagement, digital champions, and staff peer support
Education	- Embed digital literacy and data-quality awareness in healthcare curricula
	- Provide ongoing digital training for current healthcare professionals
Research	- Compare real-world outcomes of interoperable vs non-interoperable systems
	- Gather clinician and patient feedback to guide user-centred strategies
	- Explore AI/ML for automated data integration and quality monitoring
	- Map automated vs manual data workflows to identify inefficiencies

Interoperable systems and real-time data flows offer significant potential to improve patient outcomes, decision-making, and operational efficiency. Facilitating accurate, timely data integration access could reduce the burden of dual recording and help minimise data inconsistencies. To achieve this, the following policy, practice research and education recommendations should be considered, and are summarised in Table 15 (p.73)

Policy Recommendations

To create the conditions for technical advancement and sustainable healthcare infrastructure, it is recommended that national and organisational policies prioritise system modernisation, funding allocation, and NHS-wide collaboration. This study identifies the reliance on legacy systems as one of the most significant barriers to interoperability, contributing to operational inefficiencies. The adoption of modern scalable systems and interoperable data standards could facilitate consistent, accurate data exchange across platforms, alleviating the long-term costs associated with supporting multiple systems. It is also recommended that digital implementations be carried out in a structured, collaborative manner to mitigate risk of data loss, inconsistencies, or service disruptions, ensuring that technically capable systems are provisioned in line with operational needs.

It is recommended that sustained investment be directed towards both digital infrastructure and workforce capacity. Targeted funding could help build expertise within digital teams to effectively support growing data requirements, support staff training, and reduce the risk of under-resourced initiatives that stall transformation. It is suggested funding prioritise interoperability, particularly to reduce the administrative burden of dual recording and workarounds, which could improve data quality and system efficiency. This should include allocating clinicians the time, resource, and support necessary to record and use data efficiently to support long-term digital adoption and hopefully foster a culture of data-driven decision-making.

Arguably, health board dependency on DHCW may have limited local proactive innovation, contributing to delays and inefficiencies in achieving seamless data exchange and optimal patient care. It is therefore recommended that greater collaboration across healthcare sectors and cross-sector frameworks encourage knowledge sharing, helping to accelerate progress, promote innovation, and strengthen partnerships across NHS organisations. Policies supporting secure data sharing and defining interoperability requirements could foster organisational alignment, build confidence, and enable informed procurement decisions. Implementing standardised data-sharing agreements across health and social care is also recommended, ensuring clinicians have access to the necessary patient information to enhance care continuity, and facilitate effective use of advanced analytics.

Practice Recommendations

Practical recommendations at the organisational level could support policy recommendations and help embed interoperability into routine operations. Promoting digital literacy through training, alongside cultural change initiatives that emphasise the value of data, may strengthen confidence in digital tools and foster lasting, data-driven practices in healthcare. Leadership engagement and visible support for digital transformation could help reduce resistance to change. Recommended approaches include appointing digital champions and encouraging collaboration between clinical and digital teams to build trust in new systems and improve adoption. These actions may also contribute to a more unified, data-driven organisational culture

Education Recommendations

As touched upon in the policy and practice recommendations, education and training are likely requirements for successful adoption and effective use of digital health tools. Embedding digital literacy and data quality into healthcare education programmes could help prepare future professionals to operate in increasingly digital environments. For the existing workforce, ongoing training may increase professionals' confidence in using new technologies, promote consistent data entry practices, and improve the overall quality and utility of recorded data.

Research Recommendations

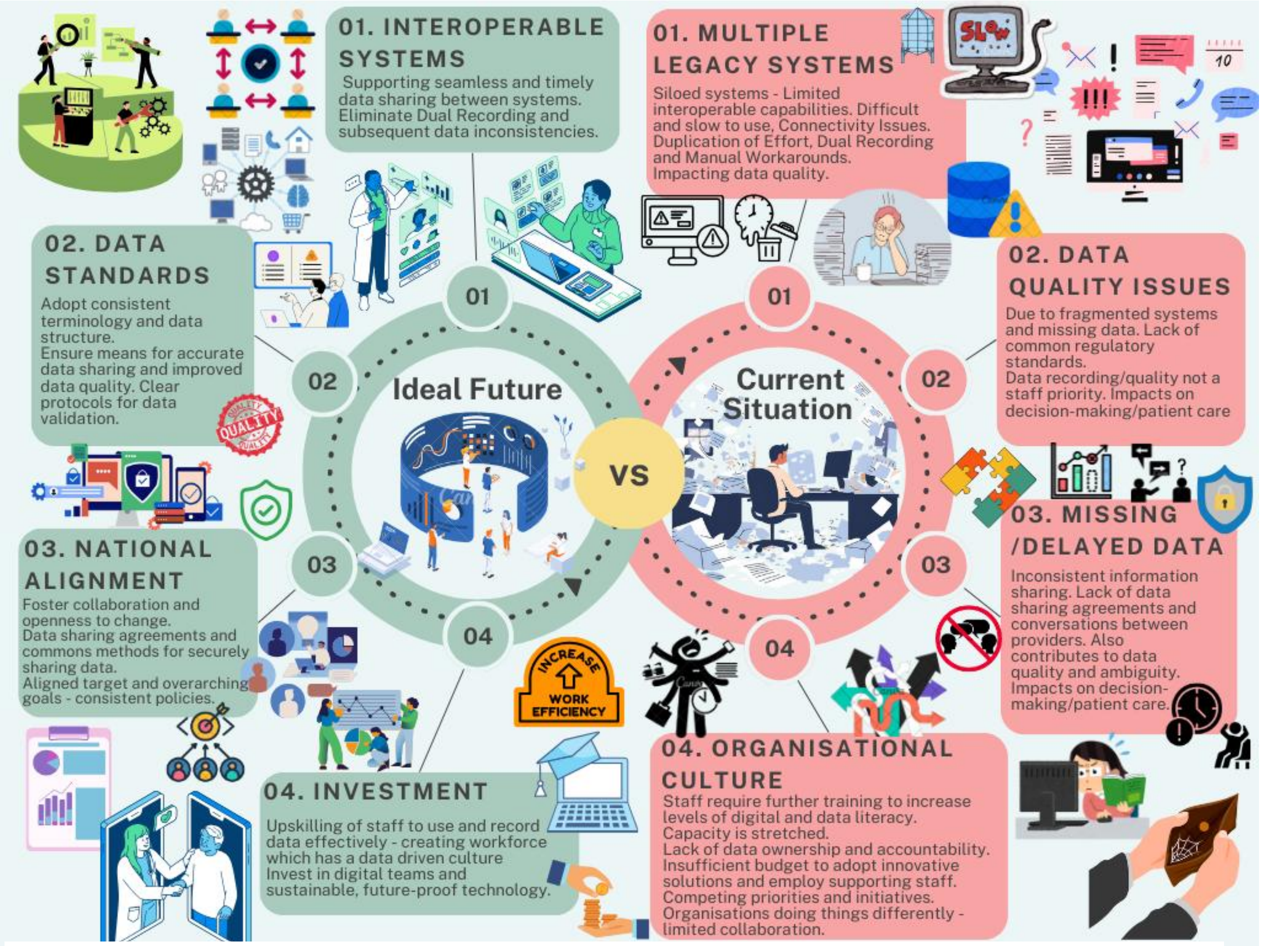
Although this study stresses some key challenges and opportunities, the question around how interoperable systems and real-time data flows support NHS delivery remain largely unanswered. Camacho *et al.* (2024) also suggest that further investigation into the effectiveness of interoperable standards compared to non-interoperable scenarios would be valuable within the NHS context. This type of research could provide empirical evidence on the potential real-world impact interoperable systems have, supporting future investments and policy decisions. A financial assessment is also recommended to understand implementation costs relative to long-term savings from efficiency gains.

A comprehensive review of NHS systems is recommended to assess current interoperability, system capability, and the true extent of FHIR adoption. This may identify where interoperability efforts are succeeding and where gaps remain, supporting targeted strategies to enhance data sharing and alignment across systems. Comparative analysis between regions may reveal best practices and scalability potential, including implementation successes and challenges.

As discussed within this study's limitations, further qualitative research into data exchange involving clinicians and patients, could help inform future ethical strategies aligned with clinical need. Another area which may benefit from future research is an evaluation of the proportion of automated versus manual data processing occurring across NHS organisations. This assessment would help identify inefficiencies, highlight areas for improvement, and provide insights into the impact of automation on data accuracy and operational efficiency. There is also the potential for research to explore how AI-driven solutions can automate data integration, identify inconsistencies, and improve data quality.

Concluding Remarks

The insights and recommendations from this study have the potential to help guide future transformation toward a more digitally advanced, data-driven healthcare organisation. The rich picture in Figure 17 visually contrasts the identified challenges in healthcare data management with an ideal future, characterised by seamless real-time data sharing between health and care providers, reduced administrative burden, improved patient care coordination, predictive analytics for early intervention, and a digitally literate workforce that embraces technology for improving efficiency and patient outcomes.



01. INTEROPERABLE SYSTEMS
Supporting seamless and timely data sharing between systems. Eliminate Dual Recording and subsequent data inconsistencies.

02. DATA STANDARDS
Adopt consistent terminology and data structure.
Ensure means for accurate data sharing and improved data quality. Clear protocols for data validation.

03. NATIONAL ALIGNMENT
Foster collaboration and openness to change.
Data sharing agreements and commons methods for securely sharing data.
Aligned target and overarching goals - consistent policies.

04. INVESTMENT
Upskilling of staff to use and record data effectively - creating workforce which has a data driven culture
Invest in digital teams and sustainable, future-proof technology.

01. MULTIPLE LEGACY SYSTEMS
Siloed systems - Limited interoperable capabilities. Difficult and slow to use, Connectivity Issues. Duplication of Effort, Dual Recording and Manual Workarounds.
Impacting data quality.

02. DATA QUALITY ISSUES
Due to fragmented systems and missing data. Lack of common regulatory standards.
Data recording/quality not a staff priority. Impacts on decision-making/patient care

03. MISSING /DELAYED DATA
Inconsistent information sharing. Lack of data sharing agreements and conversations between providers. Also contributes to data quality and ambiguity. Impacts on decision-making/patient care.

04. ORGANISATIONAL CULTURE
Staff require further training to increase levels of digital and data literacy.
Capacity is stretched.
Lack of data ownership and accountability.
Insufficient budget to adopt innovative solutions and employ supporting staff.
Competing priorities and initiatives.
Organisations doing things differently - limited collaboration.

Figure 177. Healthcare Data Interoperability: Current Challenges vs. Ideal Future

Ultimately, the future of healthcare may depend on the willingness to embrace digital transformation; by prioritising interoperability and harnessing real-time data, organisations could drive innovation, setting a standard for NHS Wales and beyond.

References

Ackoff, R.L. (1989). ‘From data to wisdom’. *Journal of Applied Systems Analysis.* 16, pp.3-9 Available at: https://www.researchgate.net/profile/Hung-Chun-Huang/post/Original_paper_of_From_data_to_wisdom_by_Ackoff_1989/attachment/67c507474e041854a3da49ac/AS%3A11431281312886827%401740965703715/download/From+data+to+wisdom.pdf?_tp=eyJjb250ZXh0Ijp7ImZpcnN0UGFnZSI6InB1YmxpY2F0aW9uIiwicGFnZSI6InF1ZXN0aW9uIn19 [Accessed 31 March 2025].

Adaba, G. B., and Kebebew, Y. (2018). ‘Improving a health information system for real-time data entries: An action research project using socio-technical systems theory’. *Informatics for Health and Social Care*, 43(2), pp.159–171. DOI: https://doi.org/10.1080/17538157.2017.1290638

Aerts, A., and Bogdan-Martin, D. (2021). ‘Leveraging data and AI to deliver on the promise of digital health’. *International Journal of Medical Informatics*, 150, 104456. DOI: 10.1016/j.ijmedinf.2021.104456

Ahmed, S. K. (2025). ‘Sample size for saturation in qualitative research: Debates, definitions, and strategies’. *Journal of Medicine, Surgery, and Public Health*, *5*, 100171. DOI: https://doi.org/10.1016/j.glmedi.2024.100171

Alhmoud, B., Melley, D., Khan, N., Bonicci, T., Patel, R., and Banerjee, A. (2022). ‘Evaluating a novel, integrative dashboard for health professionals’ performance in managing deteriorating patients: a quality improvement project’. *BMJ Open Quality*, *11*(4). DOI: https://doi.org/10.1136/bmjoq-2022-002033

Armbrust, M., Ghodsi, A., Xin, R. and Zaharia, M. (2021). ‘Lakehouse: A New Generation of Open Platforms that Unify Data Warehousing and Advanced Analytics.’ In CIDR,*11th Conference on Innovative Data Systems Research, Virtual Event, January 11-15, 2021, Online Proceedings.* Available at: https://www.cidrdb.org/cidr2021/papers/cidr2021_paper17.pdf [Accessed 31 March 2025].

Atenas, J., Havemann, L., and Timmermann, C. (2023). ‘Reframing data ethics in research methods education: a pathway to critical data literacy’. *International Journal of Educational Technology in Higher Education*, 20(1). DOI: https://doi.org/10.1186/s41239-023-00380-y

Audit Wales (2020) *Welsh Community Care Information System Report of the Auditor General for Wales - October 2020* [online] Available at: https://www.audit.wales/sites/default/files/WCCIS-Eng_10.pdf [Accessed 9 February 2022].

Audit Wales (2018) *Primary Care Services – Powys Teaching Health Board* [online] Available at: https://www.audit.wales/sites/default/files/powys_primary_care_services_final_5.pdf [Accessed 1 May 2025].

Aveyard, H. (2023) *Doing a literature review in health and social care: a practical guide*. 5th ed. Jefferson: McGraw-Hill Education.

Ayaz, M., Pasha, M., Alzahrani, M., Budiarto, R., and Stiawan, D. (2021). 'The Fast Health Interoperability Resources (FHIR) Standard: Systematic Literature Review of Implementations, Applications, Challenges and Opportunities'. *JMIR medical informatics*, *9*(7), e21929.DOI: https://doi.org/10.2196/21929

Ayilara, O.F., Zhang, L., Sajobi, T.T., Sawatzky, R., Bohm, E., and Lix, L.M. (2019) 'Impact of missing data on bias and precision when estimating change in patient-reported outcomes from a clinical registry'. *Health and Quality of Life Outcomes,* 17, 106. DOI: https://doi.org/10.1186/s12955-019-1181-2

Baird, A.G. and Wright, N (2006) 'Poor access to care: rural health deprivation?' *British Journal of General Practice,* 56(529), pp.567 – 568.

Banerjee, A., Drumright, L.N., and Mitchell, A.R.J. (2018). 'Can the NHS be a learning healthcare system in the age of digital technology?' *BMJ Evidence-Based Medicine*, 23(5), pp.161–164. DOI: https://doi.org/10.1136/bmjebm-2018-110953

Basford, T. and Schaninger, B. (2016) 'The four building blocks of change'. *McKinsey Quarterly*, McKinsey & Company. Available at: https://www.mckinsey.com/business-functions/people-and-organizational-performance/our-insights/the-four-building-blocks--of-change [Accessed 04 November 2021].

Bauchinger, L., Reichenberger, A., Goodwin-Hawkins, B., Kobal, J., Hrabar, M., and Oedl-Wieser, T. (2021) 'Developing Sustainable and Flexible Rural-Urban Connectivity through Complementary Mobility Services'. *Sustainability* 13(3), 1280. DOI: https://doi.org/10.3390/su13031280

Beasant, L., Cullen, F., Thomas, E., Kandiyali, R., Shield, J.P.H., Mcgregor, D., West, N., and Ingram, J. (2023). 'Flash glucose monitoring in young people with type 1 diabetes - A qualitative study of young people, parents and health professionals: 'It makes life much easier'. *BMJ Open*, 13(4). DOI: https://doi.org/10.1136/bmjopen-2022-070477

Beauchamp, T.L, and Childress, J.F. (2013) *Principles of bioethics*. 7th edn. Oxford University Press.

Benge, C. L., Onwuegbuzie, A. J., and Robbins, M. E. (2012). 'A Model for Presenting Threats to Legitimation at the Planning and Interpretation Phases in the Quantitative, Qualitative, and Mixed Research Components of a Dissertation'. *International Journal of Education*, 4(4). DOI: https://doi.org/10.5296/ije.v4i4.2360

Beyer, M.A. and Laney, D. (2012) *The Importance of "Big Data": A Definition*. Gartner. Available at: https://www.gartner.com/doc/2057415 [Accessed: 20 October 2024]

Biesta, G. (2010). 'Pragmatism and the Philosophical Foundations of Mixed Methods research' in Tashakkori, A. and Teddlie, C. (eds), *SAGE Handbook of Mixed Methods in Social and Behavioral Research*, 2nd edn, pp.95-117. Thousand Oaks, CA: Sage.

Bird, M., McGillion, M., Chambers, E.M., Dix, J., Fajardo, C.J., Gilmour, M., Levesque, K., Lim, A., Mierdel, S., Ouellette, C., Polanski, A.N., Reaume, S.V., Whitmore, C., and Carter N. (2021) 'A generative co-design framework for healthcare innovation: development and application of an end-user engagement framework'. *Research Involvement and Engagement,* 7, 12. DOI: https://doi.org/10.1186/s40900-021-00252-7

Bird, S., Klein, E. and Loper, E. (2009), *Natural Language Processing with Python*: *analyzing text with the natural language toolkit*. O'Reilly Media Inc.

Birgelen, D. (2024) 'From data silos to seamless integration: How the cloud is reshaping healthcare data management'. *The Journal of mHealth*. Available at: https://thejournalofmhealth.com/from-data-silos-to-seamless-integration-how-the-cloud-is-reshaping-healthcare-data-management/ [Accessed: 20 March 2025].

Booth, A. (2004) 'Formulating answerable questions'. In A. Booth and A. Brice (eds), *Evidence-Based Practice for Information Professionals: A handbook.* London: Facet, pp.59-66.

Bowling, A. (2023). *Research Methods in Health,* 5th edn. Maidenhead: Open University Press.

Braunstein, M. (2022) *Health Informatics on FHIR: How HL7's API is Transforming Healthcare.* Cham, Switzerland. Springer International Publishing.

Brooke, J. (1995) 'SUS: A quick and dirty usability scale', *Usability Evaluation Industries.* Available at: https://www.researchgate.net/publication/228593520_SUS_A_quick_and_dirty_usability_scale [Accessed: 20 October 2024]

Brown, K.M., Elliott, S.J., Leatherdale, S.T. and Robertson-Wilson, J. (2015) 'Searching for rigour in the reporting of mixed methods population health research: a methodological review', *Health Education Research*, 30(6), pp.811–839, DOI: https://doi.org/10.1093/her/cyv046

Critical Appraisal Skills Programme *(2023). CASP Qualitative studies Checklist.* Available at: https://casp-uk.net/casp-tools-checklists/qualitative-studies-checklist/ [Accessed 4 November 2024].

Camacho, E.M., Gavan, S., Keers, R.N., Chuter, A., and Elliott, R.A. (2024). 'Estimating the impact on patient safety of enabling the digital transfer of patients' prescription information in the English NHS'. *BMJ Quality and Safety*. DOI: https://doi.org/10.1136/bmjqs-2023-016675

Catlow, J., Sharp, L., Wagnild, J., Lu, L., Bhardwaj-Gosling, R., Ogundimu, E., Kasim, A., Brookes, M., Lee, T., McCarthy, S., Gray, J., Sniehotta, F., Valori, R., Westwood, C., McNally, R., Ruwende, J., Sinclair, S., Deane, J., and Rutter, M. (2024). 'Nationally Automated Colonoscopy Performance Feedback Increases Polyp Detection: The NED APRIQOT Randomized Controlled Trial'. *Clinical Gastroenterology and Hepatology*, *22*(9), pp.1926–1936. DOI: https://doi.org/10.1016/j.cgh.2024.03.048

Chalmers, R. and Brannan, G.D. (2023) *Organizational Culture*. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; Available from: https://www.ncbi.nlm.nih.gov/books/NBK560543/

Chen, Y. and Banerjee, A. (2020) 'Improving the digital health of the workforce in the COVID-19 context: an opportunity to future-proof medical training'. *Future Healthcare Journal*, **7**(3), pp.189-192. DOI: 10.7861/fhj.2020-0162

Check, J., and Schutt, R.K. (2012) 'Survey research'. In: Check, J., and Schutt., R.K. (eds). *Research methods in education*. Thousand Oaks, CA: Sage Publications.

Choun, D., and Petre, A. (2022). *Digital Health and Patient Data: Empowering Patients in the Healthcare Ecosystem*. 1st ed. Productivity Press. DOI: https://doi.org/10.4324/9781003215868

Chu, A., Kumar, A., Depoorter, G., Franklin, B.D., and Mcleod, M. (2022). 'Learning from electronic prescribing errors: A mixed methods study of junior doctors' perceptions of training and individualised feedback data'. *BMJ Open*, *12*(12). DOI: https://doi.org/10.1136/bmjopen-2021-056221

Clark, T., Foster, L., Sloan, L. and Bryman, A. (2021). *Bryman's Social Research Methods*. 6th edn. Oxford: Oxford University Press.

Codeacademy (n.d) *Normalization*. Available at: https://www.codecademy.com/article/normalization [Accessed 4 April 2025]

Cohen, J. (1990). 'Things I have learned (so far)'. *American Psychologist,* 45(12), pp.1304–1312. DOI: https://doi.org/10.1037/0003-066X.45.12.1304

Creswell, J. W. and Creswell, J. D. (2023) *Research design: qualitative, quantitative, and mixed methods approaches*. 6th Ed. Los Angeles, SAGE Publications Ltd.

Creswell, J.W (2013) *Qualitative Inquiry and Research Design*. 3rd Ed. London, Sage Publications. pp.174-175.

Cresswell, K., Williams, R., and Sheikh, A. (2020). 'Developing and Applying a Formative Evaluation Framework for Health Information Technology Implementations: Qualitative

Investigation'. *Journal of medical Internet research*, *22*(6), e15068. DOI: https://doi.org/10.2196/15068

Crowther, L., Pearson, M., Cummings, H., and Crooks, M. G. (2022). 'Towards codesign in respiratory care: Development of an implementation-ready intervention to improve guideline-adherent adult asthma care across primary and secondary care settings (The SENTINEL Project)'. *BMJ Open Respiratory Research*, *9*(1). DOI: https://doi.org/10.1136/bmjresp-2021-001155

Curtis, E.A. and Drennan, J. (2013) *Quantitative health research: issues and methods*. Maidenhead: McGraw-Hill Education/Open University Press.

Cushnan, D., Berka, R., and Jacob, J. (2021). 'Towards nationally curated data archives for clinical radiology image analysis at scale: Learnings from national data collection in response to a pandemic'. *Digital Health*, *7*. DOI: https://doi.org/10.1177/20552076211048654

CVUHB (2024) *FOI-24-145 IT Spent and HIMSS Final Response. Freedom of Information Disclosure Log*. Available at: cavuhb.nhs.wales/about-us/governance-and-assurance/freedom-of-information/disclosure-log/2024/months/february/foi-24-145-it-spend-and-himss-final-response-pdf/ [Accessed 4 November 2024].

Dahlström, P., Desmet, D. and Singer, M. (2017) *The seven decisions that matter in a digital transformation: A CEO's guide to reinvention*. Available at: https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/the-seven-decisions-that-matter-in-a-digital-transformation [Accessed: 20 October 2024]

Darzi, A., 2024. *Independent Investigation of the National Health Service in England*. Updated 25 September 2024. London: Department of Health and Social Care. Available at: https://assets.publishing.service.gov.uk/media/66f42ae630536cb92748271f/Lord-Darzi-Independent-Investigation-of-the-National-Health-Service-in-England-Updated-25-September.pdf [Accessed 4 November 2024].

Das, S., and Hussey, P. (2022). 'How ontology can be used to achieve semantic interoperability in healthcare'. *The European Journal of Public Health*, *32*(3). DOI: https://doi.org/10.1093/eurpub/ckac129.363

Databricks (2024) *PTHB derives insights to bolster care with Databricks.* Customer Story 5th February 2034 Available at: https://www.databricks.com/customers/pthb [Accessed: 21 April 2024].

Davidson, B., Portillo, K.M.F., Wac, M., McWilliams, C., Bourdeaux, C., and Craddock, I. (2022). 'Requirements for a Bespoke Intensive Care Unit Dashboard in Response to the COVID-19 Pandemic: Semistructured Interview Study'. *JMIR Human Factors*, *9*(2). DOI: https://doi.org/10.2196/30523

Dawadi, S., Shrestha, S. and Giri, R.A. (2021) ‘Mixed-Methods Research: A Discussion on its Types, Challenges, and Criticisms’, *Journal of Practical Studies in Education*, 2(2), pp.25–36. DOI: https://doi.org/10.46809/jpse.v2i2.20.

Dawoodbhoy, F.M., Delaney, J., Cecula, P., Yu, J., Peacock, I., Tan, J., and Cox, B. (2021). ‘AI in patient flow: applications of artificial intelligence to improve patient flow in NHS acute mental health inpatient units’. *Heliyon*, *7*(5). DOI: https://doi.org/10.1016/j.heliyon.2021.e06993

De Mauro, A., Greco, M., and Grimaldi, M. (2016). ‘A formal definition of Big Data based on its essential features’. *Library Review*. 65. pp.122-135. DOI:10.1108/LR-06-2015-0061.

Department of Health and Social Care (2022) *Health and social care integration: joining up care for people, places and populations*. Policy Paper, Department of Health & Social Care. Available at: https://www.gov.uk/government/publications/health-and-social-care-integration-joining-up-care-for-people-places-and-populations/health-and-social-care-integration-joining-up-care-for-people-places-and-populations [Accessed 11 October 2024]

Department of Health and Social Care (2022a) *Data saves lives: reshaping health and social care with data*. Policy Paper, Department of Health & Social Care. Available at: https://www.gov.uk/government/publications/data-saves-lives-reshaping-health-and-social-care-with-data/data-saves-lives-reshaping-health-and-social-care-with-data [Accessed 7 November 2024]

Department of Health and Social Care (2022b) *Secure data environment for NHS health and social care data - policy guidelines.* Policy Paper, Department of Health & Social Care. Available at: https://www.gov.uk/government/publications/secure-data-environment-policy-guidelines/secure-data-environment-for-nhs-health-and-social-care-data-policy-guidelines [Accessed 7 November 2024]

Department of Health and Social Care (2019) *NHSX: new joint organisation for digital, data and technology.* News Story. https://www.gov.uk/government/news/nhsx-new-joint-organisation-for-digital-data-and-technology [Accessed 7 November 2024]

Digital Health and Care Wales (2024) *Data Analytics Advancement: NHS Executive National Data Analytics Platform*. Available at: https://dhcw.nhs.wales/news/latest-news/data-analytics-advancement-nhs-executive-national-data-analytics-platform/#:~:text=29%20May%202024,Board%20and%20Trust%20data%20providers](https://dhcw.nhs.wales/news/latest-news/data-analytics-advancement-nhs-executive-national-data-analytics-platform/#:~:text=29%20May%202024,Board%20and%20Trust%20data%20providers [Accessed: 8 November 2024].

Digital Health Intelligence (2023) *Digital Maturity Review of NHS Acute Trusts*. Available at: https://digitalhealthintelligence.net/wp-content/uploads/2023/07/DHI-Digital-Maturity-Review-of-NHS-Acute-Trusts-1.pdf [Accessed: 11 November 2024].

Dillman, D.A., Smyth, J.D., and Christian, L.M. (2014). *Internet, mail, and mixed-mode surveys: The tailored design method.* 4th Edn. Hoboken, NJ: Wiley.

Doorenbos, A.Z. (2014). 'Mixed Methods in Nursing Research: An Overview and Practical Examples'. *Kango kenkyu. The Japanese journal of nursing research*, *47*(3), pp.207–217.

Dovgyi, S., and Kopiika, O. (2022). 'Standard Model of System Architecture of Enterprise IT Infrastructure'. In: Ilchenko, M., Uryvsky, L., Globa, L. (eds) *Progress in Advanced Information and Communication Technology and Systems*. MCiT 2021. Lecture Notes in Networks and Systems, vol 548. Springer, Cham. DOI: https://doi.org/10.1007/978-3-031-16368-5_9

Dudovskiy, J. (2020) *Research Philosophy - Research Methodology*. Available at: https://research-methodology.net/research-philosophy/ [Accessed: 26 November 2024].

Dwyer, J., Liang, Z. and Thiessen, V. (2019). *Project management in health and community services : getting good ideas to work.* 3rd Edn. Sydney: Allen & Unwin.

Elizondo, A. (2024). 'Governance intricacies in implementing regional shared care records: A qualitative study in the national health service, England'. *Health Informatics Journal*, *30*(4). DOI: https://doi.org/10.1177/14604582241290709

Federation for Informatics Professionals (2024) *DRAFT FEDIP Response to the Tony Blair Institute Report: "Preparing the NHS for the AI Era: A Digital Health Record for Every Citizen."* Available at: https://c2f91b66-85f1-477d-9fd6-66053db82524.usrfiles.com/ugd/c2f91b_05da1934337542508ba6b6feb2c495e1.pdf?utm_source=ActiveCampaign&utm_medium=email&utm_content=Seeking%20Your%20Insights%20on%20Transforming%20NHS%20with%20AI%20and%20Digital%20Health%20Records&utm_campaign=FDLs%20Blair%20Report&utm_source=ActiveCampaign&utm_medium=email&utm_content=Lend%20Your%20Voice%3A%20Feedback%20Needed%20on%20Digital%20Health%20Transformation%20Report&utm_campaign=Practitioner%20Re%20BLAIR [Accessed 4 November 2024]

Federation for Informatics Professionals (2022) *An Introduction To; The Federation for Informatics Professionals in Health and Social Care*. [Online Brochure] Available at: https://www.fedip.org/the-fedip-brochure [Accessed 4 November 2024]

Feely, K., Edbrooke, L., Bower, W., Mazzone, S., Merolli, M., Staples, J., and Martin, A. (2023) 'Allied health professionals' experiences and lessons learned in response to a big bang electronic medical record implementation: A prospective observational study', *International Journal of Medical Informatics*, 176, 105094. DOI: https://doi.org/10.1016/J.IJMEDINF.2023.105094

Finch, T.L., Girling, M., May, C.R., Mair, F.S., Murray, E., Treweek, S., McColl, E., Steen, I.N., Cook, C., Vernazza, C.R., Mackintosh, N., Sharma, S., Barbery, G., Steele, J. and Rapley, T. (2018) 'Improving the normalization of complex interventions: part 2 - validation of the NoMAD instrument for assessing implementation work based on normalization process theory (NPT)'. *BMC Medical Research Methodology,* 18, 135. DOI: https://doi.org/10.1186/s12874-018-0591-x

Floridi, L., and Taddeo, M. (2016). 'What is data ethics?' In *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*. 374(2083). Royal Society of London. DOI: https://doi.org/10.1098/rsta.2016.0360

Friedman, C.P., Wyatt, J. and Ash, J.S.(2022) *Evaluation methods in biomedical and health informatics*. 3rd edn. Cham, Switzerland: Springer. DOI: 10.1007/978-3-030-86453-8.

Gale, N.K., Heath, G., Cameron, E., Rashid, S., and Redwood, S. (2013).' Using the framework method for the analysis of qualitative data in multi-disciplinary health research'. *BMC Medical Research Methodology,* 13, 117 DOI: https://doi.org/10.1186/1471-2288-13-117

Gartner (n.d.) 'Definition of Legacy Application or System'. *Gartner Information Technology Glossary.* Available at: https://www.gartner.com/en/information-technology/glossary/legacy-application-or-system [Accessed 16 April 2025]

Ghafur, S., O'Brien, N., Howitt, P., Painter, A., O'Shaughnessy, J., and Darzi, A. (2023) *NHS Data - Maximising its impact for all.* Institute of Global Health Innovation, Imperial College London. DOI: https://doi.org/10.25561/103404 Available at: https://spiral.imperial.ac.uk/bitstream/10044/1/103404/8/NHS%20Data%20-%20Maximising%20Impact%20for%20All.pdf [Accessed 11 October 2024]

Gill, S.L. (2020) 'Qualitative Sampling Methods', *Journal of Human Lactation*, 36(4), pp.579–581. DOI: https://doi.org/10.1177/0890334420949218

Gliklich, R.E. and Leavy, M.B. (2019). Data Standards. In: R.E. Gliklich, M.B. Leavy and N.A. Dreyer, (2019). *Tools and Technologies for Registry Interoperability, Registries for Evaluating Patient Outcomes: A User's Guide,* 3rd Edn*, Addendum 2* [Internet]. Rockville (MD): Agency for Healthcare Research and Quality (US), Chapter 3. Available at: https://www.ncbi.nlm.nih.gov/books/NBK551886/ [Accessed 21 November 2024].

Goldacre, B. Bardsley, M., Benson, T., Cheema, K., Chinn, R., Coughlan, E., Dougan, S., Farr, M., Hawkins, L., Jonas, A., Kinnear, A., McInnes, M., Mohammed, M.A., Morton, C., Pasumarthy, R., Pettinger, N., Rowland, B., Sebire, N., Stroner, P., Tennison, J., Warnakula, S., Watson, O., Wright, E., Young, H., and Morley, J. (2020). 'Bringing NHS data analysis into the 21st century'. *Journal of the Royal Society of Medicine*, 113(10), pp.383–388. DOI: 10.1177/0141076820930666

Golinelli, D., Toscano, F., Bucci, A., and Carullo, G. (2018) 'Transferring Health Big Data within the European Legal Framework: What Role for National Healthcare Services?' *J Law Med*. Dec;26(2), pp.488-493. PMID: 30574732.

Goulding, L., Adamson, J., Watt, I. and Wright, J. (2015), 'Lost in hospital: a qualitative interview study that explores the perceptions of NHS inpatients who spent time on clinically inappropriate hospital wards'. *Health Expect,*18, pp.982-94. DOI https://doi.org/10.1111/hex.12071

Gray, A., Mackie, C., Price, W., Coupe, E., Holgate, S., Watson, E., Eatough, J., Fletcher, N., Hodson, K., and Mantzourani, E. (2020). 'Learning from Practice: How East Lancashire Hospitals' Pharmacy Service Has Embraced Information Technology'. *Pharmacy*, *8*(4),177. DOI: https://doi.org/10.3390/pharmacy8040177

Gorard, S. (2006). *Using everyday numbers effectively in research: not a book about statistics*, London: Continuum.

Groves, R.M., and Peytcheva, E. (2008). 'The impact of nonresponse rates on nonresponse bias: A meta-analysis'. *Public Opinion Quarterly, 72*(2), pp.167–189. DOI: https://doi.org/10.1093/poq/nfn011

Hackenberger, B.K. (2019). 'Data by data, Big Data'. *Croatian Medical Journal*, *60*(3), pp.290–292. https://doi.org/10.3325/cmj.2019.60.290

Halcomb, E., Massey, D. and Gunowa, N. (2023) *Navigating the Maze of Research: Enhancing Nursing and Midwifery Practice*, Elselvier Health Sciences. Available at: https://books.google.co.uk/books?hl=en&lr=&id=TbSvEAAAQBAJ&oi=fnd&pg=PA125&dq=initiation+in+mixedmethods+research+methodology&ots=E1Oi3BszkU&sig=DxPWVJzoBqrvEGiPDmP_7IVjvmzM#v=onepage&q&f=false [Accessed: 20 January 2025)

Hammer, M.J. (2017) 'Ethical Considerations for Data Collection Using Surveys'. *Oncol Nurs Forum*. Mar 1;44(2), pp.157-159. DOI: 10.1188/17. PMID: 28222089.

Hammond, W.E., Bailey, C., Boucher, P., Spohr, M., and Whitaker, W. (2010) 'Connecting information to improve health'. *Health Aff*; 29, pp.284–288. DOI: 10.1377/hlthaff.2009.0903

Harrison, R. L., Reilly, T. M., and Creswell, J. W. (2020). 'Methodological Rigor in Mixed Methods: An Application in Management Studies'. *Journal of Mixed Methods Research,* 14(4), pp.473-495. DOI: https://doi.org/10.1177/1558689819900585

Hariton, E., and Locascio, J.J. (2018). 'Randomised controlled trials - the gold standard for effectiveness research: Study design: randomised controlled trials'. *BJOG: an international journal of obstetrics and gynaecology*, *125*(13), 1716. DOI: https://doi.org/10.1111/1471-0528.15199

Health Data Hub (2021) *Position paper on health data governance*. Available at: https://www.healthdatacollaborative.org/fileadmin/uploads/hdc/Documents/2021/Data_Governance_Summit_Part_2/COMBES_St%C3%A9phanie_Summary_HealthDataHubPositionPaper2.pdf [Accessed 14 November 2024]

Health Level 7 International (2023) *FHIR Overview HL7 FHIR Release 5.* Available at: https://www.hl7.org/fhir/overview.html [Accessed 11 October 2024]

Helou, M. A., DiazGranados, D., Ryan, M. S., and Cyrus, J. W. (2020). 'Uncertainty in Decision Making in Medicine: A Scoping Review and Thematic Analysis of Conceptual Models'. *Academic medicine: journal of the Association of American Medical Colleges*, *95*(1), pp.157–165. DOI: https://doi.org/10.1097/ACM.0000000000002902

Herlitz, L., Crellin, N., Vindrola-Padros, C., Ellins, J., Georghiou, T., Litchfield, I., Massou, E., Ng, P.L., Sherlaw-Johnson, C., Sidhu, M.S., Tomini, S.M., Walton, H., and Fulop, N.J. (2023). 'Patient and staff experiences of using technology-enabled and analogue models of remote home monitoring for COVID-19 in England: A mixed-method evaluation'. *International Journal of Medical Informatics*, *179*. DOI: https://doi.org/10.1016/j.ijmedinf.2023.105230

Hersh, W. and Wright, A. (2008) 'What workforce is needed to implement the health information technology agenda? Analysis from the HIMSS analytics database'. *AMIA Annu Symp Proc*. pp.303–307. PMID:18998990; PubMed Central PMCID: PMCPMC2656033

HIMSS (2024). *Understanding National Interoperability Frameworks and Networks*. Available at: https://gkc.himss.org/sites/hde/files/media/file/2024/08/02/himss_interopframeworks_ebook-2.pdf [Accessed: 8 November 2024].

HIMSS (2024a) *'Interoperability in healthcare'*, HIMSS Global Conference & Exhibition. Available at: https://gkc.himss.org/resources/interoperability-healthcare [Accessed: 8 November 2024].

HIMSS (2024b). *The value of clinical informatics in improving care delivery and outcomes*. HIMSS. Available at: https://www.himss.org/sites/hde/files/media/file/2024/04/18/wp_value-of-clinical-informatics-1.pdf [Accessed: 8 November 2024].

HIMSS (2023) *Electronic Medical Record Adoption Model (EMRAM) Achievement Assessment Report Powys Teaching Local Health Board*. Unpublished.

HIMSS (2022) *Electronic Medical Record Adoption Model (EMRAM) Criteria | HIMSS*. Available at: https://www.himss.org/sites/hde/files/2022-01/emram-criteria.pdf [Accessed: 20 November 2024].

Honavar S.G. (2020). 'Electronic medical records - The good, the bad and the ugly'. *Indian journal of ophthalmology*, *68*(3), pp.417–418. DOI: https://doi.org/10.4103/ijo.IJO_278_20

Hulsen T. (2020). 'Sharing Is Caring-Data Sharing Initiatives in Healthcare'. *International journal of environmental research and public health*, *17*(9), 3046. DOI: https://doi.org/10.3390/ijerph17093046

IBM (2011) *From Stretched to Strengthened: Insights from the Global Chief Marketing Officer Study – Automotive Point of View*. Available at: https://public.dhe.ibm.com/software/fr/pdf/CMO_Study_Automotive_Point.pdf [Accessed 26 October 2024]

Idowu, E.A.A., Teo, J., Salih, S., Valverde, J., and Yeung, J. A. (2023). 'Streams, rivers and data lakes: an introduction to understanding modern electronic healthcare records'. *Clinical Medicine, Journal of the Royal College of Physicians of London.* 23(4), pp.409–413. Royal College of Physicians. DOI: https://doi.org/10.7861/clinmed.2022-0325

Jager, A., Papoutsi, C., and Wong, G. (2023). 'The usage of data in NHS primary care commissioning: a realist evaluation'. *BMC Primary Care*, *24*(1). DOI: https://doi.org/10.1186/s12875-023-02193-4

Jamieson S. (2004) 'Likert scales: how to (ab)use them'. *Med Educ.* Dec;38(12):1217-8. DOI: 10.1111/j.1365-2929.2004.02012.x. PMID: 15566531.

John, M. (2024) *Digital Strategy Update. Agenda Item 2.4.* Swansea Bay University Health Board. Available at: https://view.officeapps.live.com/op/view.aspx?src=https%3A%2F%2Fsbuhb.nhs.wales%2Fabout-us%2Fkey-documents-folder%2Fworkforce-and-od-committee-papers%2Fworkforce-and-od-committee-february-2024%2F2-4-digital-strategy-update-docx%2F&wdOrigin=BROWSELINK [Accessed on 4 November 2022]

Johnson, J. L., Adkins, D., and Chauvin, S. (2020). 'A Review of the Quality Indicators of Rigor in Qualitative Research'. *American journal of pharmaceutical education*, *84*(1), 7120. DOI: https://doi.org/10.5688/ajpe7120

Johnson, R.B., and Onwuegbuzie, A.J. (2004). 'Mixed Methods Research: A Research Paradigm Whose Time Has Come'. *Educational Researcher*, 33(7), pp.14-26. DOI: https://doi.org/10.3102/0013189X033007014

Jones, L.A, Nelder, J.R., Fryer, J.M., Alsop, P.H., Geary, M.R., Prince, M., and Cardinal, R.N. (2022). 'Public opinion on sharing data from health services for clinical and research purposes without explicit consent: An anonymous online survey in the UK'. *BMJ Open*, *12*(4). DOI: https://doi.org/10.1136/bmjopen-2021-057579

Justinia, T. (2017) 'The UK's National Programme for IT: Why was it dismantled?' *Health Services Management Research*. 30(1), pp.2-9. DOI:10.1177/0951484816662492

Khanbhai, M., Flott, K., Manton, D., Harrison-White, S., Klaber, R., Darzi, A., and Mayer, E. (2021). 'Identifying factors that promote and limit the effective use of real-time patient experience feedback: A mixed-methods study in secondary care'. *BMJ Open*, *11*(12). DOI: https://doi.org/10.1136/bmjopen-2020-047239

Kapadi, A., Price, G., Faivre-Finn, C., Holley, R., Wicks, K., Banfill, K., Webster, G., Franks, K., McDonald, F., Johnson, D., and French, D.P. (2024). ‘Feasibility of implementing a rapid-learning methodology to inform radiotherapy treatments: Key professional stakeholders’ views’. *BMJ*, *3*(1). DOI: https://doi.org/10.1136/bmjonc-2023-000226

Kelly, L. M., and Cordeiro, M. (2020). ‘Three principles of pragmatism for research on organizational processes’. *Methodological Innovations*, 13(2). DOI: https://doi.org/10.1177/2059799120937242

Kernebeck, S., Jux, C., Busse, T.S., Meyer, D., Dreier, L.A., Zenz, D., Zernikow, B., and Ehlers, J.P. (2022) ‘Participatory Design of a Medication Module in an Electronic Medical Record for Paediatric Palliative Care: A Think-Aloud Approach with Nurses and Physicians’. *Children*, 9(1), pp.82. DOI: https://doi.org/10.3390/children9010082

Klein, R. (2013) ‘The NHS in the age of anxiety: rhetoric and reality—an essay by Rudolf Klein’ *BMJ*; 347:f5104 DOI:10.1136/bmj.f5104

Khong, I., Yusuf, N.A., Nuriman, A., and Yadila, A.B. (2023). ‘Exploring the Impact of Data Quality on Decision-Making Processes in Information Intensive Organizations’. *APTISI Transactions on Management (ATM),* 7(3), pp.253-260.

Krejcie, R.V., and Morgan, D.W., (1970). ‘Determining Sample Size for Research Activities’. *Educational and Psychological Measurement.*

Lau, R.S., Boesen, M.E., Richer, L., and Hill, M.D. (2024). ‘Siloed mentality, health system suboptimization and the healthcare symphony: a Canadian perspective’. *Health research policy and systems*, *22*(1), 87. DOI: https://doi.org/10.1186/s12961-024-01168-w

Lewin, K. (1958) ‘Group decisions and social change’. In Maccoby, E.E. *et al.* (ed), *Readings in Social Psychology,* New York, NY: Holt. Rinehart & Winston, pp.192-211

Li, E., Lounsbury, O., Clarke, J., Ashrafian, H., Darzi, A., and Neves, A.L. (2023). ‘Perceptions of chief clinical information officers on the state of electronic health records systems interoperability in NHS England: a qualitative interview study’. *BMC Medical Informatics and Decision Making*, *23*(1). DOI: https://doi.org/10.1186/s12911-023-02255-8

Litchfield, I., Kingston, B., Narga, D., and Turner, A. (2022). ‘The move towards integrated care: Lessons learnt from managing patients with multiple morbidities in the UK’. *Health Policy*, *126*(8), pp.777–785. DOI: https://doi.org/10.1016/j.healthpol.2022.05.010

Liu, M., Li, S., Yuan, H., Ong, M.E.H., Ning, Y., Xie, F., Saffari, S.E., Shang, Y., Volovici, V., Chakraborty, B., and Liu, N. (2023). ‘Handling missing values in healthcare data: A systematic review of deep learning-based imputation techniques’. *Artificial Intelligence in Medicine, 142*, 102587. DOI: https://doi.org/10.1016/j.artmed.2023.102587

Lloyd-Jones, N. (2024). *How is the NHS in Wales funded?* NHS Confederation. Online Article. Available at: https://www.nhsconfed.org/articles/how-nhs-wales-funded-0 [Accessed 27 March 2025]

Lincoln, Y.S. and Guba, E.G. (1985). *Naturalistic Inquiry*. Newbury Park, CA: Sage Publications

Mate, K. (2022) *On the Quintuple Aim: Why expand beyond the Triple Aim?* Institute for Healthcare Improvement (IHI) Available at: https://www.ihi.org/insights/quintuple-aim-why-expand-beyond-triple-aim [Accessed 5 December 2024].

Martínez-Mesa, J., González-Chica, D.A., Duquia, R.P., Bonamigo, R.R., and Bastos, J.L. (2016) 'Sampling: how to select participants in my research study?' *Anais Brasileiros de Dermatologia,* 91(3), pp.326–330. Available at: https://doi.org/10.1590/ABD1806-4841.20165254.

Martínez-Mesa, J., González-Chica, D.A., Bastos, J.L., Bonamigo, R.R., and Duquia, R.P. (2014). 'Sample size: how many participants do I need in my research?'. *Anais brasileiros de dermatologia*, *89*(4), pp.609–615. DOI: https://doi.org/10.1590/abd1806-4841.20143705

McGowan, L.J., Graham, F., Lecouturier, J., Goffe, L., Echevarria, C., Kelly, M.P., and Sniehotta, F.F. (2024). 'The Views and Experiences of Integrated Care System Commissioners About the Adoption and Implementation of Virtual Wards in England: Qualitative Exploration Study'. *Journal of Medical Internet Research*, *26*(1). DOI: https://doi.org/10.2196/56494

McVey, L., Alvarado, N., Greenhalgh, J., Elshehaly, M., Gale, C. P., Lake, J., Ruddle, R. A., Dowding, D., Mamas, M., Feltbower, R., and Randell, R. (2021). 'Hidden labour: the skilful work of clinical audit data collection and its implications for secondary use of data via integrated health IT'. *BMC Health Services Research*, *21*(1). DOI: https://doi.org/10.1186/s12913-021-06657-0

Memon, M.A., Ting, H., Cheah, J.H., Thurasamy, R., Chuah, F., and Cham, T.H. (2020). 'Sample Size For Survey Research: Review And Recommendations'. *Journal of Applied Structural Equation Modeling*. 4(2). DOI: https://jasemjournal.com/wp-content/uploads/2020/08/Memon-et-al_JASEM_-Editorial_V4_Iss2_June2020.pdf

Microsoft Azure (2025) *What is data integration?* Available at: https://azure.microsoft.com/en-us/resources/cloud-computing-dictionary/what-is-data-integration [Accessed: 24 January 2025].

Mishra, P., Pandey, C.M., Singh, U., Keshri, A., and Sabaretnam, M. (2019). 'Selection of appropriate statistical methods for data analysis'. *Annals of cardiac anaesthesia*, *22*(3), pp.297–301. DOI: https://doi.org/10.4103/aca.ACA_248_18

Mistry, P., Maguire, D., Chikwira, L., and Lindsay, T. (2022). *Interoperability is more than technology: The role of culture and leadership in joined-up care*. The King's Fund. Available at: https://assets.kingsfund.org.uk/f/256914/x/c48bd5a1a2/interoperability_more_than_technology_2022.pdf [Accessed on 4 November 2022]

Mitchell, A. (2020). ‘A NICE perspective on computable biomedical knowledge’. *BMJ Health and Care Informatics*, *27*(2). DOI: https://doi.org/10.1136/bmjhci-2019-100126

Michie, S., van Stralen, M.M. and West, R. (2011) ‘The behaviour change wheel: A new method for characterising and designing behaviour change interventions’. *Implementation Science*, 6(1), pp.1–12. DOI: 10.1186/1748-5908-6-42

Morgan, D. (2017). *Re: Are open-ended questions on surveys enough for a study to be considered mixed-method?* Available at: https://www.researchgate.net/post/Are-open-ended-questions-on-surveys-enough-for-a-study-to-be-considered-mixed-method/5876899493553b3c65122d82/citation/download. [Accessed: 5 December 2024

Morgan, D. (2014). ‘Pragmatism as a paradigm for mixed methods research’. In *Pragmatism as a paradigm for mixed methods research*. SAGE Publications, Inc., pp. 25-44. DOI: https://doi.org/10.4135/9781544304533

Morris, A.C. Ibrahim, Z., Moghraby, O.S., Stringaris, A., Grant, I.M., Zalewski, L., McClellan, S., Moriarty, G., Simonoff, E., Dobson, R.J.B., and Downs, J. (2023). ‘Moving from development to implementation of digital innovations within the NHS: myHealthE, a remote monitoring system for tracking patient outcomes in child and adolescent mental health services’. *Digital Health*, *9*. DOI: https://doi.org/10.1177/20552076231211551

Morse, J.M. (1991) ‘Approaches to qualitative-quantitative methodological triangulation’. *Nursing Research,* 40(2), pp.120–123.

Mortensen, B., Borkowski, N., O'Connor, S.J., Patrician, P.A., and Weech-Maldonado, R. (2019) ‘The Relationship Between Hospital Interdepartmental Transfers and Patient Experience’. *Journal of Patient Experience*. 7(2), pp.263-269. DOI:10.1177/2374373519836467

Moule, P. (2020) *Making Sense of Research in Nursing, Health and Social Care*. 7th edn. SAGE Publications Ltd

Murtagh, F., and Legendre, P. (2014) ‘Ward’s Hierarchical Agglomerative Clustering Method: Which Algorithms Implement Ward’s Criterion?’ *Journal of Classification* **31**, pp.274–295. DOI: https://doi.org/10.1007/s00357-014-9161-z

National Academy of Medicine (2018) ‘Procuring Interoperability: Achieving High-Quality, Connected, and Person-Centered Care’. *The Learning Health System Series I. WHY INTEROPERABILITY IS ESSENTIAL IN HEALTH CARE*. Washington, DC: National

Academies Press (US),Available at: https://www.ncbi.nlm.nih.gov/books/NBK594855/ [Accessed 26 October 2024]

National Audit Office (2020). *Digital transformation in the NHS*. Available at: www.nao.org.uk/reports/the-use-of-digital-technology-in-the-nhs [Accessed on 4 November 2022]

Naqvi, D. Malik, A., Al-Zubaidy, M., Naqvi, F., Tahir, A., Tarfiee, A., Vara, S., and Meyer, E. (2019). 'The general practice perspective on barriers to integration between primary and social care: A London, United Kingdom-based qualitative interview study'. *BMJ Open*, *9*(8). DOI: https://doi.org/10.1136/bmjopen-2019-029702

NHS Digital (2023). *NHS Interoperability Framework*. Available at: https://digital.nhs.uk/services/interoperability-toolkit/developer-resources/nhs-interoperability-framework [Accessed 23 October 2024].

NHS Digital. (2024). *FHIR UK Core*. Available at: https://digital.nhs.uk/services/fhir-uk-core. [Accessed 23 October 2024].

NHS Digital Academy. (2023). *About the NHS Digital Academy*. Available at: https://digital-transformation.hee.nhs.uk/digital-academy/the-nhs-digital-academy/about-the-nhs-digital-academy/about-the-nhs-digital-academy [Accessed 23 January 2025].

NHS England (2019). *The NHS long term plan*. Available at: www.longtermplan.nhs.uk/publication/nhs-long-term-plan [Accessed on 4 November 2024].

NHS Wales (2024) 'NHS NUMBER'*. NHS Wales Data Dictionary* Available at: https://www.datadictionary.nhs.uk/attributes/nhs_number.html [Accessed on 4 November 2024].

NHS Wales (2022) 'Main Specialty'*. NHS Wales Data Dictionary.* Available at: https://www.datadictionary.wales.nhs.uk/index.html#!WordDocuments/mainspecialtyconsultant.htm [Accessed on 11 April 2025].

NLTK Project (2023) *Natural Language Toolkit*. Available at: https://www.nltk.org/ [Accessed: 4 November 2024]

Noble, H., and Heale, R. (2019). 'Triangulation in research, with examples'. *Evidence-based nursing*, *22*(3), pp.67–68. DOI: https://doi.org/10.1136/ebnurs-2019-103145

Norman G. (2010) 'Likert scales, levels of measurement and the "laws" of statistics'. *Adv Health Sci Educ Theory Pract.*;15(5), pp.625–632. DOI: https://doi.org/10.1007/s10459-010-9222-y

Nundy, S., Cooper, L.A., Mate, K.S. (2022) 'The quintuple aim for health care improvement: A new imperative to advance health equity'. *JAMA* ;327(6), pp.521-522.

Okorie, G.N., Udeh, C.A., Adaga, E.M., DaraOjimba, O.D., and Oriekhoe, O.I. (2024). ‘Ethical considerations in data collection and analysis: a review: investigating ethical practices and challenges in modern data collection and analysis’. *International Journal of Applied Research in Social Sciences*, *6*(1), pp.1–22. DOI: https://doi.org/10.51594/ijarss.v6i1.688

ONS (2024*) Population estimates for England and Wales: mid-2022*. Available at: https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/bulletins/populationestimatesforenglandandwales/mid2022 [Accessed 15 May 2024]

Onwuegbuzie, A. J., and Teddlie, C. (2003). ‘A framework for analyzing data in mixed methods research’. In Tashakkori, A. and Teddlie, C. (Eds.), *Handbook of mixed methods in social and behavioral research.* pp.351-383. Thousand Oaks, CA: Sage

Oreščanin, D., and Hlupic, T. (2021). ‘Data Lakehouse - a Novel Step in Analytics Architecture’. *44th International Convention on Information, Communication and Electronic Technology (MIPRO),* pp.1242-1246. DOI: 10.23919/MIPRO52101.2021.9597091

Pankhurst, T., Atia, J., Evison, F., Gallier, S., Lewis, J.M., McKee, D., Ryan, S., Sapey, E., Ball, S., and Coleman, J.J. (2021). ‘Rapid adaptation of a local healthcare digital system to COVID-19: The experience in Birmingham (UK)’. *Health Policy and Technology*, *10*(4). DOI: https://doi.org/10.1016/j.hlpt.2021.100568

Page, M.J., McKenzie, J.E., Bossuyt, P.M., Boutron, I., Hoffmann, T.C., Mulrow, C.D., Shamseer, L., Tetzlaff, J.M., Akl, E.A., Brennan, S.E., Chou, R., Glanville, J., Grimshaw, J.M., Hróbjartsson, A., Lalu, M.M., Li, T., Loder, E.W., Mayo-Wilson, E., McDonald, S., McGuinness, L.A., Stewart, L.A., Thomas, J., Tricco, A.C., Welch, V.A., Whiting, P., and Moher, D. (2021) ‘The PRISMA 2020 statement: an updated guideline for reporting systematic reviews’. *BMJ*, 372(71). DOI: https://doi.org/10.1136/bmj.n71

Patel, J. (2019). ‘Bridging Data Silos Using Big Data Integration’. *International Journal of Database Management Systems*, *11*(3), 01–06. DOI: https://doi.org/10.5121/ijdms.2019.11301

Patton, M. Q. (2015). *Qualitative Research and Evaluation Method*s. 4th edn. London: Sage.

Perlman, M. and Elsner, N. (2019). *Digital health technology: Global case studies of health care transformation.* Deloitte Insights Available at: https://www2.deloitte.com/us/en/insights/industry/health-care/digital-health-technology.html [Accessed: 7 November 2024]

Phiri, M. (2022) 'Exponential Growth of Data'. *Medium*, 19 November. Available at: https://medium.com/@mwaliph/exponential-growth-of-data-2f53df89124 [Accessed 15 May 2024]

Plantera, F. (2024) 'Scaling Interoperability across levels of governance and states in Brazil.' *X-Road.* Available as: https://x-road.global/xroad-case-studies-library/2024/10/21/scaling-interoperability-across-states-for-national-digital-transformation-in-brazil [Accessed 14 September 2024]

Plantera, F. (2024a) 'CamDX is Cambodia's national data exchange solution – and it's based on X-Road®.' *X-Road.* Available as: https://x-road.global/xroad-case-studies-library/2024/10/21/camdx-is-cambodias-national-data-exchange-solution-and-its-based-on-x-road [Accessed 14 September 2024]

Ponto, J. (2015). 'Understanding and Evaluating Survey Research'. *Journal of the advanced practitioner in oncology*, *6*(2), pp.168–171.

Pope, R. Zenonos, A., Bryant, W., Spiridou, A., Key, D., Patel, S., Robinson, J., Styles, A., Rockenbach, C., Bicknell, G., Rajendran, P., Taylor, A.M., and Sebire, N.J. (2024). 'Real-world learnings for digital health industry–NHS collaboration: Life sciences vision in action'. *Future Healthcare Journal*, *11*(3), 100168. DOI: https://doi.org/10.1016/j.fhj.2024.100168

Powys County Council (2014) 'CBH0032 - Evidence on Cross-border health arrangements between England and Wales (parliament.uk)' *Written evidence submitted by Powys County Council (CBH 32)*. Available at: https://committees.parliament.uk/writtenevidence/52670/html/

PTHB (2024) *Financial pressures on the NHS prompt change in PTHB escalation status*. Available at: https://pthb.nhs.wales/news/health-board-news/financial-pressures-on-the-nhs-prompt-change-in-pthb-escalation-status/ [Accessed 14 November 2024]

PTHB (2023) *Integrated Shared Care Report Executive Report* – Internal PTHB Report. Unpublished.

Qlik (n.d.) *What is Real-Time Data? Definition & Best Practices*. Available at: https://www.qlik.com/us/streaming-data/real-time-data#:~:text=Real%2Dtime%20data%20refers%20to,bandwidth%20can%20create%20a%20lag. [Accessed 18 April 2025]

Saldaña, J. (2009). *The coding manual for qualitative researchers.* Sage Publications Ltd.

Savitz, L.A., Greene, S.M., Gould, M.K., and Luft, H.S. (2024). 'The Right Stuff: Getting the right data at the right time and using that data to drive evidence-based practice and policy'. *Learning Health Systems* .8(S1). John Wiley and Sons Inc. DOI: https://doi.org/10.1002/lrh2.10432

Schein, E.H. (1984) ‘Coming to a new awareness of organizational culture’. *MIT Sloan Management Review.* Available at: https://sloanreview.mit.edu/article/coming-to-a-new-awareness-of-organizational-culture

Scott, J. (1990). *A Matter of Record.* Cambridge: Polity

Scott, P. J. (2016). ‘Mixed methods: A paradigm for holistic evaluation of health IT’. *In Evidence-Based Health Informatics: Promoting Safety and Efficiency through Scientific Methods and Ethical Policy*. 222, pp.102–113. IOS Press. DOI: https://doi.org/10.3233/978-1-61499-635-4-102

Scott, T. Mannion, R., Marshall, M., and Davies, H. (2003) ‘Does organisational culture influence health care performance? A review of the evidence’. *Journal of Health Services Research and Policy*. 8(2). DOI: 10.1258/135581903321466085

Sexton-Radek, K. and Simmons, L. (2018) ‘Reverse Scoring’. In: *The SAGE Encyclopedia of Educational Research, Measurement, and Evaluation*. Thousand Oaks, CA: SAGE Publications, Inc. pp.1433-1434 Available at: https://doi.org/10.4135/9781506326139 [Accessed 4 Apr 2025].

Shahid, N. U., and Sheikh, N. J. (2021). ‘Impact of Big Data on Innovation, Competitive Advantage, Productivity, and Decision Making: Literature Review’. *Open Journal of Business and Management*, *09*(02), pp.586–617. DOI: https://doi.org/10.4236/ojbm.2021.92032

Sharma, V., Augustine, T., Ainsworth, J., and van der Veer, S.N. (2022). ‘The evaluation of digital transformation in renal transplantation in the United Kingdom: A national interview study’. *International Journal of Medical Informatics*, *164*. DOI: https://doi.org/10.1016/j.ijmedinf.2022.104800

Schober, P., Boer, C., and Schwarte, L.A. (2018) ‘Correlation Coefficients: Appropriate Use and Interpretation’. *Anesthesia & Analgesia,* 126(5). pp.763-1768, DOI: 10.1213/ANE.0000000000002864

Schwarz, M., Coccetti, A., Draheim, M., and Gordon, G. (2020) ‘Perceptions of allied health staff of the implementation of an integrated electronic medical record across regional and metropolitan settings’, *Australian Health Review,* 44(6), pp.965–972. DOI: https://doi.org/10.1071/AH19024

Siegel, S. (2024). ‘Global Health Care Outlook: Trends, Opportunities, and Investment’. *Deloitte* Available at: https://www.deloitte.com/global/en/Industries/life-sciences-health-care/analysis/global-health-care-outlook.html [Accessed: 4 November 2024]

Silsand, L. (2017). ‘The ‘holy grail’ of interoperability of health information systems: Challenges and implications’. *Lecture Notes in Business Information Processing*, 294, pp.140–154. DOI:https://doi.org/10.1007/978-3-319-64695-4_11

Smith, D.A., Wang, T., Freeman, O., Crichton, C., Salih, H., Matthews, P.C., Davies, J., Várnai, K.A., Woods, K., Jones, C.R., Glampson, B., Mulla, A., Mercuri, L., Shaw, A.T., Drumright, L.N., Romão, L., Ramlakan, D., Higgins, F., Weir, A., Nastouli, E., Agarwal, K., Gelson, W., Cooke, G.S., and Barnes, E. (2020). 'National Institute for Health Research Health Informatics Collaborative: Development of a pipeline to collate electronic clinical data for viral hepatitis research'. *BMJ Health and Care Informatics*, *27*(3).DOI: https://doi.org/10.1136/bmjhci-2020-100145

Starmer, K., (2024). *PM speech on the NHS: 12 September 2024*. GOV.UK. Available at: https://www.gov.uk/government/speeches/pm-speech-on-the-nhs-12-september-2024 [Accessed 4 November 2024].

StatsWales (2024) *NHS staff by staff group and year*. Available at: https://statswales.gov.wales/Catalogue/Health-and-Social-Care/NHS-Staff/NHS-Staff-Summary/nhsstaff-by-staffgroup-year [Accessed: 31 October 2024].

Sullivan, R. Williams, J., Robinson, J., Thompson, I., McNicoll, I., and Nichols, T. (2023). 'Pathways to interoperable electronic patient records for health and social care: Part 1: for those involved directly in care'. *Future Healthcare Journal*, *10*(3), pp.232–237. DOI: https://doi.org/10.7861/fhj.2023-0071

Sullivan, G.M., and Artino, A.R. (2013). 'Analyzing and interpreting data from likert-type scales'. *Journal of graduate medical education*, *5*(4), pp.541–542. DOI: https://doi.org/10.4300/JGME-5-4-18

Rastogi, N. (2022). 'Healthcare's new frontier: The digital front door'. *BMJ Innovations*. 8(2). DOI: https://doi.org/10.1136/bmjinnov-2021-000874

RedHat (2020). *What is a REST API?* Available at: https://www.redhat.com/en/topics/api/what-is-a-rest-api [Accessed 11 October 2024]

Rees, J. (2025) ''Too much paperwork' in referring patients to England', *BBC News*. Available at: https://www.bbc.co.uk/news/articles/cn4zjn2p2keo#comments [Accessed: 20 March 2025].

Riley, S., Burhouse, A. and Nicholas, T. (2021) 'National Health Service (NHS) trust boards adopt statistical process control reporting: the impact of the Making Data Count Training Programme'. *BMJ Leader*, **0**, pp.1-6. DOI: 10.1136/leader-2020-000357

Ristevski, B. and Chen, M. (2018) 'Big data analytics in medicine and healthcare'. *J Inter Bioinform* 15(3):20170030 DOI: 10:1515/jib-2017-0030

Robson, C. (2024). *Real World Research: A Resource for Users of Social Research Methods in Applied Settings*. 5th edn. Chichester: Wiley

Roehrs, A., da Costa, C.A., Righi, R.D.R., Rigo, S.J., and Wichman, M.H. (2019). ‘Toward a Model for Personal Health Record Interoperability’. *IEEE journal of biomedical and health informatics*, *23*(2), pp.867–873. DOI: https://doi.org/10.1109/JBHI.2018.2836138

Tashakkori, A., and Teddlie, C. (1998). *Mixed methodology: Combining qualitative and quantitative approaches.* Sage Publications, Inc.

Thimbleby, H. (2022). ‘NHS Number Open Source Software: Implications for Digital Health Regulation and Development’. *ACM Transactions on Computing for Healthcare*, *3*(4). DOI: https://doi.org/10.1145/3538382

Tobin, G.A., and Begley, C.M. (2004). ‘Methodological rigour within a qualitative framework’. *Journal of advanced nursing*, *48*(4), pp.388–396. DOI: https://doi.org/10.1111/j.1365-2648.2004.03207.x

Trivedi, S., Hall, S., Inglis, F., and Chaudhry, A. (2023). ‘Proof-of-concept solution to create an interoperable timeline of healthcare data’. *BMJ Health and Care Informatics*, *30*(1). DOI: https://doi.org/10.1136/bmjhci-2023-100754

Uhrenfeldt, L. Aagaard, H., Hall, E.O.C., Fegran, L., Ludvigsen, M.S., and Meyer, G. (2013) ‘A qualitative meta-synthesis of patients′ experiences of intra- and inter-hospital transitions’. *Journal of Advanced Nursing* 69(8), pp.1678–1690 DOI: https://doi.org/10.1111/jan.12134

Unwin, A. (2020).’ Why Is Data Visualization Important? What Is Important in Data Visualization?’ *Harvard Data Science Review*, 2(1). DOI: https://doi.org/10.1162/99608f92.8ae4d525

UWTSD (2022a) *Research Ethics & Integrity Code of Practice.* University of Wales Trinity Saint David. Available at: https://www.google.co.uk/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&ved=2ahUKEwiBqa_PgKSGAxXJTkEAHS6LBhEQFnoECBsQAQ&url=https%3A%2F%2Fwww.uwtsd.ac.uk%2Fmedia%2F2041%2Fdownload%3Fattachment&usg=AOvVaw2dqYdHyxRvIuouJ30mmQT8&opi=89978449 [Accessed 11 October 2024]

UWTSD (2022b) *Research Data Management Policy*. University of Wales Trinity Saint David. Available at: https://www.google.co.uk/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&cad=rja&uact=8&ved=2ahUKEwiBqa_PgKSGAxXJTkEAHS6LBhEQFnoECD4QAQ&url=https%3A%2F%2Fwww.uwtsd.ac.uk%2Fmedia%2F1675%2Fdownload%3Fattachment&usg=AOvVaw3wEJB5IVBLaOvHbcrtqL9o&opi=89978449 [Accessed 11 October 2024]

Vezyridis, P., and Timmons, S. (2019). ‘Resisting big data exploitations in public healthcare: free riding or distributive justice?’ *Sociology of Health and Illness*, *41*(8), 1585–1599. DOI: https://doi.org/10.1111/1467-9566.12969

Virtanen, P., Gommers, R., Oliphant, T.E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S.J., Brett, M., Wilson, J., Millman, K.J., Mayorov, N., Nelson, A.R.J., Jones, E., Kern, R., Larson, E., Carey, C.J., Polat, İ., Feng, Y., Moore, E.W., VanderPlas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero, E.A., Harris, C.R., Archibald, A.M., Ribeiro, A.H., Pedregosa, F., van Mulbregt, P., and SciPy 1.0 Contributors. (2020). 'SciPy 1.0: fundamental algorithms for scientific computing in Python'. *Nature methods*, *17*(3), pp.261–272. DOI: https://doi.org/10.1038/s41592-019-0686-2

Vollmer, S., Mateen, B.A., Bohner, G., Király, F.J., Ghani, R., Jonsson, P., Cumbers, S., Jonas, A., McAllister, K.S.L., Myles, P., Grainger, D., Birse, M., Branson, R., Moons, K.G.M., Collins, G.S., Ioannidis, J.P.A., Holmes, C., and Hemingway, H. (2020) 'Machine learning and artificial intelligence research for patient benefit: 20 critical questions on transparency, replicability, ethics, and effectiveness'. *BMJ,* 368:l6927. DOI: https://doi.org/10.1136/bmj.l6927

Wain, L. M. (2021). 'Does integrated health and care in the community deliver its vision? A workforce perspective'. *Journal of Integrated Care*, *29*(2), pp.170–184. DOI: https://doi.org/10.1108/JICA-10-2020-0061

Warren, L. R., Clarke, J., Arora, S., and Darzi, A. (2019). 'Improving data sharing between acute hospitals in England: An overview of health record system distribution and retrospective observational analysis of inter-hospital transitions of care'. *BMJ Open*, 9(12). DOI: https://doi.org/10.1136/bmjopen-2019-031637

Wasti, S.P., Simkhada, P., van Teijlingen, E.R., Sathian, B., and Banerjee, I. (2022). 'The Growing Importance of Mixed-Methods Research in Health'. *Nepal journal of epidemiology*, *12*(1), pp.1175–1178. DOI: https://doi.org/10.3126/nje.v12i1.43633

Welsh Government (2025) *Health Board Allocations: Revenue Tables*. Available at: https://www.gov.wales/sites/default/files/publications/2025-01/health-board-allocations-revenue-tables.xlsx [Accessed 27 March 2025]

Welsh Government (2023) *Digital and data strategy for health and social care in Wales.* Policy and Strategy document. Available at: https://www.gov.wales/digital-and-data-strategy-health-and-social-care-wales-html [Accessed: 26 October 2024].

Welsh Government (2023a) *Introduction of HL7 FHIR as a foundational standard in all NHS Wales Bodies*. Welsh Health Circular. Available at: https://www.gov.wales/sites/default/files/publications/2023-06/introduction-of-hl7-fhir-in-all-nhs-wales-bodies_1.pdf [Accessed: 26 October 2024].

Welsh Government (2021) *Digital Health and Care Wales*. Available at: https://www.gov.wales/digital-health-and-care-wales [Accessed: 1 May 2025].

Welsh Parliament (2025) *Welsh Government draft budget 2025-26*. Health and Social Care Committee, Senedd Cymru. Available at: https://senedd.wales/media/sw5hbw0q/cr-ld16963-Ae.pdf [Accessed: 27 March 2025].

Willson, A., and Davies, A. (2021). 'Rhetoric or reform? Changing health and social care in Wales'. *International Journal of Health Policy and Management*, *10*(6), 295–298. DOI: https://doi.org/10.34172/ijhpm.2020.53

Wilson, S., Tolley, C., Mc Ardle, R., Beswick, E., Slight, and S.P. (2023). 'Key Considerations When Developing and Implementing Digital Technology for Early Detection of Dementia-Causing Diseases Among Health Care Professionals: Qualitative Study'. *Journal of Medical Internet Research*, *25*. DOI: https://doi.org/10.2196/46711

Wood, L., and Proudlove, N. (2022). 'Doing today's work today: real-time data recording and rolling audit in an IVF clinic'. *BMJ Open Quality*, *11*(3). DOI: https://doi.org/10.1136/bmjoq-2022-001943

Wong, G., Westhorp, G., Manzano, A., Greenhalgh, J., Jagosh, J., and Greenhalgh, T. (2016) 'RAMESES II reporting standards for realist evaluations', *BMC Medicine,* 14, 96. DOI: https://doi.org/10.1186/s12916-016-0643-1

World Economic Forum. (2024). *Transforming Healthcare: Creating a Sustainable, Accessible, and Resilient Healthcare System*. World Economic Forum. Available at: https://www3.weforum.org/docs/WEF_Transforming_Healthcare_2024.pdf [Accessed 26 October 2024]

World Health Organization, (2015). *Interim report: placing people and communities at the centre of health services: WHO global strategy on integrated people-centred health services 2016-2026: executive summary* (No. WHO/HIS/SDS/2015.20). Available at: https://interprofessional.global/wp-content/uploads/2019/11/WHO-2015-Global-strategy-on-integrated-people-centred-health-services-2016-2026.pdf [Accessed 26 October 2024]

Wu, M. J., Zhao, K., and Fils-Aime, F. (2022). 'Response rates of online surveys in published research: A meta-analysis'. *Computers in Human Behavior Reports*, *7*. DOI: https://doi.org/10.1016/j.chbr.2022.100206

Yin, R. K. (2014). *Case Study Research: Design and Methods*. Sage Publications.

Zhang, J., Morley, J., Gallifant, J., Oddy, C., Teo, J.T., Ashrafian, H., Delaney, B., and Darzi, A. (2023). 'Mapping and evaluating national data flows: transparency, privacy, and guiding infrastructural transformation'. In *The Lancet Digital Health.* 5(10), pp.e737–e748. Elsevier Ltd. DOI: https://doi.org/10.1016/S2589-7500(23)00157-7

Zhang, J., Ashrafian, H., Delaney, B., and Darzi, A. (2023a). 'Impact of primary to secondary care data sharing on care quality in NHS England hospitals'. *Npj Digital Medicine*, 6(1). DOI: https://doi.org/10.1038/s41746-023-00891-y

Zhang, J., Sood, H., Harrison, O.T., Horner, B., Sharma, N., and Budhdeo, S. (2020). ‘Interoperability in NHS hospitals must be improved: the Care Quality Commission should be a key actor in this process’. In *Journal of the Royal Society of Medicine*. 113(3), pp.101–104. SAGE Publications Ltd. DOI: https://doi.org/10.1177/0141076819894664

Appendices

Appendix 1. HIMSS Electronic Medical Records Adoption Model (EMRAM) Model Stages (HIMSS, 2022)

Stage 7	• **Improve Patient Safety** - Evaluate and improve patient safety at your acute facilities by optimizing your EMR implementation to provide access to critical information when and where clinicians need it. • **Increase Patient Satisfaction** - Reduce time and errors in care delivery and see increased patient satisfaction. Enhance care delivery by having the right information at the right time for both the patient and the clinician. Support Clinicians - An effective EMR is one that is designed for the distinct uses of the clinicians who work with it. The EMRAM ensures the workflow and content in the digital tool meets the needs of the clinical teams while monitoring compliance with approved standards. • **Secure Data** - Effective hospital policies and governance for data security are critical components of a successful EMR implementation. The EMRAM guides the organization in policymaking for the appropriate use of the data the EMR stores and the level of access available to clinician teams and others within the organization.
Stage 6	• HIE enables Structured or Coded Data from external sources to be integrated into the Clinical Data Repository, an icon is used to indicate external data is available for clinician teams. • Patient satisfaction is measured using automated digital tools (e.g., devices, apps, web based portal) to profile the patient experience during hospitalization. Patients are able to access a subset of clinical data: Discharge status, education. Patients can submit self-reported outcomes data and are able to update their personal health status data online (e.g., medication compliance, self-risk assessment, upload medically relevant images), and report progress with care pathways or therapies (e.g., patients can document that they performed the prescribed or recommended action). • Analytics governance actively assesses outcomes data for needed changes, available in a common repository. Rates of adverse events (medical error, all types) /patient day (inpatients), and trending over a 12-month period. Rates of adverse events associated with high-risk care processes are tracked for the following: anticoagulation errors/adverse events, insulin errors/adverse events, conscious sedation errors/adverse events, incorrect blood product use, antidote use, Intravenous medication errors/adverse events. Rates of “Never Events” across the organization, and trend over a 12-month period. Medical devices are integrated into EMR (e.g., monitoring devices) in ICUs. • Clinical Governance Committee is formed and works closely with Data Governance to optimize capture of clinical care outcomes to identify quality and safety priorities

Stage 5	• More the 75 percent of clinical documentation is created using online tools and available to the clinical team members in the Clinical Data Repository. • More than 25 percent of medications are electronically identified at the bedside. Tracking timeliness of nursing care (e.g., timed medication orders) to examine workflow efficiency and productivity, and care quality. The electronic system continuously monitors at least one patient condition, such as vital signs or laboratory values, in order to automatically alert care team members about risks of patient health status deterioration. • HIE enables documents from external sources to be integrated into the Clinical Data Repository, an icon is used to indicate external data is available for clinician teams. • Emergency situations/cases have a defined documentation strategy to verify accuracy of care interventions. Secure texting in place between clinicians to enable team communications and collaboration. Bidirectional interfaces are in place to external HIE for both inbound and outbound updates. Care teams offer/provide telehealth (e.g., telephone based monitoring, care navigation) to support patient surveillance, consultation and treatment both prior to admission and post discharge. • Clinical governance assesses effectiveness of CPOE and approves changes to workflow to improve staff efficiencies. Clinical outcome targets are measured and used to prioritize changes. Patient satisfaction targets inform service improvement programs in each clinical area e.g., surgery, medicine, inpatient, outpatient. • Data analytics governance has defined outcomes data captured - numerators, denominators, multi-source data points resolved
Stage 4	• More than 50 percent of all medical orders are placed using Computerized Practitioner Order Entry (CPOE) by any clinician licensed to create orders. CPOE is supported by a clinical decision support (CDS) rules engine for rudimentary conflict checking, and orders are added to the nursing and CDR environment. Clinical outcome targets are identified in selected areas e.g., disease groups, clinical procedures and operational services. • More than 50% of all clinical documentation is created using online tools and available to the clinical team members in the Clinical Data Repository. Where publicly available, clinicians have access to a national or regional patient database to support decision making (e.g., medications, images, immunizations, lab results, etc.). • During EMR downtimes, clinicians have access to patient allergies, problem/diagnosis list, medications, and lab results. • Patient satisfaction targets are identified for each clinical program, and/or for specific patient populations segments e.g., inpatients, day cases, outpatients, emergency room. • Clinical governance committee assesses effectiveness of computerized orders and order sets e.g., efficacy, usability and compliance.

Stage 3	• More than 25 percent of clinical documentation is created using online tools and available to the clinical team members in the Clinical Data Repository. An Electronic Medication Administration Record application (eMAR) is implemented for all medications. • Access to external data sources (e.g., educational materials for clinician reference, regional or national systems, registries, immunizations and vaccination systems), available to clinicians. Clinicians have remote access to patient records (if allowed by policy). • Infrastructure for bedside point of care scanning is planned or is installed in some but not all locations. Clinical governance committee has a process to review and update Clinical Decision Support opportunities. • Role Based Access Control (manage appropriate access based on staff role). • Scheduled outages are communicated including areas impacted and duration. Preparation plans are defined for moving into downtime and recovery.
Stage 2	• Clinicians have access to CDR for results review. A clinical governance committee is formed to begin defining workflow and Clinical Decision Support objectives. Policy and procedures for bedside scanning, specimen collection, blood administration and scanning of clinically relevant paper are in place. Appropriate use, security training policies are defined. • IT Change Management includes a review of proposed changes and have a rollback plan before the change is made. Applications are prioritized by criticality (high, medium, low or similar) for business continuity.
Stage 1	• All major ancillary clinical systems are installed. The Clinical Data Repository has more than 90% of lab data available for trending analysis and Clinical Decision Support. In addition, the CDR has more than 90% of all DICOM and non-DICOM images stored in a patient centric manner and available across the hospital network with a minimum of 25% available to clinicians online. • Business Resilience plans are in place for each ancillary system describing how to communicate the scope and duration of the outages and the process to distribute results as needed.
Stage 0	• The organization has not installed all of the key ancillary department systems (laboratory, pharmacy, cardiology, radiology, etc)

Appendix 2. Literature Review Appraisal

Table mapping reviewed studies against eight thematic categories under three overarching areas.

Authors	Technical Factors			People and Organisational Factors			Macroenvironmental Factors	
	System Architecture	Data Standards and Protocols	Data Quality, Consistency and Presentation	Leadership and Engagement	Workforce Expertise, Training and Digital Literacy	Operational Culture and Capacity Challenges	**Governance, Privacy and Ethical Considerations**	**National Policy, Aligned Governmental/ Organisational Directives**
Adaba & Kebebew (2017)	X				X			
Alhmoud *et al.* (2022)					X			
Banerjee *et al.* (2018)	X		X	X				
Beasant *et al.* (2023)			X				X	
Camacho *et al.* (2024)	X			X		X		
Catlow *et al.* (2024)				X	X			
Chu *et al* (2022)					X	X		
Crowther *et al.* (2022)				X		X		
Cushnan et al. (2021)	X	X				X	X	X
Davidson *et al.* (2022)	X				X	X		
Dawoodbhoy *et al*. (2021)			X				X	
Elizondo (2024)							X	X
Golinelli *et al.* (2018)							X	
Herlitz *et al.* (2023)	X				X			
Jager *et al.* (2023)		X	X	X				
Jones *et al.* (2022)							X	
Kapadi *et al*. (2024)	X	X	X					X
Khanbhai *et al.* (2021)			X	X	X	X		
Li *et al.* (2023)	X	X	X			X		X
Litchfield et al. (2022)	X			X			X	X
McGowan et al. (2024)	X							
McVey et al. (2021)	X	X	X					
Mitchell (2020)		X						X
Morris et al. (2023)	X			X			X	X
Naqvi et al. (2019)	X					X		
Pankhurst et al. (2021)				X				
Pope et al (2024)	X			X	X		X	
Schmitt et al. (2023)							X	X
Sharma et al (2023)	X			X				X
Sheikh et al (2021)		X					X	X
Smith et al (2020)	X	X	X				X	
Sullivan et al (2023)	X	X			X		X	X
Symons et al (2019)	X							
Thimbleby (2022)		X			X			
Trivedi et al (2023)		X		X				
Vezyridis & Timmons (2019)							X	
Wain (2021)	X				X			
Warren et al (2019)	X							
Wilson et al (2023)			X	X	X	X	X	
Wood & Proudlove (2022)	X			X		X		
Xiao et al (2022)	X	X						
Zhang et al (2023)	X					X	X	
Zhang et al (2023a)	X						X	X
	24	**12**	**10**	**14**	**12**	**11**	**17**	**12**

Table summarising reviewed literature

Author (Year)	Title of Paper	Method of Research	Main Findings	Additional Appraisal Summary
Adaba and Kebebew (2017)	Improving a health information system for real-time data entries: An action research project using socio-technical systems theory.	Action research (AR) informed by socio-technical systems theory, conducted within the theatre department of an NHS England Trust. Two iterative feedback cycles with 31 qualitative interviews of 19 staff who were part of AR project team, staff observations staff over several days' phases in addition to document reviews.	Existing PAS design unsuitable. Collaborative co-design fostered ownership in the re-design process and resulted in a system which was more acceptable to end-users. Requirement to address organisational cultural issues for example training on new technology as well as education on data quality for example how reliable and timely data can be of benefit. Real-time data entry reduced staff workload, mitigated errors, and improved data quality, addressing frustrations and inefficiencies stemming from delayed data recording.	Multi-method of data gathering used to triangulate findings. Informed by STS theory for collaboration. Context-specific and potentially not generalisable. Ethical approval from health trust mentioned.
Alhmoud *et al.* (2022)	Evaluating a novel, integrative dashboard for health professionals' performance in managing deteriorating patients	Mixed methods with pre- and post-implementation data from clinical performance metrics using PDSA Model. Covering 5 hospitals within the UK. Purposive sample of 3 staff for virtual face-to-face interviews.	Demonstrates the potential of dashboards in enhancing decision-making and patient outcomes in critical situations, found improvements in data sources. Clinicians need training and guidance on using digital solutions to enhance functionality	Practical but constrained by small qualitative sample size; needs broader validation and testing of generalisability. While the direct focus isn't interoperability, it explores practical tools for real-time data use. Ethically approved by HRA.
Banerjee *et al.* (2018)	Can the NHS be a learning healthcare system in the age of digital technology?	Conceptual analysis leveraging policy reports, NHS case studies, and literature on digital transformation in healthcare systems. Influenced by 3 workshops in 2016 for conferences in Oxford, Swansea and Barcelona.	Explores challenges and opportunities for the NHS to adopt a digital-first approach to become a learning healthcare system.	Well-argued but anecdotal; lacks supporting evidence from current NHS case studies.
Beasant *et al.* (2023)	Flash glucose monitoring in young people with type 1 diabetes - A qualitative study of young people, parents, and health professionals	Semi-structured virtual interviews with 34 participants, including young patients, parents, and healthcare professionals, analysed and mapped onto normalisation process theory for thematic analysis. Participants were purposively sampled, but based on social medica volunteers and 1 NHS diabetic centre in South West England.	Having real-time data for flash glucose monitoring can empower and improve quality of life for young patients and their families. Health professionals were enthusiastic about the technology and supportive of its continued use. Concerns were raised into potential of widening health inequalities for patients without technology, accuracy of some of the results and the overwhelming nature of having the continuous data.	Rich qualitative insights but focus is narrow (specific device use), and while cultural barriers are discussed, the study doesn't address interoperability system-wide. Study could benefit from quantitative validation. Ethically approved by University of Bristol.
Camacho *et al.* (2024)	Estimating the impact on patient safety of enabling the digital transfer of patients' prescription information in the English NHS	Using published literature and expert opinion, used a probabilistic mathematical model on the data of medication errors to estimate those that led to ham and effect of information sharing. PICO Framework for analysis	Estimated reduction in medication errors would reduce patient episodes experiencing harm and admission costs. Interoperable solutions could reduce medication errors by 30-50%. Doesn't consider cost of setting up interoperable systems but notes could be offset by potential estimated savings. Interoperable system alone is not enough, requires buy-in and positive staff relationships and culture of agreement in how it is used.	Based on a number of assumptions to generate indicative estimates. Lack of data from certain transfer settings and wider concepts of harm.
Catlow *et al.* (2024)	Nationally Automated Colonoscopy Performance Feedback Increases Polyp Detection: The NED APRIQOT Randomized Controlled Trial	Randomised controlled trial carried out across 36 centres, with 541 endoscopists. Pre-intervention procedures were used to establish a baseline, the intervention period ran for 9 months and post-intervention for a further 3 go establish maintenance of intervention effects. Randomised 1:1 intervention or control across the centres.	Automated performance feedback did not increase case-mix detection performance significantly, but did significantly improve traditional polyp detection rates. Engaged clinicians benefitted the most. Performance improvements weren't sustained postintervention suggesting ongoing feedback is required.	Large RCT. Robust data analysis, equitable pre and post intervention data sets. Limited to colonoscopy; broader application needs testing.
Chu *et al.* (2022)	Learning from electronic prescribing errors: A mixed methods study of junior doctors' perceptions of training and individualised feedback data.	Sequential mixed methods approach including surveys (25 out of 89 responses) which then informed semi structured interviews (5 participants) and focus group discussions (7 participants) across 3 hospitals in 1 NHS trust where all junior doctors were invited to participate.	That current feedback was lacking and informal, but the junior doctors displayed readiness to receive real-time feedback data but in manageable amounts and motivational. Requirement for more training was also requested but lack of agreement in how the data could best be presented to enable learning.	Both qualitative and quantitative analysis provides insights into perceptions of what clinicians want from electronic prescribing systems. However, it is context specific. Had a relatively low response rate from a single organisation but note saturation was reached. Registered as a service evaluation but conducted ethically.

Crowther *et al.* (2022)	Towards codesign in respiratory care: development of an implementation-ready intervention to improve guideline-adherent adult asthma care across primary and secondary care settings (The SENTINEL Project)	Reflective perspective on adapted experience-based co-design process used to develop intervention made of 5 pillars in Hull and East Yorkshire. Patients and staff were consulted through separate exploratory meetings (7 clinical and 3 patient meetings), 1 staff feedback event and 1 co-design event. Total of 15 participants. 6 GP practices were invited to participate.	Co-design enabled evidence for acceptance of proposed ideas, and development of practices provides an understanding of their acceptability and implementation obstacles. Concerns relating to additional workload and repetitive nature of conducting reviews impacting the quality was raised as well as concern over real-time data sharing highlighting poor practices. However real-time data was still preferable to compare, support and learn from practices. Implementation will be different due to different processes and demographics and change will take time.	Qualitative data identified novel viewpoints and included mix of patients and staff despite limited participation. Limited scalability analysis. A critical reflection, to an extent is hypothetical, as intervention still requires evaluation, long-term impact or wider implementation not assessed.
Cushnan *et al.* (2021)	Towards nationally curated data archives for clinical radiology image analysis at scale: Learnings from national data collection in response to a pandemic	Case study of the National COVID-19 Chest Imaging Database (NCCID) data collection initiative during COVID-19.	Empasis on future opportunities and learning across multiple themes; Information governance processes should be clarified and standardised to reduce barriers in participation. Different regulations delay facilitation of national data collection. Also needs to ensure patient and public engagement due to concerns over data storage and use so to provide ethical assurance. Collaboration and linking databases improve the quality and completeness of data collected. National co-ordination is required to overcome fragmented data landscape. Infrastructure such as trusted research environments (TREs) enabled secure access and analysis. TRE environments require users to have high performing compute for analysis. Automation enables mass data collection reducing manual burden on staff – supported by data standards but difficult due to heterogeneity of systems so reliant on national-linkage. Funding to support data collection and infrastructure engineering is key to sustaining them.	Focus covered wide range of factors. Authors were consultants on the NCCID, large number of collaborators. Lacks details of empirical evidence.
Davidson *et al.* (2022)	Requirements for a Bespoke Intensive Care Unit Dashboard in Response to the COVID-19 Pandemic: Semistructured Interview Study	Semi-structured interviews with 6 ICU staff across 2 hospitals to identify dashboard features and usability requirements. Used NVivo for qualitative coding and thematic analysis.	ICU dashboards must balance real-time data visualisation with usability but are perceived to reduce staff cognitive load and clinical errors. Dashboards need to be flexible, mobile, and customisable. Data-entry needs to be consistent. Staff engagement in design increases adoption but requires iterative refinement.	Strong alignment with technical and cultural themes. Wider ICU testing required but also the generalisability to non-pandemic and non-emergency contexts is unclear. Small sample size but across multiple clinical roles.
Dawoodbhoy *et al*. (2021)	AI in patient flow: applications of artificial intelligence to improve patient flow in NHS acute mental health inpatient units	Narrative literature review of 72 articles, followed by 20 semi-structured interviews with two selected groups of AI (n=11) and MH (n=9) experts, in addition to snowball sampling. Pilot interview was carried out first with a consultant to validate structure and design. Thematic analysis used to analyse data and identify 5 final themes.	Common themes in patient flow issues. AI could improve patient flow by streamlining admin tasks and optimisation of resources. Real-time data could support clinical decision-making. AI requires collaborative investment and infrastructure. Concerns around data-use, regulation and transparency remain. Frustrations around fragment systems.	Detailed themes and components. Small participant sample didn't cover range a roles. Lack of quantitative data prevented triangulation.
Elizondo (2024)	Governance intricacies in implementing regional shared care records: A qualitative study in the national health service, England	Longitudinal (5 year) period of interoperability program in Southwest England. Multi-method qualitative study included 50 interviews with stakeholders involved in implementing regional care records, 6 governance meeting observations (12 hours) and document analysis. Pilot interviews with 5 participants to refine and validate and additional follow up interview with 6 participants. Interviews were analysed using NVivo to identify themes, and data saturation reached at 40 interviews.	Governance challenges include varying stakeholder priorities, insufficient policy alignment, and data-sharing resistance. Necessity for continuous and iterative discussion between stakeholders to establish effective governance of integrated healthcare infrastructures. Tensions exist between regional and national objectives. Concerns over accountability of shared care records and balance of patient privacy and seamless data sharing. Technical integration of diverse systems is a challenge, variations in IT infrastructure exacerbate issues.	In depth methodology and rich qualitative method. Strong focus on governance barriers but lacks technical depth. The study may have focused more on high-level decision-makers, potentially underrepresenting frontline staff perspectives.
Golinelli *et al.* (2018)	Transferring Health Big Data within the European Legal Framework: What Role for National Healthcare Services?	Analysis and interpretation of European Legal Framework for the transfer and re-use of data.	Highlights some historical instances where large scale data transfer faced criticism, and therefore considers the legal and ethical framework governing health organisations and the need for updated, adequate and innovative data policies given the rise in digital opportunities such as big data. Focuses on aspects that are less obvious and not regulated under GDPR such as the public authorities position to decide on the selection and transfer of data to third parties' data. Suggest NHS act in a structured and consistent manner, by adopting stronger policies regarding health data creation, analysis and trade, contributing to their own financial sustainability	Strong governance focus, but other dimensions are not addressed. It is primarily based on perception of policy rather than based on any empirical study. .

Herlitz *et al.* (2023)	Patient and staff experiences of using technology-enabled and analogue models of remote home monitoring for COVID-19 in England: A mixed-method evaluation	Mixed-methods evaluation including surveys (1069 patients and carers, and 292 staff) and interviews (61 patients and carers, 58 staff) covering 21 sites across England. Compared tech-enabled to analogue models of monitoring. NVivo used for qualitative analysis and SPSS for statistical analysis to triangulate quantitative survey data.	Patients with tech-enabled monitoring tended to be younger with higher level of education. Staff found adoption of mixed models to be beneficial and supported clinical decision making but suggested functionality improvements. Human contact was still important and accommodated different needs, with some patients more likely to relay symptoms over the phone. Interoperability and usability were important with efficiency gains from real=time data, and tech-models considered more scalable. Barriers to patient submitting data were mainly related to ill health or forgetting, and reliance on family member to relay readings especially in those lacking digital skills. Additional support for patients, training and digital literacy encourage among staff could widen participation.	Provides insights into barriers and enablers of remote monitoring from staff and patients using mixed concurrent methodology. Underpinned by the Planning and Evaluation Remote Consultation Services (PERCS) Framework with underpinning principles of healthcare quality and ethics, and detailed methodology. Cost implications weren't measured, low response rates not necessarily representative.
Jager *et al.* (2023)	The usage of data in NHS primary care commissioning: a realist evaluation.	Realist evaluation to create a context-mechanism-outcome configuration for data usage in primary care following the RAMESES reporting standard. Analysis of theories, and qualitative interviews (n=30) from purposive sample based on geographic location and commissioning job titles (n=23) and academic experts in evidence-based commissioning (n=7). Additionally included analysis and transcribing of recorded CCG meetings (n=51). Used a constant comparison method to compare and contrast findings and Pawson's method of reconciliation if contradictory.	Commissioners would be more inclined to use data if it indicated a potential cost saving. Sometimes data was believed to be incorrect due to the way it was coded or reported. Data errors were presumed to be because of gaps in knowledge or capabilities. Trusted and credible sources of data made its use more likely. When commissioners perceived that data was imposed on them or presented in an unappealing manner, they sometimes disengaged or looked for flaws in the data. Lack of capacity for data analysis meant it was not used, with belief that external support lacked NHS-specific knowledge and expertise, so analysis perceived as less valid. Combined datasets and sharing data in a visually appealing manner increased engagement – those which weren't interoperable with consistent definitions were difficult to use.	Framework and theory-based approach for methodology and analysis. Large amount of primary data analysed and found consistency across data sources. Supplementary information accessible. Ethically approved. Would be useful to also collect perspectives of primary care providers and patients on data usage and sharing in commissioning processes.
Jones *et al.* (2022)	Public opinion on sharing data from health services for clinical and research purposes without explicit consent: An anonymous online survey in the UK.	Anonymous online survey open to all UK residents, recruited from health sites and media – 29,275 participants. Survey was co-designed with patients and carers from the research advisory group. Questions sought views on different types of data such as physical or mental health, the extent to which data was identifiable or linkable, the destination which data was shared and the structure such as coded or free-text and randomised the framing of questions to risk vs benefit perspectives. Statistical analysis was conducted on Likert scale questions and thematic analysis on free-text questions.	Most of the public supported data sharing for direct clinical care without explicit consent. 30% believed there was already free sharing of identifiable data across the NHS. There was also willingness to share both physical (slightly more so) and mental health data for clinical purposes to a local NHS service, but still supported broader geographically. However, there was strong opposition for sharing data to other companies and sharing of structured rather than free-text data was preferred. De--identified linkage to non-health data as generally supported.	Online sampling methodology with many public responses provides applicable results across multiple geographies and demographics. Potentially biased towards internet users, but questions were co-designed. Descriptive statistics and demographic stratification provided insight into variations in attitudes across different groups.
Kapadi *et al.* (2024)	Feasibility of implementing a rapid-learning methodology to inform radiotherapy treatments: Key professional stakeholders' views.	Qualitative virtual semi-structured interviews were conducted with 23 radiotherapy stakeholders from 5 geographically and environmentally diverse UK radiotherapy centres. Purposive sampling was used following consultation with clinical leads. Interviews were not equitable between the sites but carried out until saturation reached. NVivo thematic analysis reviewed by multiple researchers produced 4 themes.	Rapid learning was supported as a gold standard and robust method for addressing evidence gaps and complementing practice. Effectiveness of rapid-learning was viewed as being dependent on having quality data, and variable set-up between cancer centres being a challenge for potential implementation. Scepticism was raised over accuracy and robustness of datasets – noting issues of incomplete and missing datasets, as well as different formats. Accessibility and digital literacy of patients was raised as a concern for quality of PROMS data if it was collected at all. Importance was given to integrating data, data standardisation and accessibility through developing informatic infrastructures but would require time, investment and organisational commitment. It was felt national-level guidance would be required to inform and support implementation, as was method clarity, educational support and training for staff.	Perspectives sought from diverse range of radiotherapy stakeholders across different cancer centres. Deeper variance analysis could have been beneficial as would having an equitable number of interviews across each site.
Khanbhai *et al.* (2021)	Identifying factors that promote and limit the effective use of real-time patient experience feedback: A mixed-methods study in secondary care.	Semi-structured interviews (n=15) with healthcare staff and stakeholders within a large London NHS Trust until saturation met. Purposeful sampling of staff within patient experience team as well as clinical staff. Transcribed verbatim and reviewed by participant before thematic analysis using NVivo following the framework method.	Identifies concerns in using real-time feedback in secondary care due to inefficient flow of data and lack of time to analyse reports due because of the volume of free-text data. Current process is unstructured, data is sent to NHS England monthly and limited analysis and visualisation carried out by an external provider. Capacity and resource concerns restrict regular consistent collection and relied on manual transference of paper feedback. Multiple formats of data and staff have insufficient access or training to understand data presented therefore are not engaged. Suggestions are that structural/organisational factors be tackled alongside micro-level factors with strategic focus prioritising data utilisation over collection, and that data be presented accessibly, straightforward and engaging.	Iterative qualitative analysis method. Collaborative group with patient and public involvement. Ethically approved. Restricted specifically to Friends and Family Test feedback initiative and stakeholders involved in its reposting so a wider sample would be beneficial.

Li *et al.* (2023)	Perceptions of chief clinical information officers on the state of electronic health records systems interoperability in NHS England: a qualitative interview study.	Qualitative online semi-structured interviews (n=15) carried out with chief clinical information officers in NHS England using convenience sampling and identification through a Digital Academia network followed by snowball sampling. Verbatim transcriptions were then thematically analysed by two researchers.	There were differing perceptions around what interoperability meant. Generally, perceptions were that limited EHR interoperability contributed to inability to access and transfer data to a central source especially in terms of sharing between primary and secondary care, and certain clinical specialties. This data fragmentation, as a result is perceived to cause a lack of clarity and impact on patient safety, suboptimal care coordination, duplication of efforts, and defensive practice. Barriers also included system usability difficulties, and institutional, data management, and financial-related challenges. Identified need for focus on mandating data standards, user-centred design, greater patient involvement, and encouraging inter-organisational collaboration. Facilitators identified were themes of recognition of need and possibility amongst healthcare workers, expectation from patients, and centralised nature, strategic levers, and national oversight. Barriers identified were themes systems usability, institutional, data-related, and vendor/finance-related barriers. Requirements identified were themes the need for common data standards, to address existing EHR systems usability issues, to incorporate patients in accessing their clinical records, and the need to promote greater interorganisational collaboration.	Methodology targets perceptions from a relevant technical group. Sample size and representativeness across multiple NHS regions but not all regions were represented and relatively small sample size. Used COREQ best practice guides.
Litchfield *et al.* (2022)	The move towards integrated care: Lessons learnt from managing patients with multiple morbidities in the UK.	Qualitative study – 28 semi-structured interviews from both managerial and clinical roles. Snowball sampling with initial participants relevant to the researchers from West Midlands and Southeast UK. Data analysis carried us using NVivo, presented in SELFIE Framework, and triangulated with secondary data.	Focused on integrated working. Continuing challenges with accuracy and timeliness of communication between primary and secondary care. Shared responsibility and collaborative leadership are required for truly integrated care, also by way of training of workforce. Lack of data sharing and interoperability is a key barrier to integrated care and prevents timely access of the patient record. No centralised or unified selection of interoperable software – procurement of systems undertaken independently based on local needs and priorities and a more strategic approach is required. GDPR adds complexity to data use for integrated care, described as a massive barrier which impacts evidence-based integrated care.	Clear aim and data collected was appropriate. The qualitative approach had a relevant sampling strategy and appropriate measurements, with options of in-person and virtual interviews.
McGowan *et al.* (2024)	The Views and Experiences of Integrated Care System Commissioners About the Adoption and Implementation of Virtual Wards in England: Qualitative Exploration Study.	Qualitative semi-structured online interviews with 20 commissioners across NHS England (purposive sampled with subsequent snowball sampling). Open-ended questions co-developed. Thematic analysis using NVivo in second stage, with framework approach, following COREQ guidelines.	Focused on Virtual Ward implementation. Participants implied limited evidence for advantages of using technology, particularly to remotely monitoring, and felt it could actually lead to uncertainty and nervousness among clinical staff members and resistance to change. It was felt that it was mandated to be a technology-led solutions rather than led based on pathway needs.	Followed appropriate frameworks and approach to data collection and analysis. Ethical approval
McVey *et al.* (2021)	Hidden labour: the skilful work of clinical audit data collection and its implications for secondary use of data via integrated health IT.	Qualitative 54 semi-structured interviews with staff in five English NHS hospitals, including 20 staff involved substantively with audit data collection. Identified purposive and snowball methods, also ethnographic observations took place on wards, in 'back offices' and meetings (102 h). Thematic analysis informed by framework method, and NVivo, synthesised narrative. Was part of a wider study. Interviews were 2017-18 and observation 2019-20	Large volume and complexity of labour to obtain required data, which was diverse and distributed. Observed varying methods of how staff spent time gathering and checking data from a range of sources, often copying information from digital systems to paper forms, before rekeying it into local databases or web portal. Difficulties in providing systems that could integrate appeared to be linked to resource limitations in the hospitals and the dated technology used. Staff did not always trust the quality of data in their hospitals' digital systems thus the use of paper forms -development of deep understanding of the data built and maintained trust in its quality. Suggested that automating the most labour-intensive parts of data integration needs to balance with designing interfaces that empower users to assess integration outcomes and to use their own skill and ingenuity to address problems like data quality	Part of a larger study. Gap between interviews and observations. Ethical approval

Mitchell (2020)	A NICE perspective on computable biomedical knowledge.	Perspective on Mobilising Computable Biomedical Knowledge (MCBK) from a NICE perspective and based on MCBK UK Workshop discussions from 2019. Agency for Healthcare Research and Quality (AHRQ) 4 levels of knowledge used to draw conclusions, based on work/interviews already done by NICE.	NICE is mainly at level 1 of the AHRQ knowledge hierarchy with content being narrative and unstructured. Adding structure and standard clinical codes to NICE guidelines found to have significant methodological implications and an impact step to develop guidance and an aim for NICE, due to understanding that structured data and knowledge enable the concepts of a continually learning healthcare system. Technically there are challenges of agreeing which existing formalisms, coding and information standards for representing clinical knowledge could be used to share knowledge. Practically there are questions about who has responsibility for applying codes and interpreting guidance. Lack of clarity in what levels of structure and coding are needed.	Perspective discussion, however NICE is known for evidence-based and rigorous methods.
Morris *et al.* (2023)	Moving from development to implementation of digital innovations within the NHS: myHealthE, a remote monitoring system for tracking patient outcomes in child and adolescent mental health services.	Retrospective summary of experience for developing and implementing myHealthE for CAMHS by a Kings Health Partnership Group. Specification developed by MHE Team and digital commercial supplier to help design. Implementation was a 12-week single-blinded parallel group pilot RCT separate study paper. CFIR framework used to pre-determine impact of implementation.	Anticipated challenges included Legacy systems, resource barriers and lack of clinical engagement. Overall took 23 months to build - found there was no API in main EPR to allow data transfer so looked to use NDL Automate Robotics Processing Application. Lack of guidelines for development of new cloud infrastructure and relied on inter-departmental relationships to develop a standard pipeline. Lack of data security and IG, Insufficient resources to develop a protype were also barriers and required recruitment of Centre for Translation Informatics (CTI) operations board for IG and technical Officers. Clinical disengagement arose from project delays.	
Naqvi *et al.* (2019)	The general practice perspective on barriers to integration between primary and social care: A London, United Kingdom-based qualitative interview study.	Qualitative semi-structured interviews with 18 GPs and 7 practice managers within London. Constructivist paradigm, purposive sampling - open questions, with two pilot interviews to ensure refined questions. Transcribed verbatim. Thematic analysis - based on Clarke and Braun 6-Stage Framework and handwritten methods to generate codes per transcript line. Findings checked with participants to allow for feedback. Findings reported using SRQR checklist.	Identified three themes of barriers to integration: accessing social services, interprofessional relationships and infrastructure. Issues with contacting staff from other sectors creates delays in referrals for patient care and perpetuates existing logistical challenges. Hostile working culture between sectors has resulted in silo working mentalities. Staff are overworked, multidisciplinary team meetings are inefficient, and the poor relationships across sectors cause a diffusion of responsibility. Lack of interoperability between information systems, lack of pooled budgets and misaligned incentives between managerial staff compound the infrastructural divide between both sectors. Lack of interoperability between systems means communication is limited to emails and phone calls, which often leads to patient confidentiality issues and delays.	Appropriate methodology for aim - mixed face-to-face and online interviews based on preference. Ethical approval.
Pankhurst *et al.* (2021)	Rapid adaptation of a local healthcare digital system to COVID-19: The experience in Birmingham (UK).	Case study experience within UHB of rapid digital development during the COVID-19 Pandemic	Integration between the EHR and laboratory systems provided up-to-date information to help identify COVID-19 patients. Required EHR system changes to be implemented quickly based on evidence and clinical need. Staff engagement help to maintain buy-in. Avoidance of support alert and data entry fatigue could result in important information being missed so need to be clinically important.	Data sources in methodology not fully detailed.
Pope *et al.* (2024)	Real-world learnings for digital health industry–NHS collaboration: Life sciences vision in action.	Discussion of partnership formed in 2021 between Great Ormond Street Hospital and pharmaceutical company and their joint learnings from first 2 years collaboration.	Limitations in IT infrastructure, which has received significant investment but lacking in-house cloud infrastructure, example of manual loading due to incompatibility. Specific technology roles are needed that liaise with existing NHS ICT teams, and understand NHS regulations and approvals needed to deliver the vision supported by cloud computing - proposed as an important, often-overlooked dimension.	Governance of partnership and DPIA discussed highlights appropriate and legally sensitive collaboration.
Schmitt *et al.* (2023)	What does it take to create a European Health Data Space?	Based on theoretical insight, elaborates on the European Union's EHDS proposal. Discussions from the International expert workshop	Emphasises the need for a flexible governance framework and the creation of use cases and trust to demonstrate the secondary use of electronic health records (EHR) data. Need to consider maturity levels of health data infrastructures in different countries and context-specific factors. Digital health literacy of healthcare providers should be ensured, and offered incentives for providing high-quality and structured data. Workshop participants suggested that reliable clinical data could be collected for research using interoperable data infrastructures without becoming a burden. Strict data protection rules and different interpretations of national and European laws build barriers to health data re-use. .	In a broader EU space, macroenvironmental factors - cross-country comparisons used in the study, base on data from a variety of reliable sources, including reports from EU bodies, health data initiatives, and governmental studies

Sharma *et al.* (2022)	The evaluation of digital transformation in renal transplantation in the United Kingdom: A national interview study.	Qualitative and framework-informed. Structured phone interviews with transplant coordinators at all 23 transplant centres in the UK. Analysed interview data thematically and synthesised results between centres. Digital Maturity Frameworks also used to analyse data and assign maturity scores as well as inform interview topics.	IT and EHRs did not readily exchange required clinical data, a range of methods for data sharing were seen but all needed manual data entry, document scanning, or storage of paper copies in patient files. Workarounds included creation of shared drives on the network and excel/access databases. Variations of IT solutions with multiple systems needing to be logged onto highlighted as the main challenge - referrals outside of the transplant centres were manual. Lacking data strategy or formal data sharing agreements across all sites. 4 had regional interoperability via shared care records.	Appropriate methodology and lead investigator background expertise. Specific context but all sites within the UK participated. Established frameworks used to contextualise findings. Acknowledged clinical staff may not have had the technical expertise to answer some questions accurately.
Sheikh *et al.* (2021)	Health information technology and digital innovation for national learning health and care systems.	Evidence based synthesis of research and opinions. Summarises findings from various sources to make recommendations for the UK's digital health strategies.	Requirements for a bottom-up innovation, addressing front-line and patient challenges, and top-down strategies such as governmental common approaches and incentives. Poor system usability contributes to errors in the process of entering and retrieving information, communication and coordination jeopardising patient safety. Contracts not specific for interoperable requirements. Digital inclusivity is important for patients. Programs within NHS England for Digital leadership and FEDIP.	
Smith *et al.* (2020)	National Institute for Health Research Health Informatics Collaborative: Development of a pipeline to collate electronic clinical data for viral hepatitis research.	Adhering to governance framework and NIHR HIC Programme and methodology across network of 25 NHS Trusts. Specifically Viral Hepatitis Theme led by Oxford NHS Trust for 5 participating sites. First clinical leads defined data fields, required to answer the initial academic questions posed by the clinical and scientific leads across centres and sample case report forms submitted to OpenClinica to refine schema definition.	Pre-existing data warehouses and systems were not always suitable for generated required dataset. Variety of EPRs resulting in different formats meaning data architecting was required for each unique site. Creation of NIHR HIC Viral Hepatitis Central Data Repository - some already structured, others free-text which needed manual intervention and anonymisation by using XML format. No fixed schedule for submissions. Reliance on duplicates and unique IDs to be managed by sites. Basic integrity checks were automated but needed manual checks following data load - rejects identified duplications, may require additional information to confirm submitted data validity. Resulting 32 tables with 349 data fields and 20 element types but due to differences between sites none were made mandatory. Differences in practice between clinicians and sites lead to heterogenous patterns of missing data.	More detail of numbers of clinicians and participating sites would be useful. Technical detail does make process reproducible and assured through the governance framework and program.
Sullivan *et al.* (2023)	Pathways to interoperable electronic patient records for health and social care: Part 1: for those involved directly in care.	Perspective summary	Interoperability has progressed. Examples for GP Patient Record Access, GP2GP programme for transferring patient data using HL7. Interoperability to progress requires understanding, investment and effective implementation at the national level. At the local level, informed adoption and effective training to facilitate a different approach to record-creation and maintenance that is aware of the need to communicate effectively with everyone who might share the record soon.	
Symons *et al.* (2019)	From EHR to PHR: Let's get the record straight.	Perspective summary	There has been growth in the number of PHR but uptake by patients and organisations is still remarkably slow and been attributed to barriers in sharing of data at a local level, clinical aversion and reluctance, patient awareness and technical integration with local information systems.	
Thimbleby (2022)	NHS Number Open Source Software: Implications for Digital Health Regulation and Development.	Perspective summary/Critical Review	relevant healthcare standards are inconsistent and written without sufficient rigor to be at all constructive for implementing digital systems. The widely recognized problems of interoperability may be traced back to diverse (and buggy) interpretations of vague standards. Calls for qualified professionals to handle digital data.	A controversial critique which raises a legitimate concern around lack of updated standards.
Trivedi *et al. (2023)*	Proof-of-concept solution to create an interoperable timeline of healthcare data.	System customised between two hospital sites which used different EHRs. Further 50 sites were artificially created with simulated data. Patient and public involvement panel provided input and feedback. Software uses FHIR-based blockchain centralised SQL server to present amalgamated data in a single view to form a clinical timeline regardless of originating EHR.	Shared patient timeline could be seen at both sites, with their own clinical narrative and the other site specific test sets. Patients could also log on and see simulated data entries and add personal input to the timeline. Patients wanted to see data in one place and write their own inputs, but had questions about data security, consent, speed and accuracy of the system - System passed security and functionality tests.	
Vezyridis and Timmons (2019)	Resisting big data exploitations in public healthcare: free riding or distributive justice?	Uses documentary analysis, actor network-theory and a previous qualitative study 2016 which collected 27 perspectives from EHR researchers, GPs and citizens who had opted out/were against the care.data programme	Support was found for the principle of re-using personal health data for wider societal benefit but was felt care.data approach was cavalier and too heavy-handed due to lack of tiered consent and met resistance as a result. Reassurance was sought that sharing data sets would be for public rather than personal or commercial use.	Theory bases approaches. Considered multiple perspectives and documentary sources. Also history of related research study around NHS patient datasets.

		through semi-structured interviews. Recruitment was through social media, newsletters and targeted emails. Interviews were recorded. Used grounded theory approach to analyse using NVivo for differences and patterns.	Vocabulary was used interchangeably and confusing. Individuals felt trust was betrayed y care.data programme. Called for collaborative governance framework where experts can provide informed opinions but the individuals having the final say.	
Wain (2021)	Does integrated health and care in the community deliver its vision? A workforce perspective.	Interpretive Design – Qualitative in-person study with 5 in-depth interviews, using open-ended questions from a purposive sample. Included 2 community nurses and 2 social workers, and 1 AHP. Colaizzi 7-Step method and descriptive phenomenology for analysis to identify 4 themes and 1 fundamental structure. Data saturation reached.	Focus on integrated care – found differing understanding on the meaning/concept. Different regulators between health and social care, different IT platforms not compatibility required and digital literacy skills. Found commonality of frustrations with lack of interoperability, and different systems open to error.	In-depth design method intended with appropriate small, localised sample. Ethical approval.
Warren *et al.* (2019)	Improving data sharing between acute hospitals in England: An overview of health record system distribution and retrospective observational analysis of inter-hospital transitions of care.	Quantitative - retrospective observational study used Hospital Episode Statistics (HES) and publicly available organisational data for Acute hospital trusts in the NHS England. To identify the frequency of use and spatial distribution of health record systems; and spatial distribution of transitions of care between health record systems. Statistics using Python, Excel and illustrated using Tableau.	152 NHS Trusts covered, showed 77% used ESRs, 23% paper records. 78% used one of 21 systems identified – 10% used multiple different EHRs and 11% in-house developed software. Given the number of patient transfers often organisations didn't share the same system - minimal coordination resulting in fragmentated patient records	Focused on Acute Care within England, which may limit generalisability to other regions and sectors. Only a quantitative perspective but statistical evidence for points covered. Local ethical approval and use of Hospital Episode Statistics data was approved by NHS Digital.
Wilson et al (2023)	Key Considerations When Developing and Implementing Digital Technology for Early Detection of Dementia-Causing Diseases Among Health Care Professionals: Qualitative Study.	18 Semi--structured interviews with 11 primary care and 7 secondary health professionals from various NHS regions - participants identified through clinical networks and snowball sampling. Piloted and refined questions before data collection - video calls were transcribed and analysed using NVivo using an inductive framework approach.	Identified themes related to clinical aspects of the health care service users as well as considerations of digital exclusion and additional needs. Health care professional themes identified concerns of overburdening the system and additional resources required. Need for engagement for the system to be accepted and usable. Output needs to be easy to interpret and integrate, and visual aids to make it understandable. Technically the system needs to evidence accuracy and validity, integrated and appraised prior to implementation.	Emphasis placed on ensuring trustworthiness and credibility in results as well as consent and ethical approval. Framework for thematic analysis evidenced.
Wood and Proudlove (2022)	Doing today's work today: real-time data recording and rolling audit in an IVF clinic.	Process-mapped processes and gathered staff views on problems and potential solutions through 4N Chart results.	Current problem identified several electronic and paper systems, with duplication of entry. Considerable delay in data entry as perceived as relatively low priority. Data not being available for the MDT reviews meant staff wasted time searching across various systems. Anticipated the main challenge would be encouraging the behavioural change of incorporating administrative data capture as part of 'real time' work	Evidence-based. Carried out PSDA cycles, SPC charts appropriate analysis method to demonstrate significance of changes.
Xiao *et al.* (2022)	Towards a systematic approach for argumentation, recommendation, and explanation in clinical decision support.	Systematic approach proposal for generic argumentation and recommendation scheme is put forward -representing clinical rationale. - using Resource Description Framework (RDF) for clinical guidelines, a rule engine developed for their interpretation, and recommendation rules represented using Semantic Web Rule Language (SWRL). Describes a case study and hypothesis testing to evaluate metrics of accuracy, variation, adherence, time, satisfaction, confidence, learning, and integration of the developed prototype CDSS.	The lack of semantic interoperation hinders the adoption of CDSSs and eventually leads to their failure. Issues such as interoperability and explain-ability burden local advocators and potential users possibly leading to the reluctance of CDSS use. Developed prototype simple and intuitive, reduces the cognitive cost. Semantic Web-orientated knowledge with RDF structures said to offer a promising, open and interoperable clinical decision support paradigm	
Zhang *et al.* (2023)	Impact of primary to secondary care data sharing on care quality in NHS England hospitals.	A retrospective longitudinal analysis across 135 Acute NHS Trusts to characterise landscape and progression of data-sharing networks in NHS England 2015-2022. Considered Acute Trust responses to a national Clinical digital maturity assessment in 2017. Carried out univariate and multivariable linear regression analyses - Covariate and outcome data taken from national aggregate secondary care datasets.	Despite the variation in vendors and technology solutions used across NHS Trusts, the ability to share primary care data with secondary care providers remained a consistent feature across all implementations. Found positive effects in emergency care pathways, where primary to secondary care data-sharing capabilities reduced A&E breach percentage and improved patient experience. - No impacts were found in analysis of patient mortality or safety incidents as outcomes. - While data-sharing shows positive effects, overall efficiency improvements depend on workforce, staffing, and capacity investments. Expanding secondary data use increases privacy risks, requiring robust governance frameworks and potential changes to the NHS opt-out system for data use.	Secondary analysis of aggregate population and organisation data from government statistical datasets, did not require ethical approval. -
Zhang *et al.* (2023a)	Mapping and evaluating national data flows: transparency, privacy, and	Focus on structured, coded patient-level records from NHS England providers that originated in primary or secondary care, excludes unstructured text records. Includes data extractors, databases, and data consumers.	Discovered a vast ecosystem of secondary uses, including 460+ non-NHS organisations, consumers also include researchers from 216 academic organisations; 143 pharmaceutical, life sciences, data analytics, and consulting companies; and 44 non-profit organisations. More than 95% of consumers collect these data indirectly	

	guiding infrastructural transformation.	Covers only systematized or single-instance data flows from April 2021 to April 2022. Excludes manually collected data and entities that only provide extraction software, storage, or backup services. Reviewed nine categories of information sources, including GDPR-related legal documents, administrative data use registers, and academic metadata registers. Sent Freedom of Information (FOI) requests to 216 secondary care trusts and 106 clinical commissioning groups to gather details on shared care record data flows. Conducted scoping reviews in the MEDLINE database to identify NHS databases and their use in observational research. Information discovery took place between April and November 2022. Synthesized findings based on NHS data strategy documents and created a typology to compare data extractors.	via data extractor intermediaries. NHS primary care data is highly fragmented across multiple databases, with duplicated NHS Digital datasets stored in various locations for onward provision. Concluding recommendation; **Enhancing Transparency** – Public access to information should not require investigative efforts. **Refining Opt-Out Mechanisms** – Patients should have control over how their data is distributed to different consumer types rather than opting out at the extraction stage. **Maximizing Existing Infrastructure** – Efforts should focus on improving and expanding the use of NHS Digital and OpenSAFELY for general applications. Mandating the use of these environments should precede the creation of new Secure Data Environments (SDEs). **Expanding Secondary Care Data Capabilities** – New data infrastructure should prioritize extracting underutilized secondary care records and increasing multimodal data availability rather than redistributing existing assets. A national federated data platform could enhance analytics across regions, improving privacy and reducing bulk transfers, but its effectiveness depends on regional participation and infrastructure readiness. **Shifting Focus from Analysis to Intervention** – Data infrastructure should emphasise real-world interventions rather than just analytical capabilities - faster data processing cycles, streamlined regulatory and governance processes, and AI-driven intelligence capabilities.	

Appendix 3. Classification of informatics study types by broad study foci, the stakeholders most concerned and the method(s) most commonly employed

Reproduced from *(Friedman et al., 2022, p.49)*

Study Type	**Study Focus**	**Resource aspect primarily studied**	**Audience/ stakeholders primarily interested in results**	**Family of methods most commonly employed**
1. Needs assessment	The problem to be solved	Need for the resource	Resource developers, funders of the resource	Qualitative
2. Design validation	Conformance of the development process to accepted standards	Design and development process	Funders of the resource, resource developers, certification agencies	Qualitative
3. Structure validation	Design of the resource in relation to its intended function	Resource static structure	Insurers, resource developers, certification agencies	Qualitative
4. Usability	Ability of users to navigate the resource to carry out key functions	Resource user interfaces	Resource developers, resource users	Both
5. Laboratory function	The potential of the resource to be beneficial	Resource performance under controlled conditions (efficacy)	Resource developers, funders, users	Quantitative
6. Field function	The potential of the resource to be beneficial in the real world	Resource performance in actual use	Resource developers, funders, users	Quantitative
7. Lab user effect	Likelihood of the resource to change user behaviour	Resource performance under controlled conditions (efficacy)	Resource developers, funders, users	Quantitative
8. Field user effect	Impact on user behaviour in the real world	Resource effectiveness	Resource users and their clients, resource purchasers and funders	Both
9. Problem impact	Effect of the resource on the health problem it was designed to solve	Resource effectiveness	The universe of stakeholders	Both
10. Organization and system	Relationships between the resource and the organizational context in which it is deployed	Broader implications of the resource	Members of the organization where study is conducted, and similar organizations and policy makers	Both

Appendix 4. Count of full time equivalent (FTE) NHS Wales staff by staff group and organisation for June 2024. (Stats Wales, 2024)

The Stats Wales data does not provide a categorical breakdown specifically for digital/technical roles, these are grouped within a broader administrative and support category, making it challenging to determine a representative sample for this subgroup.

Organisation	**All Staff**	**Staff Group**						
		Medical and dental staff	Nursing, midwifery and health visiting staff (2)	Administration and estates staff	Scientific, therapeutic and technical staff (3)	Health care assistants and other support staff (4)	Ambulance staff (5)	Other non-medical staff
Betsi Cadwaladr University LHB	**17894.29**	1027.72	8559.71	3824.07	3237.85	1225.57	1.10	18.28
Powys Teaching LHB	**2080.41**	32.58	865.85	678.95	362.38	139.39	1.00	0.27
Hywel Dda University LHB	**10280.65**	605.71	4696.40	2789.29	1806.52	362.32	9.40	11.00
Abertawe Bro Morgannwg University LHB (7)	**.**	.	.	.	.	.	.	.
Swansea Bay University LHB (7)	**12609.59**	762.33	5981.53	2555.74	2431.86	824.30	45.84	8.00
Cwm Taf University LHB (7)	**.**	.	.	.	.	.	.	.
Cwm Taf Morgannwg University LHB (7)	**11146.79**	717.54	5334.72	2457.47	1820.75	796.71	.	19.60
Aneurin Bevan University LHB	**12961.70**	822.31	6125.97	2870.57	2229.99	892.29	0.80	19.76
Cardiff and Vale University LHB	**14885.25**	979.09	6533.39	2672.44	3560.13	1125.36	7.12	7.72
Public Health Wales NHS Trust	**2282.52**	82.92	86.39	1272.73	823.61	5.07	2.00	9.81
Velindre NHS Trust (8)	**1585.78**	81.13	380.30	596.86	480.96	41.53	.	5.00
Welsh Ambulance Services NHS Trust (9)	**4029.53**	0.50	193.65	908.12	12.24	.	2908.02	7.00
Health Education and Improvement Wales	**435.62**	19.51	12.40	356.81	43.90	.	.	3.00
Digital Health & Care Wales	**1223.20**	2.46	.	1215.53	.	.	.	5.20
NHS Wales Shared Services Partnership (8)	**5397.94**	3098.70	3.10	1724.38	150.09	418.67	.	3.00

Appendix 5. Research Data Collection Mediums Strengths and Weaknesses

Reproduced from Clark *et al.* (2021) pp.227 – Number of ticks indicates the strength of the mode of administration of a questionnaire in relation to each issue.

Issues to Consider		Mode of Survey Administration				
		Structured In-Person Interview	**Structured telephone Interview**	**Postal Questionnaire**	**Email**	**Online**
Resource Issues	Is the **cost** of administration relatively low?	✓	✓✓	✓✓✓	✓✓✓	✓ (unless low cost software)
	Is the **speed** of administration relatively fast?	✓	✓✓✓	✓✓✓	✓✓✓	✓✓✓
	Is the cost of **handling** a dispersed sample relatively low?	✓ (✓✓ if clustered)	✓✓✓	✓✓✓	✓✓✓	✓✓✓
	Can the researcher **design** a questionnaire without needing much technical expertise?	✓✓✓	✓✓✓	✓✓✓	✓✓	✓
Sampling-related Issues	Does the mode of administration tend to produce a good **response** rate?	✓✓✓	✓✓	✓	✓	✓
	Can the researcher **control** who responds (so that the person targeted is the person who answers)?	✓✓✓	✓✓✓	✓✓	✓✓	✓✓
	Is the mode of administration **accessible** to all sample members?	✓✓✓	✓✓	✓✓✓	✓ (because respondents need online access)	✓ (because respondents need online access)
Questionnaire Issues	Is it suitable for long questionnaires?	✓✓✓	✓✓	✓✓	✓✓	✓✓
	Is it suitable for complex questionnaires?	✓✓✓	✓	✓✓	✓✓	✓✓
	Is it suitable for open questions?	✓✓✓	✓✓	✓	✓✓	✓✓
	Is it suitable for filter questions?	✓✓✓ (especially if CATI used)	✓✓✓ (especially if CATI used)	✓	✓	✓✓✓ (if it allows jumping)
	Does it allow control over the order in which questions are answered?	✓✓✓	✓✓✓	✓	✓	✓✓
	Is it suitable for sensitive questions?	✓	✓✓	✓✓✓	✓✓✓	✓✓✓
	Is it less likely to result in non-response to questions?	✓✓✓	✓✓✓	✓✓	✓✓	✓✓
	Does it allow the use of visual aids?	✓✓✓	✓	✓✓✓	✓✓	✓✓✓
Answering Context Issues	Does it give respondents the opportunity to consult others for information?	✓✓	✓	✓✓✓	✓✓✓	✓✓✓
	Does it minimise the impact of interviewers' characteristics (gender, class, ethnicity) where these are not relevant to the researcher aims?	✓	✓✓	✓✓✓	✓✓✓	✓✓✓
	Does it minimise the impact of the social desirability effect?	✓	✓✓	✓✓✓	✓✓✓	✓✓✓
	Does it allow control over the intrusion of others in answering questions?	✓✓✓	✓✓	✓	✓	✓
	Does it minimise the need for respondents to have certain skills to answer questions?	✓✓✓	✓✓✓	✓✓	✓ (because of the need to have online skills)	✓ (because of the need to have online skills)
	Does it enable respondents to be probed?	✓✓✓	✓✓✓	✓	✓✓	✓
	Does it reduce the likelihood of data entry errors by the researcher?	✓	✓	✓✓	✓	✓✓✓

Appendix 6. UWTSD Ethical Approval Form

EthicalApprovalForm
final.docx

APPLICATION FOR ETHICAL APPROVAL

RESEARCH STUDENTS
This form is to be completed by the student within **SIX** months for full-time students and **TWELVE** months for part time students, after the commencement of the research degree or following progression to Part Two of your course.

Once complete, submit this form via the ***MyTSD Doctoral College Portal*** at (https://mytsd.uwtsd.ac.uk).

This document is also available in Welsh.

RESEARCH STAFF ONLY
All communications relating to this application during its processing must be in writing and emailed to pgresearch@uwtsd.ac.uk , with the title ‘Ethical Approval’ followed by your name.

STUDENTS ON UNDERGRADUATE OR TAUGHT MASTERS PROGRAMMES should submit this form (and receive the outcome) via systems explained to you by the supervisor/module leader.

In order for research to result in benefit and minimise risk of harm, it must be conducted ethically. A researcher may not be covered by the University’s insurance if ethical approval has not been obtained prior to commencement.

The University follows the OECD Frascati manual definition of **research activity**: “creative work undertaken on a systematic basis in order to increase the stock of knowledge, including knowledge of man, culture and society, and the use of this stock of knowledge to devise new applications”. As such this covers activities undertaken by members of staff, postgraduate research students, and both taught postgraduate and undergraduate students working on dissertations/projects.

The individual undertaking the research activity is known as the “principal researcher”.

Ethical approval is not required for routine audits, performance reviews, quality assurance studies, testing within normal educational requirements, and literary or artistic criticism.

Please read the notes for guidance before completing ALL sections of the form.

This form must be completed and approved prior to undertaking any research activity. Please see Checklist for details of process for different categories of application.

SECTION A: About You (Principal Researcher)

1	Full Name:	Jay Hier-Jones

2	Tick all boxes that apply:		Member of staff:	☐	Honorary research fellow:	☐
	Undergraduate Student	☐	Taught Postgraduate Student	☒	Postgraduate Research Student	☒

3	Institute/Academic Discipline/Centre:	Institute of Management & Health
4	Campus:	Online/Virtual
5	E-mail address:	2111607@student.uwtsd.ac.uk
6	Contact Telephone Number:	
	For students:	
7	Student Number:	2111607
8	Programme of Study:	MSc Digital Transformation for the Health and Care Professions
9	Director of Studies/Supervisor:	Prof Phillip Scott / Dr Ben Duxbury

SECTION B: Approval for Research Activity

1	Has the research activity received approval in principle? (please check the Guidance Notes as to the appropriate approval process for different levels of research by different categories of individual)	**YES**	☐	**NO**	☒
				Date	
2	If Yes, please indicate source of approval (and date where known): ***Approval in principle must be obtained from the relevant source prior to seeking ethical approval***	Research Degrees Committee	☐		
		Institute Research Committee	☐		
		Other (write in)	☐		

SECTION C: Internal and External Ethical Guidance Materials

	Please list the core ethical guidance documents that have been referred to during the completion of this form (including any discipline-specific codes of research ethics, location-specific codes of research ethics, and also any specific ethical guidance relating to the proposed methodology). Please tick to confirm that your research proposal adheres to these codes and guidelines. You may add rows to this table if needed.	
1	**UWTSD Research Ethics & Integrity Code of Practice**	☒
2	**UWTSD Research Data Management Policy**	☒
3	*[List any other relevant documents here]*	☐

SECTION D: External Collaborative Research Activity

If there are external collaborators then you should gain consent from the contact persons to share their personal data with the university. If there are no external collaborators then leave this section blank and continue to section E.

1	Institution					
2	Contact person name					
3	Contact person e-mail address					
4	Is your research externally funded?		**YES**	☐	**NO**	☐
5	Are you in receipt of a KESS scholarship?		**YES**	☐	**NO**	☐
6	Are you specifically employed to undertake this research in either a paid or voluntary capacity?	Voluntary	**YES**	☐	**NO**	☐
7		Employed	**YES**	☐	**NO**	☐
8	Is the research being undertaken within an existing UWTSD Athrofa Professional Learning Partnership (APLP)?	If YES then the permission question below does not need to be answered.	**YES**	☐	**NO**	☐
9	Has permission to undertake the research has been provided by the partner organisation?	(If YES attach copy) If NO the application cannot continue	**YES**	☐	**NO**	☐

Where research activity is carried out in collaboration with an external organisation

10	Does this organisation have its own ethics approval system?	**YES**	☐	**NO**	☐
	If Yes, please attach a copy of any final approval (or interim approval) from the organisation (this may be a copy of an email if appropriate).				

SECTION E: Details of Research Activity

1	Indicative title:	Optimising Healthcare Data Flows and Interoperability: Exploring Challenges and Opportunities in NHS Wales		
2	Proposed start date:	Sept 2024	Proposed end date:	May 2025

Introduction to the Research (maximum 300 words per section)

Ensure that you write for a Non-Specialist Audience when outlining your response to the points below:

Purpose of Research Activity
Proposed Research Question
Aims of Research Activity
Objectives of Research Activity

Demonstrate, briefly, how **Existing Research** has informed the proposed activity and explain

What the research activity will add to the body of knowledge
How it addresses an area of importance.

3

Purpose of Research Activity

The purpose of this study is to explore the data flows and interoperability within NHS Wales, with a particular focus on Powys Teaching Health Board (PTHB). The study aims to evaluate how effectively data is shared and utilised across systems, identifying factors that may affect operational efficiency and patient care. This study seeks to assess existing practices, barriers, and opportunities related to data interoperability across NHS Wales, exploring technical, human and organisational factors. In line with Robson's (2024) view that evaluation is both an assessment and a tool for assisting in improvements, the study hopes identification of potential strategies could be used to further optimise data flows, enhancing decision-making processes and patient outcomes.

Insights from existing research highlight that effective data sharing and interoperability are critical for optimising decision-making processes and improving patient outcomes (Choun and Petre, 2022; Perlman and Elsner, 2019, 2024; Welsh Government, 2023). Welsh Government's Digital and Data Strategy for Health and Social Care (2023) emphasises the need for standardised data infrastructures and cross-sector collaboration to enable seamless sharing of patient information, however there are key barriers to achieving this such as fragmented data flows and a lack of system interoperability (Mistry *et al.* 2022; FEDIP, 2024).

The research activity will contribute to the body of knowledge by providing an exploration of data flows and interoperability within NHS Wales and aspires to provide valuable evidence to inform the development of strategies and solutions that align with national policy frameworks, particularly in the context of rural healthcare settings like PTHB.

Choun, D., and Petre, A. (2022). Digital Health and Patient Data: Empowering Patients in the Healthcare Ecosystem (1st ed.). Productivity Press. https://doi.org/10.4324/9781003215868

FEDIP (2024) *DRAFT FEDIP Response to the Tony Blair Institute Report: "Preparing the NHS for the AI Era: A Digital Health Record for Every Citizen."* Available at: https://c2f91b66-85f1-477d-9fd6-66053db82524.usrfiles.com/ugd/c2f91b_05da1934337542508ba6b6feb2c495e1.pdf?utm_source=ActiveCampaign&utm_medium=email&utm_content=Seeking%20Your%20Insights%20on%20Transforming%20NHS%20with%20AI%20and%20Digital%20Health%20Records&utm_campaign=FDLs%20Blair%20Report&utm_source=ActiveCampaign&utm_medium=email&utm_content=Lend%20Your%20Voice%3A%20Feedback%20Needed%20on%20Digital%20Health%20Transformation%20Report&utm_campaign=Practitioner%20Re%20BLAIR
[Accessed 4 November 2024]

	Mistry, P., Maguire, D., Chikwira, L., and Lindsay, T. (2022). *Interoperability is more than technology: The role of culture and leadership in joined-up care*. The King's Fund. Available at: https://assets.kingsfund.org.uk/f/256914/x/c48bd5a1a2/interoperability_more_than_technology_2022.pdf [Accessed on 4 November 2022] Perlman, M. and Elsner, N. (2019). *Digital health technology: Global case studies of health care transformation.* Deloitte Insights Available at: https://www2.deloitte.com/us/en/insights/industry/health-care/digital-health-technology.html [Accessed: 7 November 2024] Robson, C. (2024). *Real World Research: a Resource for Users of Social Research Methods in Applied Settings*. (5th edn.). Chichester: Wiley Welsh Government (2023) *Digital and data strategy for health and social care in Wales.* Policy and Strategy document. Available at: https://www.gov.wales/digital-and-data-strategy-health-and-social-care-wales-html [Accessed: 26 October 2024]. (this box should expand as you type)
4	**Research Question** What factors impact the effectiveness of data flow and interoperability in supporting healthcare delivery within the NHS? (this box should expand as you type)
5	**Aims of Research Activity** This research aims to answer the research question by evaluating the current state of data flows and interoperability within Powys Teaching Health Board (PTHB). The study seeks to understand how these processes operate in practice, identify factors that influence their effectiveness, and provide insights into their role in supporting healthcare delivery and improving operational outcomes. (this box should expand as you type)
6	**Objectives of Research Activity** To achieve the research aim, the following objectives will look to identify; • ***What are the current data flow processes at Powys Teaching Health Board (PTHB), and how do they function in practice?*** *Why: To gain insight into existing data mechanisms and their operational role within healthcare delivery.* • ***How do healthcare professionals perceive the interoperability of the systems they use, and what impact does this have on their work?*** *Why: To explore user experiences and gather insights into how system interoperability affects the efficiency of healthcare delivery.* • ***In what ways do data quality and accessibility influence decision-making and***

	patient outcomes? *Why: To assess the importance of data reliability and availability in supporting informed clinical decisions and patient care.* • ***What factors are perceived as barriers or enablers to improving data flows and interoperability?*** *Why: To identify key influences on data management practices, which could help understand areas for potential enhancement in the future.* (this box should expand as you type)
	Proposed methods (maximum 600 words) Provide a brief summary of all the methods that **may** be used in the research activity, making it clear what specific techniques may be used. If methods other than those listed in this section are deemed appropriate later, additional ethical approval for those methods will be needed. You do not need to justify the methods here, but should instead describe how you intend to collect the data necessary for you to complete your project.
7	The principal data collection method will involve conducting a survey, complemented by existing data extracted from the Powys Teaching Health Board (PTHB) data warehouse. The proposed methods will take a mixed methods approach, using both qualitative and quantitative methods to obtain data which will support to answer the research aims and objectives. An anonymous MS Forms Questionnaire will collect qualitative and quantitative data from PTHB and other NHS staff and stakeholders. The targeted main sample population will be PTHB staff who have made requests and recently worked with the PTHB Data Engineering and Analysis team because these individuals will be best placed to explain and evaluate the use data within their role. The questionnaire will be circulated with an information consent form explaining the use and security of response data and participants right to withdraw. Participants of the questionnaire will be assigned unique reference numbers to ensure they are not identifiable. Likert questions within the questionnaire will be used for quantitative metrics, while qualitative questions will be analysed to identify themes between responses and highlight commonalities between individuals' viewpoints. The questionnaire aims to capture whether staff experience or believe there are significant benefits to interoperable and real-time data flows. It looks to understand how well-versed staff are with using data, whilst identifying if there are any perceived barriers or blockers with regards to data accessibility, reliability or usability within their role. An additional documentary sources will be used to obtain quantitative data from data which already exists within the PTHB Data Engineering & Analysis Team' Databricks Platform. Structured Query Language (SQL) scripts will be used to collect aggregated non-personable identifiable data. This part of the data collection and analysis will focus on three main areas: Firstly, summarised data will be collected to identify instances where demographic errors are present for the same individuals across multiple systems. Secondly, the study will identify the number of duplicated records across different systems, focusing on instances where activities were logged for the same individual by the same healthcare professional on the same date but recorded in different systems. Thirdly, data will be extracted to compare the number of manual data processing tasks undertaken by the PTHB Data Engineering & Analysis Team with the number of automated

	processes the team has implemented. Manual tasks will be identified from the work items recorded within the Azure DevOps system, while workflows running within the Databricks platform will reveal the extent of automated data processes. The documentary data analysis of will focus on four key areas: • Demographic Consistency: The summarised demographic data will be examined to identify any discrepancies for the same individuals across multiple systems, highlighting potential data quality issues. • Data Duplication: Analysis will be made to identify any instances of duplicated records across systems, to help assess the extent of data duplication and its impact on data integrity. • Manual Interventions: The logs from Azure DevOps will be analysed to evaluate the frequency and number of manual data processing tasks performed by the PTHB Data Engineering & Analysis Team. This will provide insights into the time and resources required for manual interventions due to limitations in data flow and interoperability. • Automated Processes: The study will assess the current automated processes in place, evaluating their effectiveness in ensuring timely and up-to-date data flow. (this box should expand as you type)
	Location of research activity Identify all locations where research activity will take place.
8	Electronically MS Forms for NHS Staff and within Powys Teaching Health Boards (PTHB) Cloud databases. (this box should expand as you type)
	Research activity outside of the UK If research activity will take place overseas, you are responsible for ensuring that local ethical considerations are complied with and that the relevant permissions are sought. Specify any local guidelines (e.g. from local professional associations/learned societies/universities) that exist and whether these involve any ethical stipulations beyond those usual in the UK (provide details of any licenses or permissions required). Also specify whether there are any specific ethical issues raised by the local context in which the research activity is taking place, for example, particular cultural and/or legal sensitivities or vulnerabilities of participants. If you live in the country where you will do the research then please state this.
9	N/A *(this box should expand as you type)*

10	Use of documentation not in the public domain: Are any documents **NOT** publicly available?	**NO**	☐
		YES	☒
11	If Yes, please provide details here of how you will gain access to specific documentation that is not in the public domain and that this is in accordance with the current data protection law of the country in question and that of England and Wales. Data held within Powys Teaching Health Boards (PTHB) will be accessed based on approval from the health boards research and development hub. All data will be aggregated and not personally identifiable data. This study is unlikely to meet any thresholds which require special approval. *(this box should expand as you type)*		

	Does your research relate to one or more of the seven aims of the Well-being of Future Generations (Wales) Act 2015?	**YES**	**NO**
12	A prosperous Wales	☐	☒
13	A resilient Wales	☐	☒
14	A healthier Wales	☒	☐
15	A more equal Wales	☒	☐
16	A Wales of cohesive communities	☐	☒
17	A Wales of vibrant culture and thriving Welsh language	☐	☒
18	A globally responsible Wales	☐	☒
19	If YES to any of the above, please give details:		
	Whilst not explicitly, the Act's goals and principles inherently support the need for effective data management to achieve its well-being objectives. The healthier Wales aim of the act aims to improve physical and mental well-being. For healthcare professionals, having the right data at the right time can lead to better patient outcomes, more efficient care delivery, and improved public health strategies and this is what this project hopes to facilitate. Similarly, this could be applied to some of the other principles for example, ensuring that all professionals have access to the same high-quality data can reduce inequalities in service delivery and ensure that all communities benefit from the same level of care and resources. its goals and principles inherently support the need for effective data management to achieve its well-being objectives. (this box should expand as you type)		

SECTION F: Scope of Research Activity

	Will the research activity include:	**YES**	**NO**
1	Use of a questionnaire or similar research instrument?	☒	☐
2	Use of interviews?	☐	☒
3	Use of focus groups?	☐	☒
4	Use of participant diaries?	☐	☒
5	Use of video or audio recording?	☐	☒
6	Use of computer-generated log files?	☐	☒
7	Participant observation with their knowledge?	☐	☒
8	Participant observation without their knowledge?	☐	☒
9	Access to personal or confidential information without the participants' specific consent?	☐	☒
10	Administration of any questions, test stimuli, presentation that may be experienced as physically, mentally or emotionally harmful / offensive?	☐	☒

11	Performance of any acts which may cause embarrassment or affect self-esteem?	☐	☒
12	Investigation of participants involved in illegal activities?	☐	☒
13	Use of procedures that involve deception?	☐	☒
14	Administration of any substance, agent or placebo?	☐	☒
15	Working with live vertebrate animals?	☐	☒
16	Procedures that may have a negative impact on the environment?	☐	☒
17	Other primary data collection methods. Please indicate the type of data collection method(s) below.	☒	☐
	Details of any other primary data collection method: Access to data recorded internally to Powys Teaching Health Board Data Engineering & Analytics Team (this box should expand as you type)		

If NO to every question, then the research activity is (ethically) low risk and **may** be exempt from **some** of the following sections (please refer to Guidance Notes).

If YES to any question, then no research activity should be undertaken until full ethical approval has been obtained.

SECTION G: Intended Participants

If there are no participants then do not complete this section, but go directly to section H.

	Who are the intended participants:	**YES**	**NO**
1	Students or staff at the University?	☐	☒
2	Adults (over the age of 18 and competent to give consent)?	☒	☐
3	Vulnerable adults?	☐	☒
4	Children and Young People under the age of 18? (Consent from Parent, Carer or Guardian will be required)	☐	☒
5	Prisoners?	☐	☒
6	Young offenders?	☐	☒
7	Those who could be considered to have a particularly dependent relationship with the investigator or a gatekeeper?	☐	☒
8	People engaged in illegal activities?	☐	☒
9	Others. Please indicate the participants below, and specifically any group who may be unable to give consent.	☒	☐
	Details of any other participant groups: Staff and colleagues within Powys Teaching Health Board and potentially other Health Boards within Wales/NHS England		

	Stakeholders for systems which are in use withing Powys Teaching Health Board (this box should expand as you type)		

	Participant numbers and source Provide an estimate of the expected number of participants. How will you identify participants and how will they be recruited?	
10	How many participants are expected?	Aiming for 50 Questionnaire responses *(this box should expand as you type)*
11	Who will the participants be?	Staff and colleagues within Powys Teaching Health Board and potentially other Health Boards within Wales/NHS England Stakeholders for systems which are in use withing Powys Teaching Health Board *(this box should expand as you type)*
12	How will you identify the participants?	Targeted individuals who have submitted requests to the PTHB Data Engineering and Analytics Team, individuals within data teams in other health boards as well as by asking colleagues to share to their colleagues. *(this box should expand as you type)*

	Information for participants:	**YES**	**NO**	**N/A**
13	Will you describe the main research procedures to participants in advance, so that they are informed about what to expect?	☒	☐	☐
14	Will you tell participants that their participation is voluntary?	☒	☐	☐
15	Will you obtain written consent for participation?	☒	☐	☐
16	Will you explain to participants that refusal to participate in the research will not affect their treatment or education (if relevant)?	☒	☐	☐
17	If the research is observational, will you ask participants for their consent to being observed?	☐	☐	☒
18	Will you tell participants that they may withdraw from the research at any time and for any reason?	☒	☐	☐
19	With questionnaires, will you give participants the option of omitting questions they do not want to answer?	☒	☐	☐
20	Will you tell participants that their data will be treated with full confidentiality and that, if published, it will not be identifiable as theirs?	☒	☐	☐
21	Will you debrief participants at the end of their participation, in a way appropriate to the type of research undertaken?	☒	☐	☐
22	If NO to any of above questions, please give an explanation			
	(this box should expand as you type)			

	Information for participants:	**YES**	**NO**	**N/A**
24	Will participants be paid?	☐	☒	☐
25	Is specialist electrical or other equipment to be used with participants?	☐	☒	☐
26	Are there any financial or other interests to the investigator or University arising from this study?	☐	☒	☐
27	Will the research activity involve deliberately misleading participants in any way, or the partial or full concealment of the specific study aims?	☐	☒	☐
28	If YES to any question, please provide full details			
	(this box should expand as you type)			

SECTION H: Anticipated Risks

	Outline any anticipated risks that may adversely affect any of the participants, the researchers and/or the University, and the steps that will be taken to address them. If you have completed a full risk assessment (for example as required by a laboratory, or external research collaborator) you may append that to this form.		
1	**Full risk assessment completed and appended?**	Yes No	☐ ☒
2	**Risks to participants** For example: sector-specific health & safety, emotional distress, financial disclosure, physical harm, transfer of personal data, sensitive organisational information		
	Risk to participants: Participants may have concerns over the confidentiality of their responses and be concerned that their responses could identify them, especially if negative. This concern could pressurise a biased and more positive response opposed to honest opinions. Stress due to having already busy work commitments and perception that participating is an additional time-burden. Also potential fear of change due to the evolving digital world. *(this box should expand as you type)*	*How you will mitigate the risk to participants:* Concerns around participant confidentiality will be reduced due to following ethical guidelines, obtaining consent and issuing an information and statement of participation explaining the clear purpose of the questionnaire, that data will be anonymised and used solely for research purposes. The questionnaire will be relatively short and the researcher will offer additional time to discuss any concerns. Staff are also supported by the health boards well-being service. Efforts will be made to ensure wording of the questions are easy to understand to save any confusion. Participants will be reminded their participation is voluntary and stated in the information and statement of participation.	

		(this box should expand as you type)
3	If research activity may include sensitive, embarrassing or upsetting topics (e.g. sexual activity, drug use) or issues likely to disclose information requiring further action (e.g. criminal activity), give details of the procedures to deal with these issues, including any support/advice (e.g. helpline numbers) to be offered to participants. Note that where applicable, consent procedures should make it clear that if something potentially or actually illegal is discovered in the course of a project, it may need to be disclosed to the proper authorities	
	N/A *(this box should expand as you type)*	
4	**Risks to the investigator** For example: personal health & safety, physical harm, emotional distress, risk of accusation of harm/impropriety, conflict of interest	
	Risk to the investigator: Stress due to having already busy work commitments and perception that participating is an additional time-burden *(this box should expand as you type)*	*How you will mitigate the risk to the investigator:* Staff are also supported by the health boards well-being service, the investigator will agree protected time with their manager to undertake this investigation. *(this box should expand as you type)*
5	**University/institutional risks** For example: adverse publicity, financial loss, data protection	
	Risk to the University: Data Breaches Reputational Damage *(this box should expand as you type)*	*How you will mitigate the risk to the University:* By abiding to ethical research guidelines. Securing data within password protected devices and the UWTSD network. Limiting access to the raw data before anonymisation to the investigator and academic supervisor. *(this box should expand as you type)*
6	**Environmental risks** For example: accidental spillage of pollutants, damage to local ecosystems	

	Risk to the environment: No risks to the environment *(this box should expand as you type)*	*How you will mitigate the risk to environment:* N/A *(this box should expand as you type)*

	Disclosure and Barring Service			
	If the research activity involves children or vulnerable adults, a Disclosure and Barring Service (DBS) certificate must be obtained before any contact with such participants.	**YES**	**NO**	**N/A**
7	Does your research require you to hold a current DBS Certificate?	☐	☐	☒
8	If YES, please give the certificate number. If the certificate number is not available please write "Pending"; in this case any ethical approval will be subject to providing the appropriate certificate number.			

SECTION I: Feedback, Consent and Confidentiality

1	**Feedback** What de-briefing and feedback will be provided to participants, how will this be done and when?
	Participants will be thanked following responding to the questionnaire. *(this box should expand as you type)*
2	**Informed consent** Describe the arrangements to inform potential participants, before providing consent, of what is involved in participating. Describe the arrangements for participants to provide full consent before data collection begins. If gaining consent in this way is inappropriate, explain how consent will be obtained and recorded in accordance with prevailing data protection legislation.
	Participants will be provided with a written statement for them to agree to at the start of the MS Form, ensuring they understand the purpose of the study, how their data will be used, and the overall importance of their contribution. This will detail their right to withdraw at any point, how long and where data will be stored. *(this box should expand as you type)*
3	**Confidentiality / Anonymity** Set out how anonymity of participants and confidentiality will be ensured in any outputs. If anonymity is not being offered, explain why this is the case.
	Data collected for the study is de-identified and handled securely to protect privacy and comply with data protection regulations. Assigning each participant a number will be the approach taken for this project, and any selected quotes used will be assured that the respondent cannot be identified. *(this box should expand as you type)*

SECTION J: Data Protection and Storage

	Does the research activity involve personal data (as defined by the General Data Protection Regulation 2016 “GDPR” and the Data Protection Act 2018 “DPA”)?	**YES**	**NO**
1	***“Personal data”*** *means any information relating to an identified or identifiable natural person (‘data subject’). An identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person. Any video or audio recordings of participants is considered to be personal data.*	☐	☒
	If YES, provide a description of the data and explain why this data needs to be collected:		
2	*(this box should expand as you type)*		
	Does it involve special category data (as defined by the GDPR)?	**YES**	**NO**
3	***“Special category data”*** *means sensitive personal data consisting of information as to the data subjects’ –* *(a) racial or ethnic origin,* *(b) political opinions,* *(c) religious beliefs or other beliefs of a similar nature,* *(d) membership of a trade union (within the meaning of the Trade Union and Labour Relations (Consolidation) Act 1992),* *(e) physical or mental health or condition,* *(f) sexual life,* *(g) genetics,* *(h) biometric data (as used for ID purposes),*	☐	☒
	If YES, provide a description of the special category data and explain why this data needs to be collected:		
4	*(this box should expand as you type)*		

	Will data from the research activity (collected data, drafts of the thesis, or materials for publication) be stored in any of the following ways?	YES	NO
5	Manual files (i.e. in paper form)?	☐	☐
6	University computers?	☐	☒
7	Private company computers?	☒	☐
8	Home or other personal computers?	☐	☒
9	Laptop computers/ CDs/ Portable disk-drives/ memory sticks?	☐	☒
10	“Cloud” storage or websites?	☒	☐
11	Other – specify:	☐	☐
12	For all stored data, explain the measures in place to ensure the security of the data collected, data confidentiality, including details of backup procedures, password protection, encryption, anonymisation and pseudonymisation:		

	Data within MS Forms is encrypted and in compliance with global regulatory standards. Microsoft authentication ensures that only authorized users can view form responses and can be deleted when no longer required. Any data analysis will be anonymised by giving participant numbers and stored within secure NHS cloud-based storage for the length of time needed to complete this project. *(this box should expand as you type)*

	Data Protection		
	Will the research activity involve any of the following activities:	**YES**	**NO**
13	Electronic transfer of data in any form?	☐	☒
14	Sharing of data with others at the University outside of the immediate research team?	☐	☒
15	Sharing of data with other organisations?	☐	☒
16	Export of data outside the UK or importing of data from outside the UK?	☐	☒
17	Use of personal addresses, postcodes, faxes, emails or telephone numbers?	☐	☒
18	Publication of data that might allow identification of individuals?	☐	☒
19	Use of data management system?	☐	☒
20	Data archiving?	☐	☒
21	If YES to any question, please provide full details, explaining how this will be conducted in accordance with the GDPR and Data Protection Act (2018) (and any international equivalents, where appropriate):		
	(this box should expand as you type)		
22	List all who will have access to the data generated by the research activity:		
	Jay Hier-Jones and potentially her project supervisor. *(this box should expand as you type)*		
23	List who will have control of, and act as custodian(s) for, data generated by the research activity:		
	Jay Hier-Jones *(this box should expand as you type)*		
24	Give details of data storage arrangements, including security measures in place to protect the data, where data will be stored, how long for, and in what form. Will data be archived – if so how and if not why not.		
	Data storage within secure NHS cloud-based storage; a personal OneDrive for the length of time needed to complete this programme of study. *(this box should expand as you type)*		

25	Please indicate if your data will be stored in the UWTSD Research Data Repository (see https://researchdata.uwtsd.ac.uk/). If so please explain. (*Most relevant to academic staff)*		
	(this box should expand as you type)		
26	Confirm that you have read the UWTSD guidance on data management (see https://www.uwtsd.ac.uk/library/research-data-management/)	**YES**	☒
27	Confirm that you are aware that you need to keep all data until after your research has completed or the end of your funding	**YES**	☒

SECTION K: Declaration

	The information which I have provided is correct and complete to the best of my knowledge. I have attempted to identify any risks and issues related to the research activity and acknowledge my obligations and the rights of the participants. In submitting this application I hereby confirm that I undertake to ensure that the above named research activity will meet the University's Research Ethics and Integrity Code of Practice which is published on the website: https://www.uwtsd.ac.uk/research/research-ethics/		
1	**Signature of applicant:**		**Date:** **11/11/2024**

For STUDENT Submissions:

2	Director of Studies/Supervisor:	B. Duxbury	**Date:** 5/12/24
3	Signature:		

For STAFF Submissions:

4	Academic Director/ Assistant Dean:		**Date:**
5	Signature:		

Checklist: Please complete the checklist below to ensure that you have completed the form according to the guidelines and attached any required documentation:

☒	I have read the guidance notes supplied before completing the form.
☒	I have completed **ALL RELEVANT** sections of the form in full.
☐	I confirm that the research activity has received approval in principle
☐	I have attached a copy of final/interim approval from external organisation (**where appropriate**)
☐	I have attached a full risk assessment (where appropriate) *ONLY TICK IF YOU HAVE* ***ATTACHED A FULL RISK ASSESSMENT***

☒	I understand that it is my responsibility to ensure that the above named research activity will meet the University's Research Ethics and Integrity Code of Practice.
☒	I understand that before commencing data collection all documents aimed at respondents (including information sheets, consent forms, questionnaires, interview schedules etc.) must be confirmed by the DoS/Supervisor, module tutor or Academic Director.

RESEARCH STUDENTS ONLY
Once complete, submit this form via the **MyTSD Doctoral College Portal** at (https://mytsd.uwtsd.ac.uk).

RESEARCH STAFF ONLY
All communications relating to this application during its processing must be in writing and emailed to pgresearch@uwtsd.ac.uk , with the title 'Ethical Approval' followed by your name.

STUDENTS ON UNDERGRADUATE OR TAUGHT MASTERS PROGRAMMES should submit this form (and receive the outcome) via systems explained to you by the supervisor/module leader.

Appendix 8. Questionnaire; Participant Consent Information

Participation Information Sheet - *MSc Digital Transformation for the Health and Care Professions* - Evaluating Healthcare Data Flows and Interoperability: Exploring Challenges and Opportunities in NHS Wales

The Questionnaire

The purpose of this questionnaire is to explore your perception of data flow practices, data quality and accessibility and within your organisation. It seeks to identify positive ways data can and does inform your role well as any issues or concerns you may have, aiming to understand the impact any data challenges may have on operational efficiency and patient care. This includes exploring the role of interoperable systems, data standards, and specific technologies used for enhancing data flows and their efficiency.

The term interoperability is used to describe how systems can connect and share information between them effectively.

The Study

This questionnaire will inform the broader aim of the MSc thesis which seeks to evaluate data flows and interoperability within NHS Wales. The study seeks to:

- Assess how data is currently shared and utilised across systems.
- Identify factors that influence operational efficiency and patient outcomes.
- Explore technical, human, and organisational barriers and opportunities related to data flows and data interoperability

Insights gathered will contribute to identifying strategies and factors relating to optimised data flows and their ability to enhance decision-making processes and improve healthcare delivery and patient outcomes.

What is involved with taking part?

Completion of a short opinion-based survey regarding data challenges, concerns and improvements which should take no longer than 10-to-15 minutes. A link to the online questionnaire can be found below.

Your data

Your responses will be collected automatically into a central digital repository, this will only be seen by the author of the study and possibly by her academic supervisor. Once received, any identifiable data will be anonymised or removed and handled securely to protect privacy and comply with data protection regulations. Responses will be analysed for common themes; direct quotes may be extracted but will be used anonymously. Data will be stored in an anonymised format on a secure NHS cloud drive.

Consent

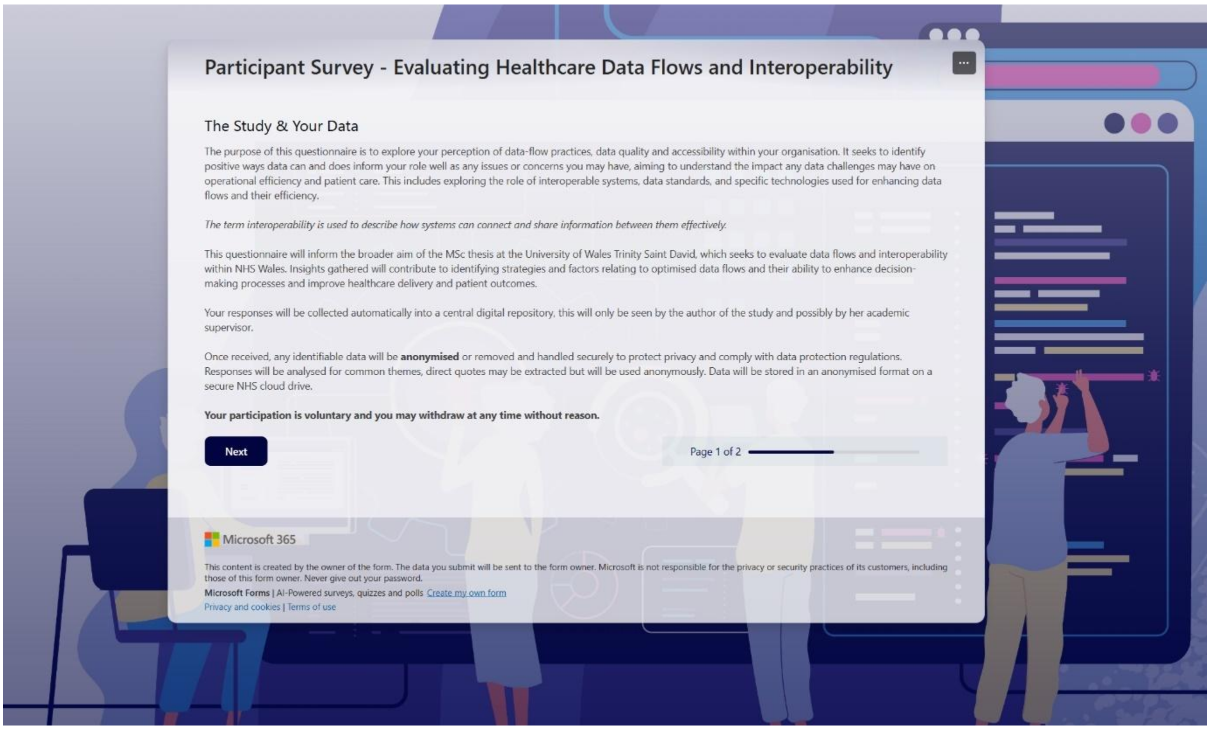
Your completion and return of the survey via Microsoft Forms will imply consent for your data to be used for the purposes of this study. Your participation is voluntary, and you may withdraw at any time without reason.

Links

Either scan the QR code or click on the link below to be taken to the survey.

https://forms.office.com/e/RBxuN9vin7

Example of Online Design



Participant Survey - Evaluating Healthcare Data Flows and Interoperability

The Study & Your Data

The purpose of this questionnaire is to explore your perception of data-flow practices, data quality and accessibility within your organisation. It seeks to identify positive ways data can and does inform your role well as any issues or concerns you may have, aiming to understand the impact any data challenges may have on operational efficiency and patient care. This includes exploring the role of interoperable systems, data standards, and specific technologies used for enhancing data flows and their efficiency.

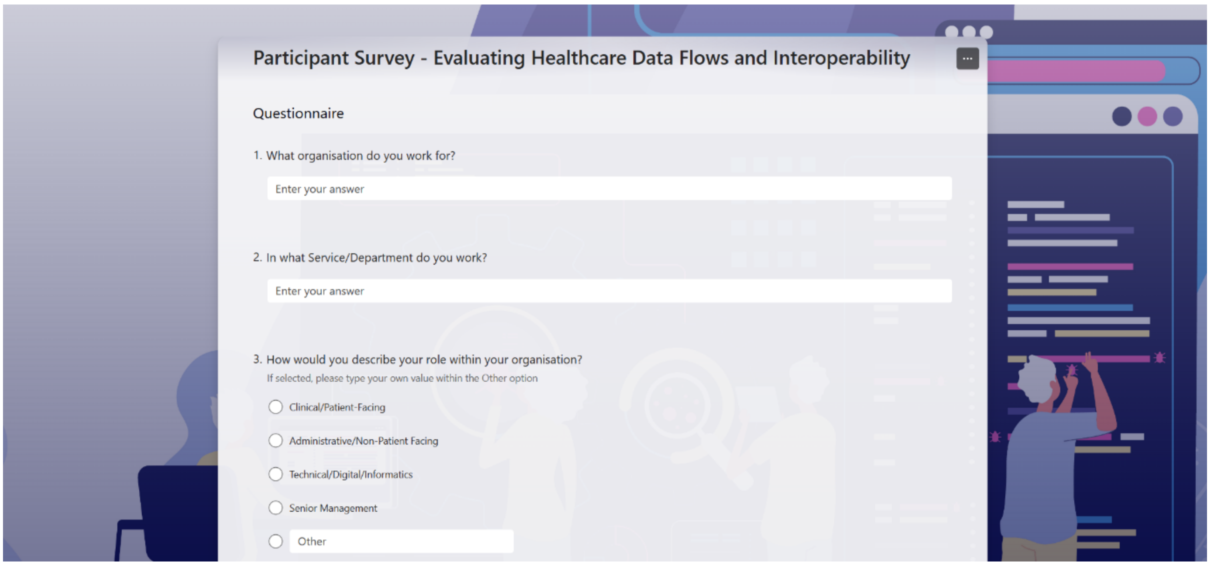
The term interoperability is used to describe how systems can connect and share information between them effectively.

This questionnaire will inform the broader aim of the MSc thesis at the University of Wales Trinity Saint David, which seeks to evaluate data flows and interoperability within NHS Wales. Insights gathered will contribute to identifying strategies and factors relating to optimised data flows and their ability to enhance decision-making processes and improve healthcare delivery and patient outcomes.

Your responses will be collected automatically into a central digital repository, this will only be seen by the author of the study and possibly by her academic supervisor.

Once received, any identifiable data will be **anonymised** or removed and handled securely to protect privacy and comply with data protection regulations. Responses will be analysed for common themes, direct quotes may be extracted but will be used anonymously. Data will be stored in an anonymised format on a secure NHS cloud drive.

Your participation is voluntary and you may withdraw at any time without reason.



Participant Survey - Evaluating Healthcare Data Flows and Interoperability

Questionnaire

1. What organisation do you work for?

Enter your answer

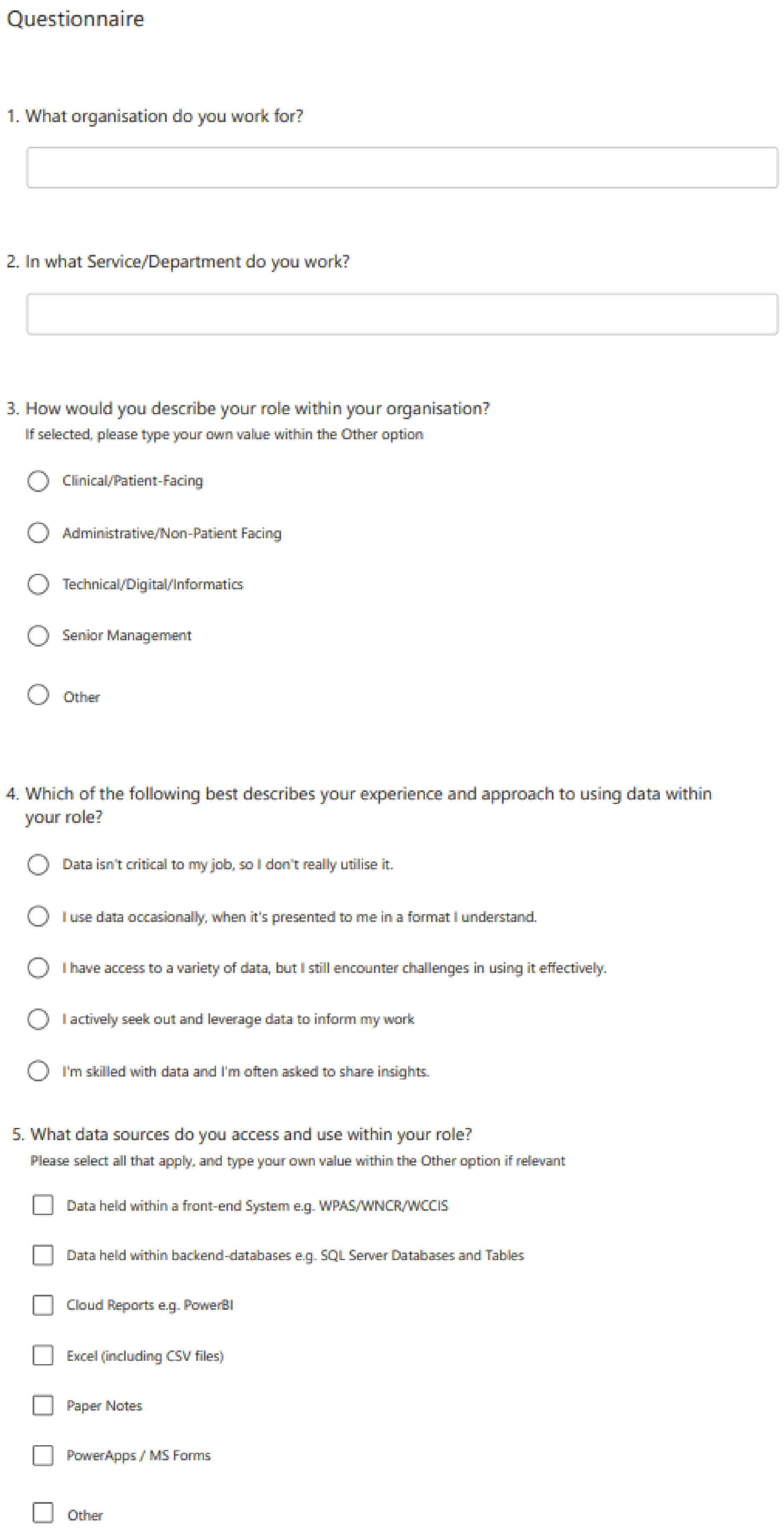
2. In what Service/Department do you work?

Enter your answer

3. How would you describe your role within your organisation?

If selected, please type your own value within the Other option

○ Clinical/Patient-Facing
○ Administrative/Non-Patient Facing
○ Technical/Digital/Informatics
○ Senior Management
○ Other



Questionnaire

1. What organisation do you work for?

2. In what Service/Department do you work?

3. How would you describe your role within your organisation?
If selected, please type your own value within the Other option

○ Clinical/Patient-Facing

○ Administrative/Non-Patient Facing

○ Technical/Digital/Informatics

○ Senior Management

○ Other

4. Which of the following best describes your experience and approach to using data within your role?

○ Data isn't critical to my job, so I don't really utilise it.

○ I use data occasionally, when it's presented to me in a format I understand.

○ I have access to a variety of data, but I still encounter challenges in using it effectively.

○ I actively seek out and leverage data to inform my work

○ I'm skilled with data and I'm often asked to share insights.

5. What data sources do you access and use within your role?
Please select all that apply, and type your own value within the Other option if relevant

☐ Data held within a front-end System e.g. WPAS/WNCR/WCCIS

☐ Data held within backend-databases e.g. SQL Server Databases and Tables

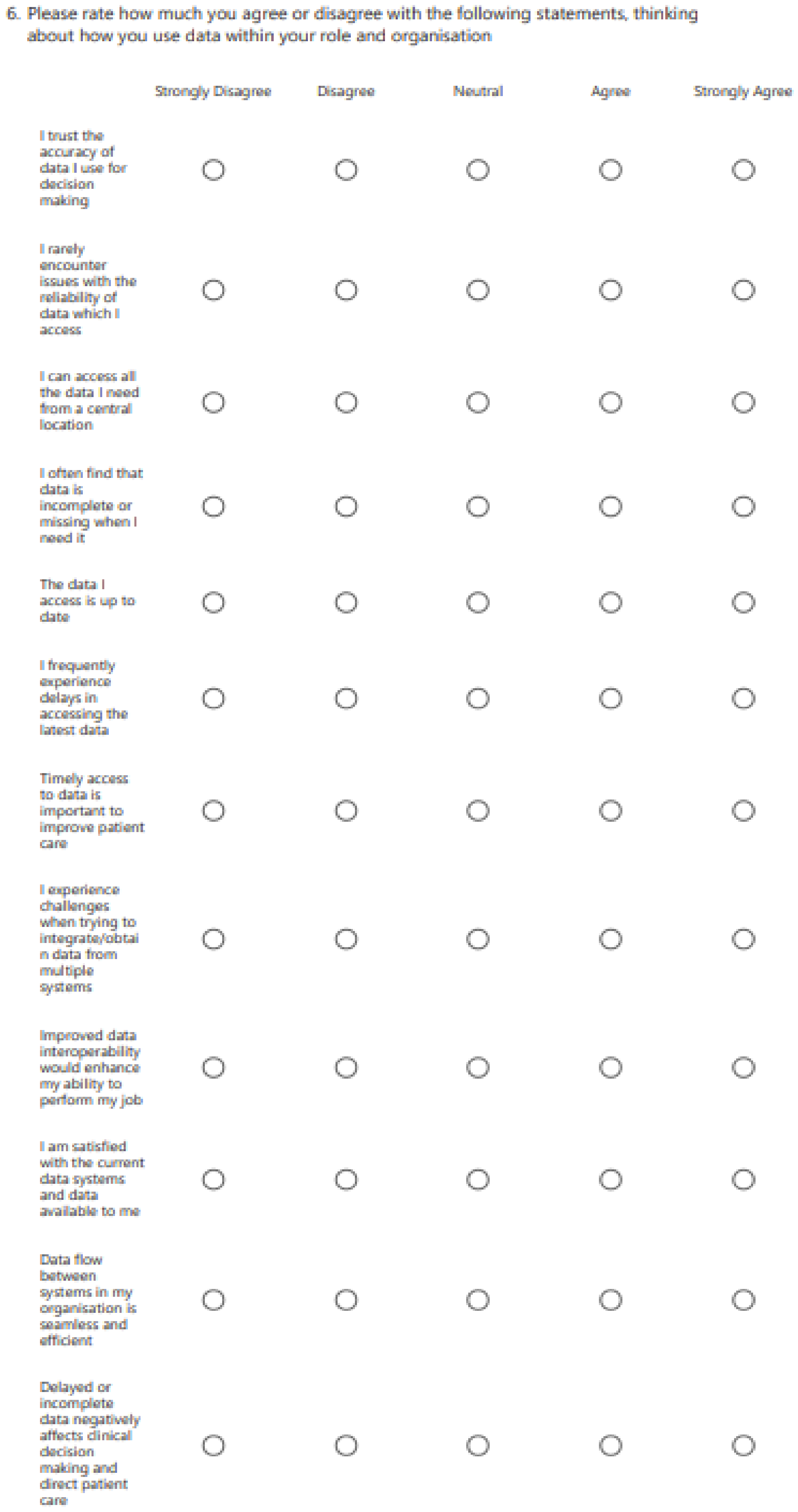
☐ Cloud Reports e.g. PowerBI

☐ Excel (including CSV files)

☐ Paper Notes

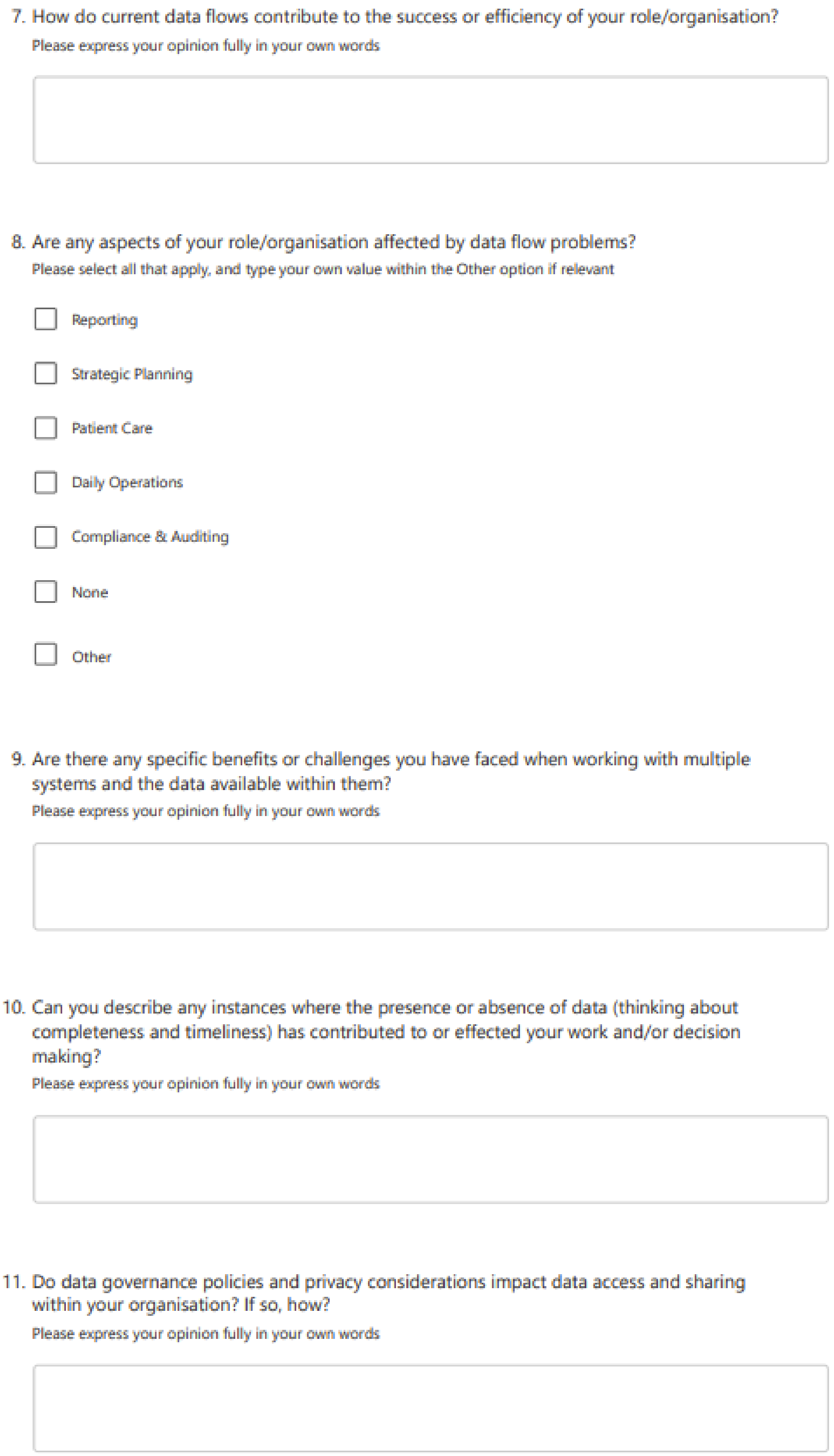
☐ PowerApps / MS Forms

☐ Other



6. Please rate how much you agree or disagree with the following statements, thinking about how you use data within your role and organisation

	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
I trust the accuracy of data I use for decision making	○	○	○	○	○
I rarely encounter issues with the reliability of data which I access	○	○	○	○	○
I can access all the data I need from a central location	○	○	○	○	○
I often find that data is incomplete or missing when I need it	○	○	○	○	○
The data I access is up to date	○	○	○	○	○
I frequently experience delays in accessing the latest data	○	○	○	○	○
Timely access to data is important to improve patient care	○	○	○	○	○
I experience challenges when trying to integrate/obtain data from multiple systems	○	○	○	○	○
Improved data interoperability would enhance my ability to perform my job	○	○	○	○	○
I am satisfied with the current data systems and data available to me	○	○	○	○	○
Data flow between systems in my organisation is seamless and efficient	○	○	○	○	○
Delayed or incomplete data negatively affects clinical decision making and direct patient care	○	○	○	○	○



7. How do current data flows contribute to the success or efficiency of your role/organisation?
Please express your opinion fully in your own words

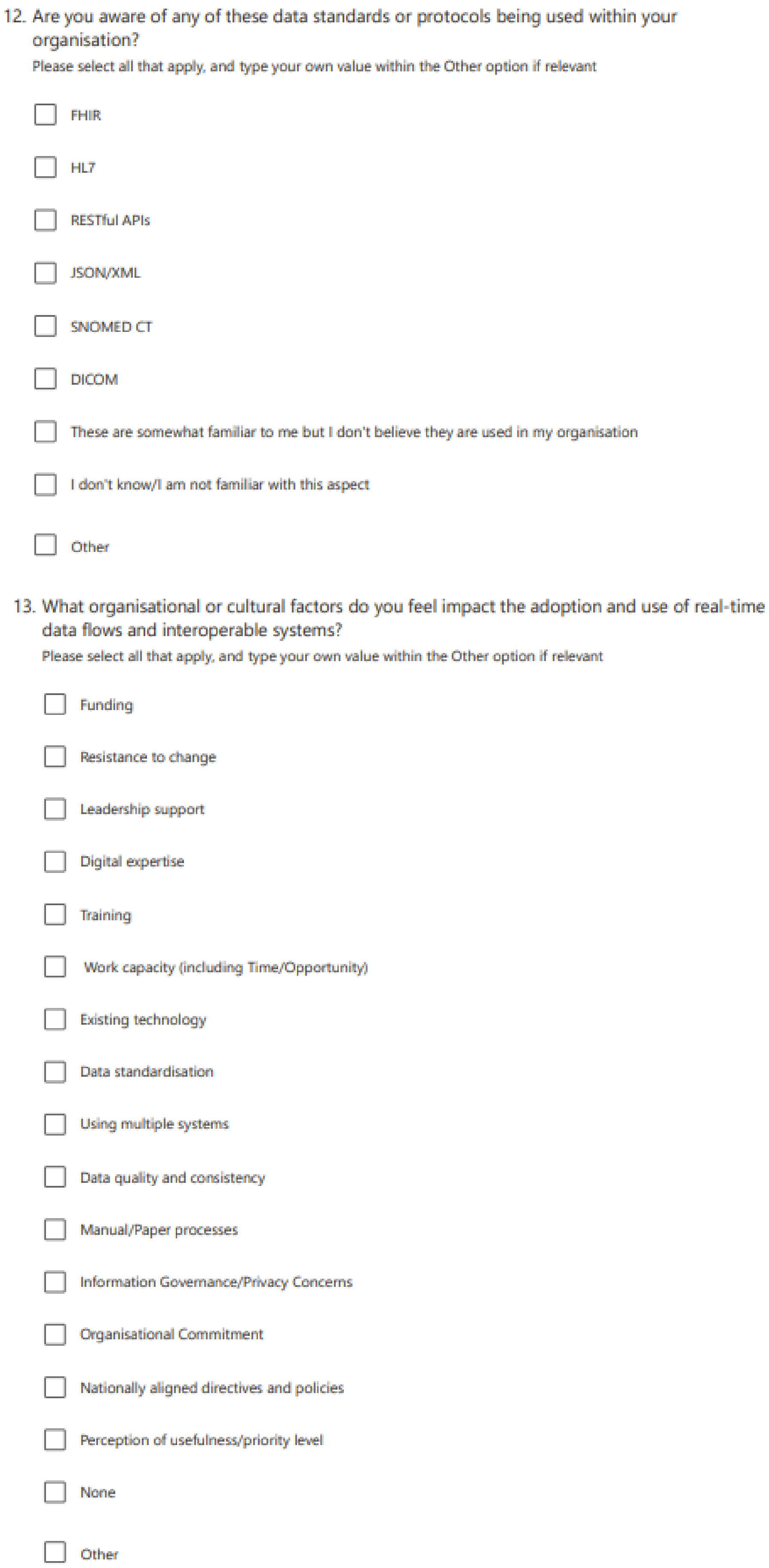
8. Are any aspects of your role/organisation affected by data flow problems?
Please select all that apply, and type your own value within the Other option if relevant

☐ Reporting
☐ Strategic Planning
☐ Patient Care
☐ Daily Operations
☐ Compliance & Auditing
☐ None
☐ Other

9. Are there any specific benefits or challenges you have faced when working with multiple systems and the data available within them?
Please express your opinion fully in your own words

10. Can you describe any instances where the presence or absence of data (thinking about completeness and timeliness) has contributed to or effected your work and/or decision making?
Please express your opinion fully in your own words

11. Do data governance policies and privacy considerations impact data access and sharing within your organisation? If so, how?
Please express your opinion fully in your own words



12. Are you aware of any of these data standards or protocols being used within your organisation?

Please select all that apply, and type your own value within the Other option if relevant

☐ FHIR

☐ HL7

☐ RESTful APIs

☐ JSON/XML

☐ SNOMED CT

☐ DICOM

☐ These are somewhat familiar to me but I don't believe they are used in my organisation

☐ I don't know/I am not familiar with this aspect

☐ Other

13. What organisational or cultural factors do you feel impact the adoption and use of real-time data flows and interoperable systems?

Please select all that apply, and type your own value within the Other option if relevant

☐ Funding

☐ Resistance to change

☐ Leadership support

☐ Digital expertise

☐ Training

☐ Work capacity (including Time/Opportunity)

☐ Existing technology

☐ Data standardisation

☐ Using multiple systems

☐ Data quality and consistency

☐ Manual/Paper processes

☐ Information Governance/Privacy Concerns

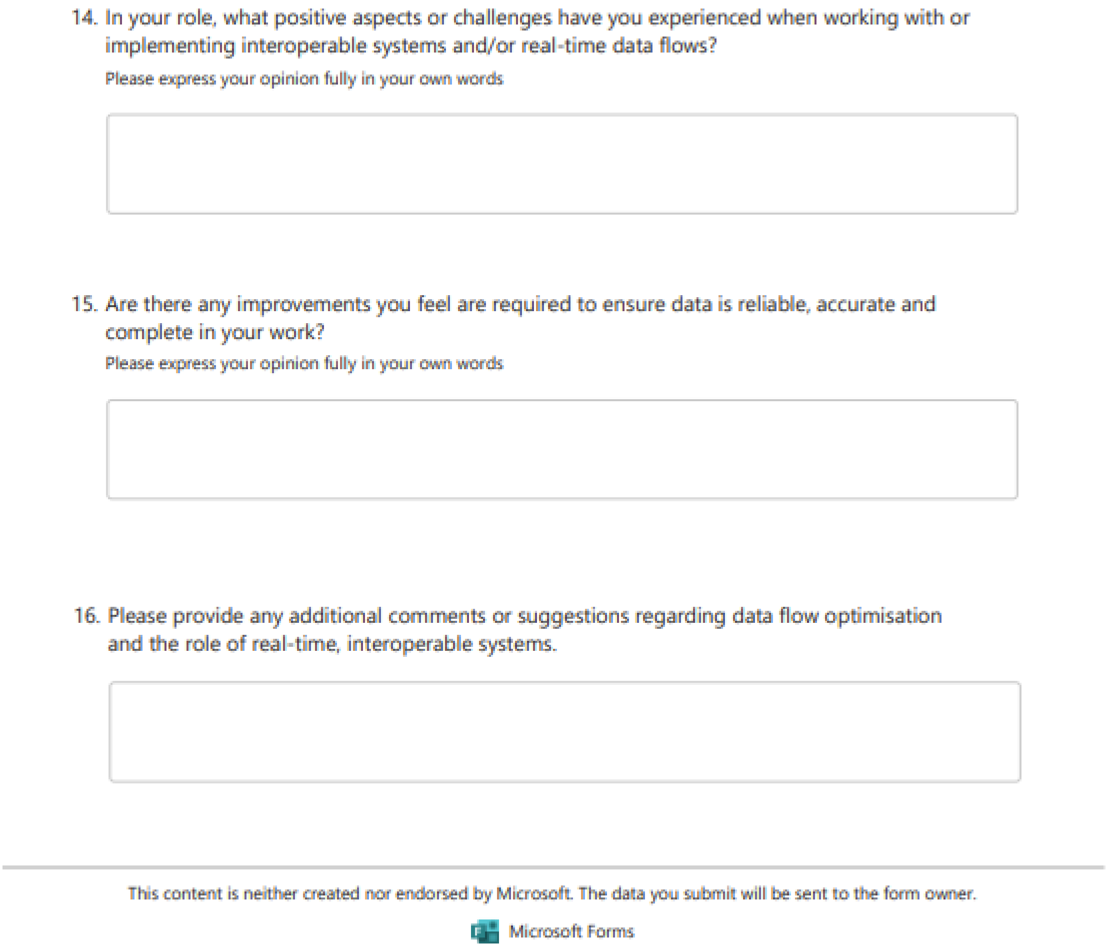
☐ Organisational Commitment

☐ Nationally aligned directives and policies

☐ Perception of usefulness/priority level

☐ None

☐ Other



14. In your role, what positive aspects or challenges have you experienced when working with or implementing interoperable systems and/or real-time data flows?

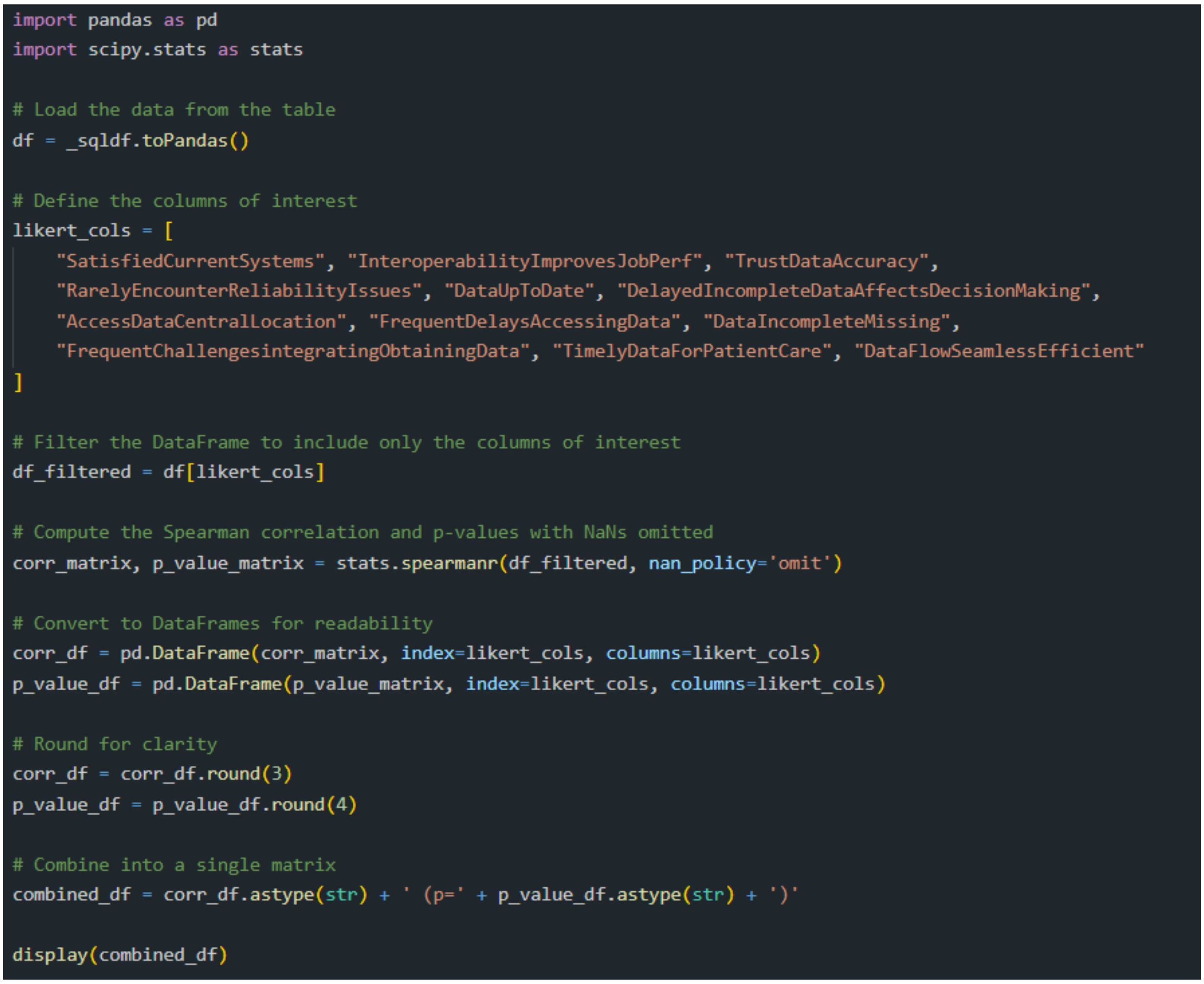
Please express your opinion fully in your own words

15. Are there any improvements you feel are required to ensure data is reliable, accurate and complete in your work?

Please express your opinion fully in your own words

16. Please provide any additional comments or suggestions regarding data flow optimisation and the role of real-time, interoperable systems.

Appendix 9. Python (Pyspark) Code and Full Results of Likert-Scale Question Pair Correlation



```
import pandas as pd
import scipy.stats as stats

# Load the data from the table
df = _sqldf.toPandas()

# Define the columns of interest
likert_cols = [
    "SatisfiedCurrentSystems", "InteroperabilityImprovesJobPerf", "TrustDataAccuracy",
    "RarelyEncounterReliabilityIssues", "DataUpToDate", "DelayedIncompleteDataAffectsDecisionMaking",
    "AccessDataCentralLocation", "FrequentDelaysAccessingData", "DataIncompleteMissing",
    "FrequentChallengesintegratingObtainingData", "TimelyDataForPatientCare", "DataFlowSeamlessEfficient"
]

# Filter the DataFrame to include only the columns of interest
df_filtered = df[likert_cols]

# Compute the Spearman correlation and p-values with NaNs omitted
corr_matrix, p_value_matrix = stats.spearmanr(df_filtered, nan_policy='omit')

# Convert to DataFrames for readability
corr_df = pd.DataFrame(corr_matrix, index=likert_cols, columns=likert_cols)
p_value_df = pd.DataFrame(p_value_matrix, index=likert_cols, columns=likert_cols)

# Round for clarity
corr_df = corr_df.round(3)
p_value_df = p_value_df.round(4)

# Combine into a single matrix
combined_df = corr_df.astype(str) + ' (p=' + p_value_df.astype(str) + ')'

display(combined_df)
```

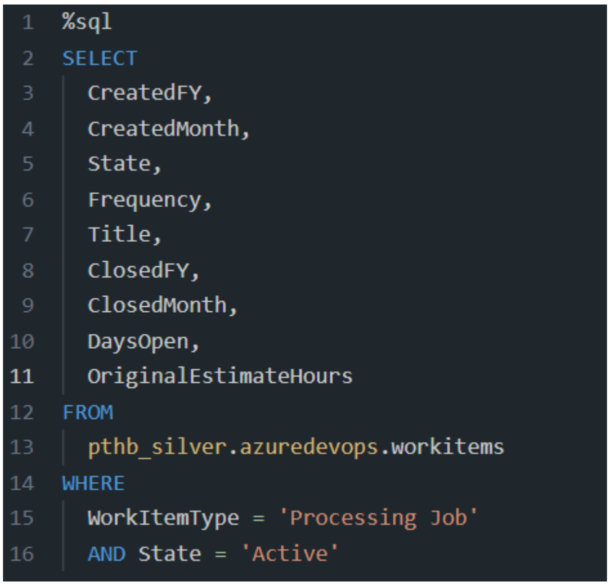
Table of Likert-Scale Pairwise Spearman Correlation Coefficient Results

	I am satisfied with the current data systems and data available to me	Improved data interoperability would enhance my ability to perform my job	I trust the accuracy of data I use for decision making	I rarely encounter issues with the reliability of data which I access	The data I access is up to date	Delayed or incomplete data negatively affects clinical decision making and direct patient care	I can access all the data I need from a central location	I often find that data is incomplete or missing when I need it	I frequently experience delays in accessing the latest data	I experience challenges when trying to integrate/obtain data from multiple systems	Timely access to data is important to improve patient care	Data flow between systems in my organisation is seamless and efficient
I am satisfied with the current data systems and data available to me	1.0 (p=0.0)	-0.292 (p=0.0043)	0.194 (p=0.0608)	0.473 (p=0.0)	0.28 (p=0.0065)	-0.182 (p=0.0798)	0.395 (p=0.0001)	0.342 (p=0.0007)	0.424 (p=0.0)	0.418 (p=0.0)	-0.189 (p=0.0684)	0.499 (p=0.0)
Improved data interoperability would enhance my ability to perform my job	-0.292 (p=0.0043)	1.0 (p=0.0)	-0.004 (p=0.9724)	-0.198 (p=0.0542)	-0.02 (p=0.8455)	0.131 (p=0.2048)	-0.336 (p=0.0009)	-0.203 (p=0.0489)	-0.273 (p=0.0074)	-0.515 (p=0.0)	0.412 (p=0.0)	-0.277 (p=0.0066)
I trust the accuracy of data I use for decision making	0.194 (p=0.0608)	-0.004 (p=0.9724)	1.0 (p=0.0)	0.434 (p=0.0)	0.26 (p=0.0114)	0.017 (p=0.8678)	0.168 (p=0.104)	0.265 (p=0.0095)	0.307 (p=0.0025)	-0.071 (p=0.4924)	-0.075 (p=0.47)	0.166 (p=0.1087)
I rarely encounter issues with	0.473 (p=0.0)	-0.198 (p=0.0542)	0.434 (p=0.0)	1.0 (p=0.0)	0.387	-0.135	0.279	0.488 (p=0.0)	0.505 (p=0.0)	0.262 (p=0.0104)	-0.067	0.318

the reliability of data which I access					(p=0.0001)	(p=0.1921)	(p=0.0062)				(p=0.5176)	(p=0.0017)
The data I access is up to date	0.28 (p=0.0065)	-0.02 (p=0.8455)	0.26 (p=0.0114)	0.387 (p=0.0001)	1.0 (p=0.0)	0.001 (p=0.9891)	0.19 (p=0.0661)	0.472 (p=0.0)	0.424 (p=0.0)	0.015 (p=0.8886)	0.002 (p=0.9855)	0.153 (p=0.1418)
Delayed or incomplete data negatively affects clinical decision making and direct patient care	-0.182 (p=0.0798)	0.131 (p=0.2048)	0.017 (p=0.8678)	-0.135 (p=0.1921)	0.001 (p=0.9891)	1.0 (p=0.0)	-0.054 (p=0.6039)	-0.014 (p=0.8917)	0.017 (p=0.8727)	-0.186 (p=0.0707)	0.263 (p=0.01)	-0.223 (p=0.03)
I can access all the data I need from a central location	0.395 (p=0.0001)	-0.336 (p=0.0009)	0.168 (p=0.104)	0.279 (p=0.0062)	0.19 (p=0.0661)	-0.054 (p=0.6039)	1.0 (p=0.0)	0.401 (p=0.0001)	0.418 (p=0.0)	0.524 (p=0.0)	-0.034 (p=0.7414)	0.335 (p=0.0009)
I frequently experience delays in accessing the latest data	0.342 (p=0.0007)	-0.203 (p=0.0489)	0.265 (p=0.0095)	0.488 (p=0.0)	0.472 (p=0.0)	-0.014 (p=0.8917)	0.401 (p=0.0001)	1.0 (p=0.0)	0.604 (p=0.0)	0.325 (p=0.0013)	-0.143 (p=0.1668)	0.297 (p=0.0034)
I often find that data is incomplete or missing when I need it	0.424 (p=0.0)	-0.273 (p=0.0074)	0.307 (p=0.0025)	0.505 (p=0.0)	0.424 (p=0.0)	0.017 (p=0.8727)	0.418 (p=0.0)	0.604 (p=0.0)	1.0 (p=0.0)	0.355 (p=0.0004)	-0.052 (p=0.6152)	0.419 (p=0.0)
I experience challenges when trying to integrate/obt	0.418 (p=0.0)	-0.515 (p=0.0)	-0.071	0.262	0.015	-0.186	0.524 (p=0.0)	0.325	0.355	1.0 (p=0.0)	-0.161	0.313 (p=0.002)

ain data from multiple systems			(p=0.4924)	(p=0.0104)	(p=0.8886)	(p=0.0707)		(p=0.0013)	(p=0.0004)		(p=0.1187)	
Timely access to data is important to improve patient care	-0.189 (p=0.0684)	0.412 (p=0.0)	-0.075 (p=0.47)	-0.067 (p=0.5176)	0.002 (p=0.9855)	0.263 (p=0.01)	-0.034 (p=0.7414)	-0.143 (p=0.1668)	-0.052 (p=0.6152)	-0.161 (p=0.1187)	1.0 (p=0.0)	-0.056 (p=0.5881)
Data flow between systems in my organisation is seamless and efficient	0.499 (p=0.0)	-0.277 (p=0.0066)	0.166 (p=0.1087)	0.318 (p=0.0017)	0.153 (p=0.1418)	-0.223 (p=0.03)	0.335 (p=0.0009)	0.297 (p=0.0034)	0.419 (p=0.0)	0.313 (p=0.002)	-0.056 (p=0.5881)	1.0 (p=0.0)

Appendix 10. SQL Code for PTHB Data Analysis and Engineering Team Azure DevOps



```sql
%sql
SELECT
  CreatedFY,
  CreatedMonth,
  State,
  Frequency,
  Title,
  ClosedFY,
  ClosedMonth,
  DaysOpen,
  OriginalEstimateHours
FROM
  pthb_silver.azuredevops.workitems
WHERE
  WorkItemType = 'Processing Job'
  AND State = 'Active'
```

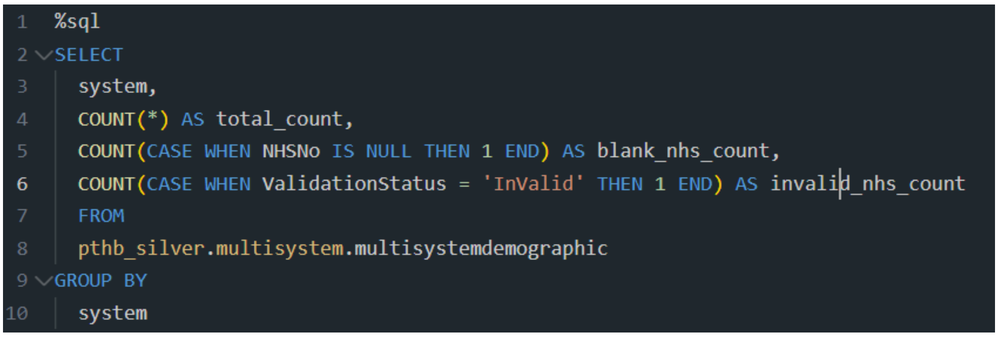
Appendix 11. SQL Code for PTHB Activity Duplicated Between WPAS and WCCIS

Duplication and join based on activity in both systems recorded on the same person, on the same day by the specialty. The results were then aggregated together and restricted to the last five complete calendar years, as displayed in Table 13.

```sql
SELECT DISTINCT WCCIS.NHSNumber,
  WCCIS.SpecGroup,
  WCCIS.TRT_DATE,
  WCCIS.ReportingConsultant,
  WCCIS.ReportTeam,
  WCCIS.ACTIVITY_TYPE,
  WCCIS.ContactType,
  wccis.ActivityCount,
  wccis.Specialty AS WCCISSpec,
  wpas.ActivityCount,
  wpas.SPEC_NAME,
  wpas.CONS_NAME,
  wpas.TREATTYPE_DESC,
  wpas.TREATTYPE_CONTACT_TYPE,
  wpas.actnotekey,
  SPLIT(IFNULL(WCCIS.ReportingConsultant, ''), ' ') [SIZE(SPLIT(IFNULL(WCCIS.ReportingConsultant, ''), ' ')) - 1] AS WCCISSurname,
  SPLIT(IFNULL(wpas.CONS_NAME, ''), ' ') [SIZE(SPLIT(IFNULL(wpas.CONS_NAME, ''), ' ')) - 1] AS WPASSurname
FROM wccissactivitysummary as wccis
INNER JOIN (SELECT *,
            COUNT(*) OVER(PARTITION BY a.NHS, a.TRT_DATE, a.SpecGroup) AS ActivityCount,
            ROW_NUMBER() OVER(PARTITION BY a.NHS, a.TRT_DATE, a.SpecGroup ORDER BY a.TRT_DATE DESC) = 1 AS MostRecentFlag FROM wpasactivitysummary as a) as wpas
  ON wccis.NHSNumber = wpas.NHS
              and wccis.SpecGroup = wpas.SpecGroup
              and wccis.TRT_DATE = wpas.TRT_DATE
  WHERE wccis.MostRecentActivityFlag = TRUE and WPAS.MostRecentFlag = TRUE
```

Appendix 12. SQL Code for PTHB Multiple EHR System Demographics

SQL Code for Count of Records Per EHR Systems and NHS Number Validity



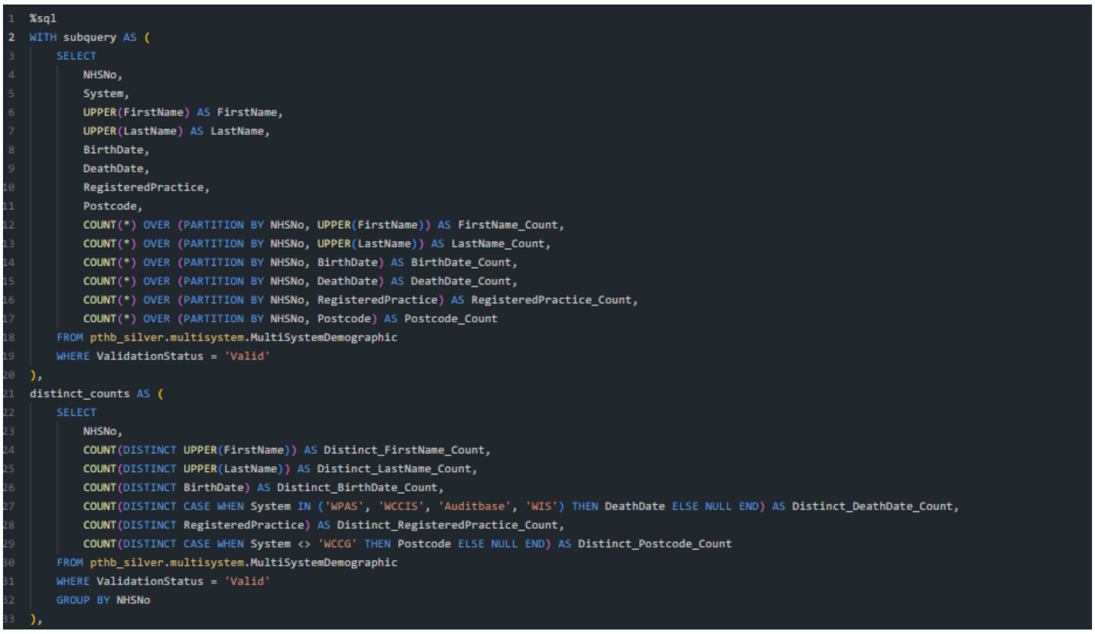
```sql
%sql
SELECT
  system,
  COUNT(*) AS total_count,
  COUNT(CASE WHEN NHSNo IS NULL THEN 1 END) AS blank_nhs_count,
  COUNT(CASE WHEN ValidationStatus = 'InValid' THEN 1 END) AS invalid_nhs_count
  FROM
  pthb_silver.multisystem.multisystemdemographic
GROUP BY
  system
```

SQL Code for Duplicate Identification Across EHR Systems

```sql
%sql
SELECT system,
  COUNT(*) - COUNT(DISTINCT NHSNo) AS NHS_duplicates_count
FROM
  pthb_silver.multisystem.multisystemdemographic
WHERE ValidationStatus = 'Valid'
GROUP BY
  system;

SELECT system,
  COUNT(*) - COUNT(DISTINCT FuzzyIdentifier) AS FuzzyIdentifier_duplicates_count
FROM
  pthb_silver.multisystem.multisystemdemographic
WHERE FuzzyIdentifier IS NOT NULL
GROUP BY
  system;
```

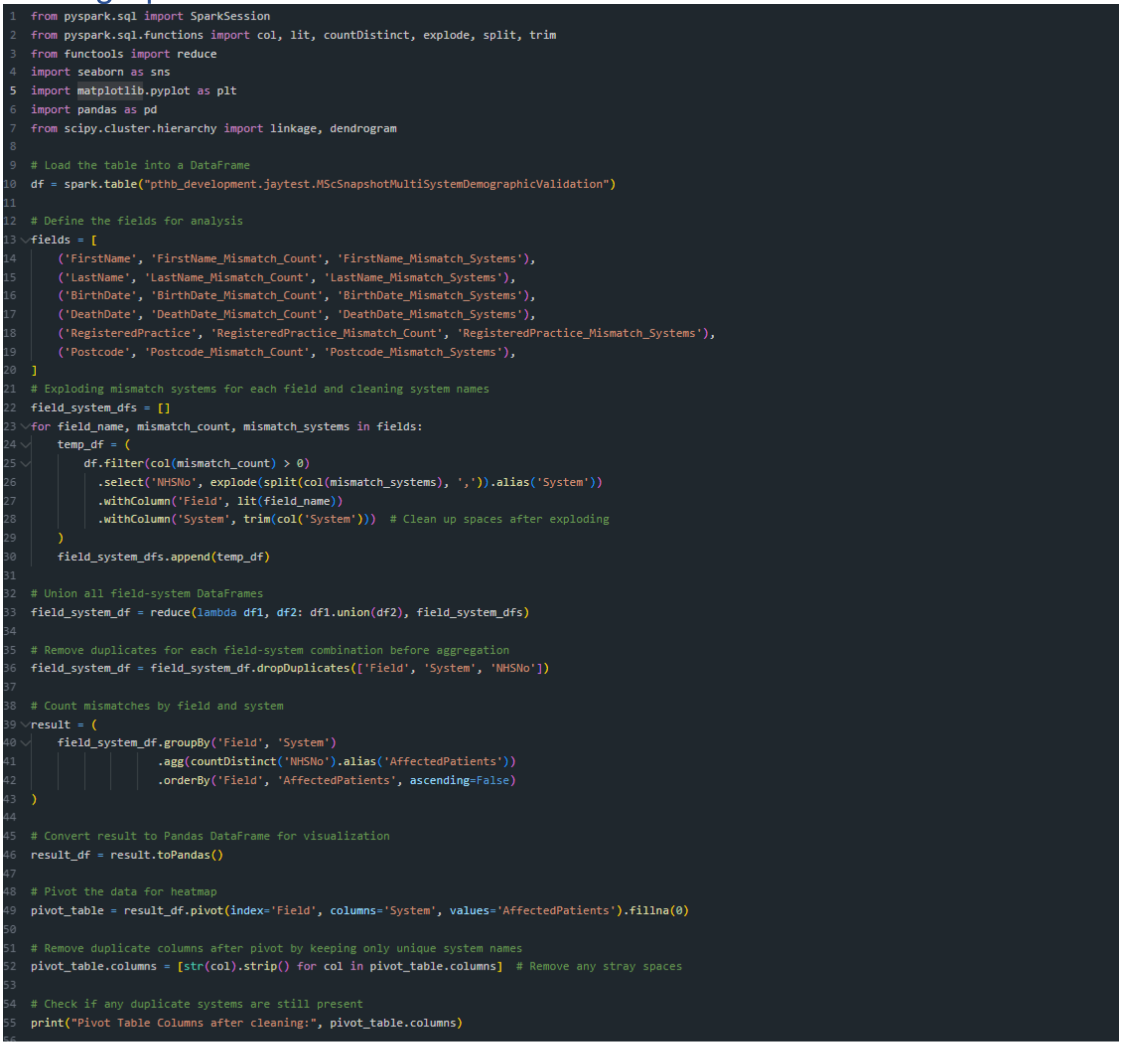
SQL Code for Demographic Matching Across EHR Systems



```sql
%sql
WITH subquery AS (
    SELECT
        NHSNo,
        System,
        UPPER(FirstName) AS FirstName,
        UPPER(LastName) AS LastName,
        BirthDate,
        DeathDate,
        RegisteredPractice,
        Postcode,
        COUNT(*) OVER (PARTITION BY NHSNo, UPPER(FirstName)) AS FirstName_Count,
        COUNT(*) OVER (PARTITION BY NHSNo, UPPER(LastName)) AS LastName_Count,
        COUNT(*) OVER (PARTITION BY NHSNo, BirthDate) AS BirthDate_Count,
        COUNT(*) OVER (PARTITION BY NHSNo, DeathDate) AS DeathDate_Count,
        COUNT(*) OVER (PARTITION BY NHSNo, RegisteredPractice) AS RegisteredPractice_Count,
        COUNT(*) OVER (PARTITION BY NHSNo, Postcode) AS Postcode_Count
    FROM pthb_silver.multisystem.MultiSystemDemographic
    WHERE ValidationStatus = 'Valid'
),
distinct_counts AS (
    SELECT
        NHSNo,
        COUNT(DISTINCT UPPER(FirstName)) AS Distinct_FirstName_Count,
        COUNT(DISTINCT UPPER(LastName)) AS Distinct_LastName_Count,
        COUNT(DISTINCT BirthDate) AS Distinct_BirthDate_Count,
        COUNT(DISTINCT CASE WHEN System IN ('WPAS', 'WCCIS', 'Auditbase', 'WIS') THEN DeathDate ELSE NULL END) AS Distinct_DeathDate_Count,
        COUNT(DISTINCT RegisteredPractice) AS Distinct_RegisteredPractice_Count,
        COUNT(DISTINCT CASE WHEN System <> 'WCCG' THEN Postcode ELSE NULL END) AS Distinct_Postcode_Count
    FROM pthb_silver.multisystem.MultiSystemDemographic
    WHERE ValidationStatus = 'Valid'
    GROUP BY NHSNo
),
```

There are some exceptions made due to data availability; Postcode is not present in the WCCG data, and Death Date can only be compared between WPAS, WCCIS, Auditbase and WIS.

Appendix 13. Python (Pyspark) Code for PTHB Multiple EHR System Demographic Correlation



```python
from pyspark.sql import SparkSession
from pyspark.sql.functions import col, lit, countDistinct, explode, split, trim
from functools import reduce
import seaborn as sns
import matplotlib.pyplot as plt
import pandas as pd
from scipy.cluster.hierarchy import linkage, dendrogram

# Load the table into a DataFrame
df = spark.table("pthb_development.jaytest.MScSnapshotMultiSystemDemographicValidation")

# Define the fields for analysis
fields = [
    ('FirstName', 'FirstName_Mismatch_Count', 'FirstName_Mismatch_Systems'),
    ('LastName', 'LastName_Mismatch_Count', 'LastName_Mismatch_Systems'),
    ('BirthDate', 'BirthDate_Mismatch_Count', 'BirthDate_Mismatch_Systems'),
    ('DeathDate', 'DeathDate_Mismatch_Count', 'DeathDate_Mismatch_Systems'),
    ('RegisteredPractice', 'RegisteredPractice_Mismatch_Count', 'RegisteredPractice_Mismatch_Systems'),
    ('Postcode', 'Postcode_Mismatch_Count', 'Postcode_Mismatch_Systems'),
]
# Exploding mismatch systems for each field and cleaning system names
field_system_dfs = []
for field_name, mismatch_count, mismatch_systems in fields:
    temp_df = (
        df.filter(col(mismatch_count) > 0)
          .select('NHSNo', explode(split(col(mismatch_systems), ',')).alias('System'))
          .withColumn('Field', lit(field_name))
          .withColumn('System', trim(col('System')))  # Clean up spaces after exploding
    )
    field_system_dfs.append(temp_df)

# Union all field-system DataFrames
field_system_df = reduce(lambda df1, df2: df1.union(df2), field_system_dfs)

# Remove duplicates for each field-system combination before aggregation
field_system_df = field_system_df.dropDuplicates(['Field', 'System', 'NHSNo'])

# Count mismatches by field and system
result = (
    field_system_df.groupBy('Field', 'System')
                   .agg(countDistinct('NHSNo').alias('AffectedPatients'))
                   .orderBy('Field', 'AffectedPatients', ascending=False)
)

# Convert result to Pandas DataFrame for visualization
result_df = result.toPandas()

# Pivot the data for heatmap
pivot_table = result_df.pivot(index='Field', columns='System', values='AffectedPatients').fillna(0)

# Remove duplicate columns after pivot by keeping only unique system names
pivot_table.columns = [str(col).strip() for col in pivot_table.columns]  # Remove any stray spaces

# Check if any duplicate systems are still present
print("Pivot Table Columns after cleaning:", pivot_table.columns)
```

```python

# Plot heatmap of mismatches by field and system
plt.figure(figsize=(10, 6))
sns.heatmap(pivot_table, annot=True, fmt="g", cmap="YlGnBu")
plt.title("Mismatch Counts by Field and System")
plt.xlabel("System")
plt.ylabel("Field")
plt.show()

# Normalize the data (e.g., calculate percentages or z-scores)
normalized_data = (pivot_table - pivot_table.mean()) / pivot_table.std()

# Calculate the correlation matrix
correlation_matrix = normalized_data.corr()  # Correlation between systems

# Visualize the correlation matrix
plt.figure(figsize=(12, 8))
sns.heatmap(correlation_matrix, annot=True, fmt=".2f", cmap="coolwarm", cbar=True)
plt.title("Correlation Between Systems Based on Mismatch Counts")
plt.xlabel("System")
plt.ylabel("System")
plt.show()

# Optional: Hierarchical clustering
linkage_matrix = linkage(correlation_matrix, method='ward')
plt.figure(figsize=(10, 6))
dendrogram(linkage_matrix, labels=correlation_matrix.columns, leaf_rotation=90)
plt.title("Hierarchical Clustering of Systems Based on Correlations")
plt.xlabel("System")
plt.ylabel("Distance")
plt.show()

# Pearson correlation between Match_Count and Mismatch_Count
from pyspark.sql.functions import corr
correlation_df = result.select(corr("AffectedPatients", "AffectedPatients").alias("Correlation"))
correlation_df.show()
```

Appendix 14. Welsh Health Board Allocations Summary: Baseline discretionary capital funding 2025-26

Welsh Government (2025) Health Board Allocations: Revenue Tables. Available at: https://www.gov.wales/sites/default/files/publications/2025-01/health-board-allocations-revenue-tables.xlsx [Accessed 27 March 2025]

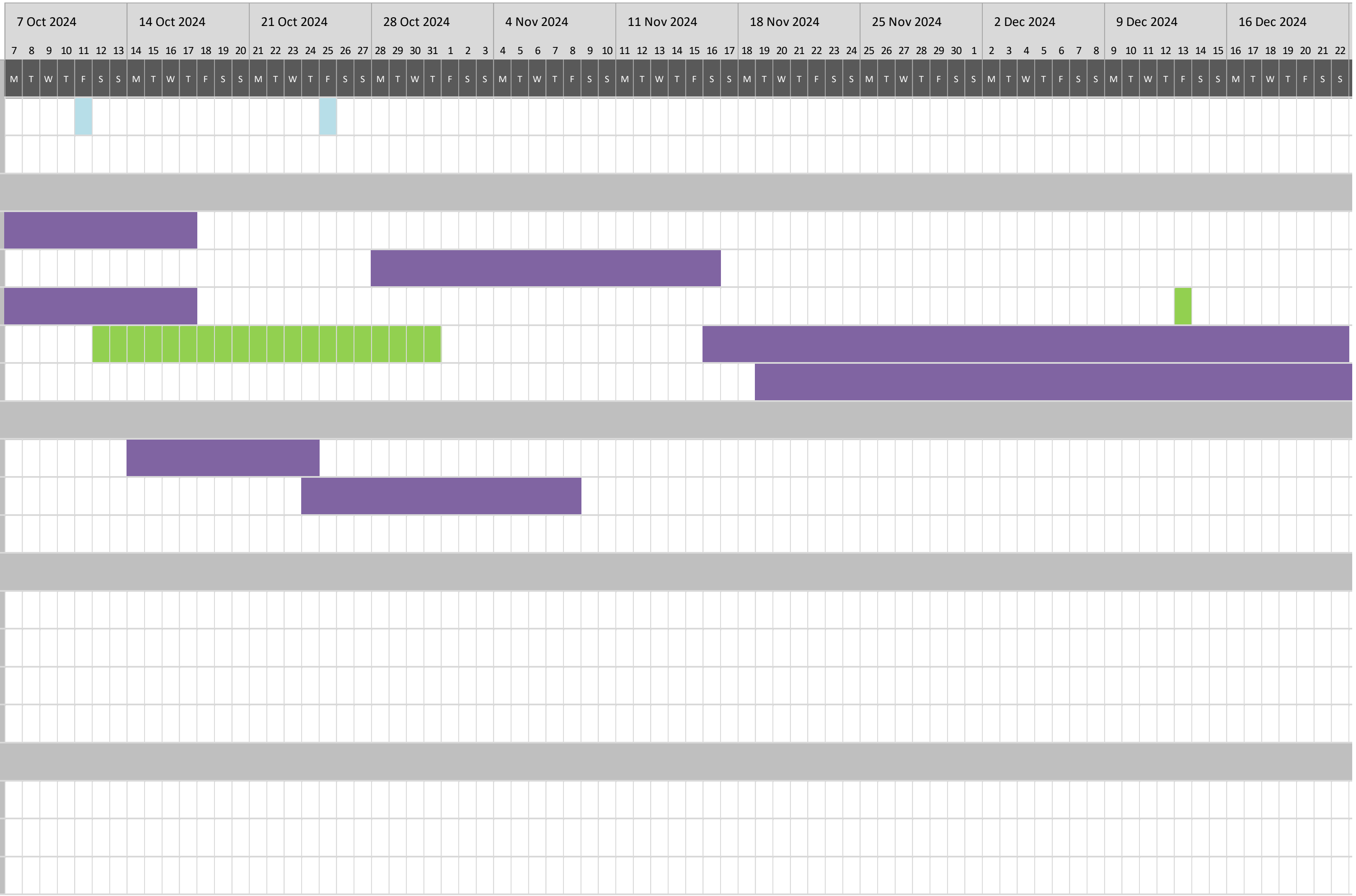
Health Boards	**2025 26 Baseline discretionary capital funding £m**
Aneurin Bevan HB	12.875
Betsi Cadwaladr University HB	17
Cardiff and Vale University HB	17
Cwm Taf Morgannwg HB	12
Hywel Dda HB	10
Powys HB	2.7
Swansea Bay HB	13.875
HEIW	0.1
Digital Health Care Wales	3.25
Total (Health Boards)	**88.8**
NHS Trusts	
Velindre	2
NHS Wales Shared Services Partnership	1.25
Public Health Wales	1.7
Welsh Ambulance Services	6.25
Total (NHS Trusts)	11.2
Total	**100**

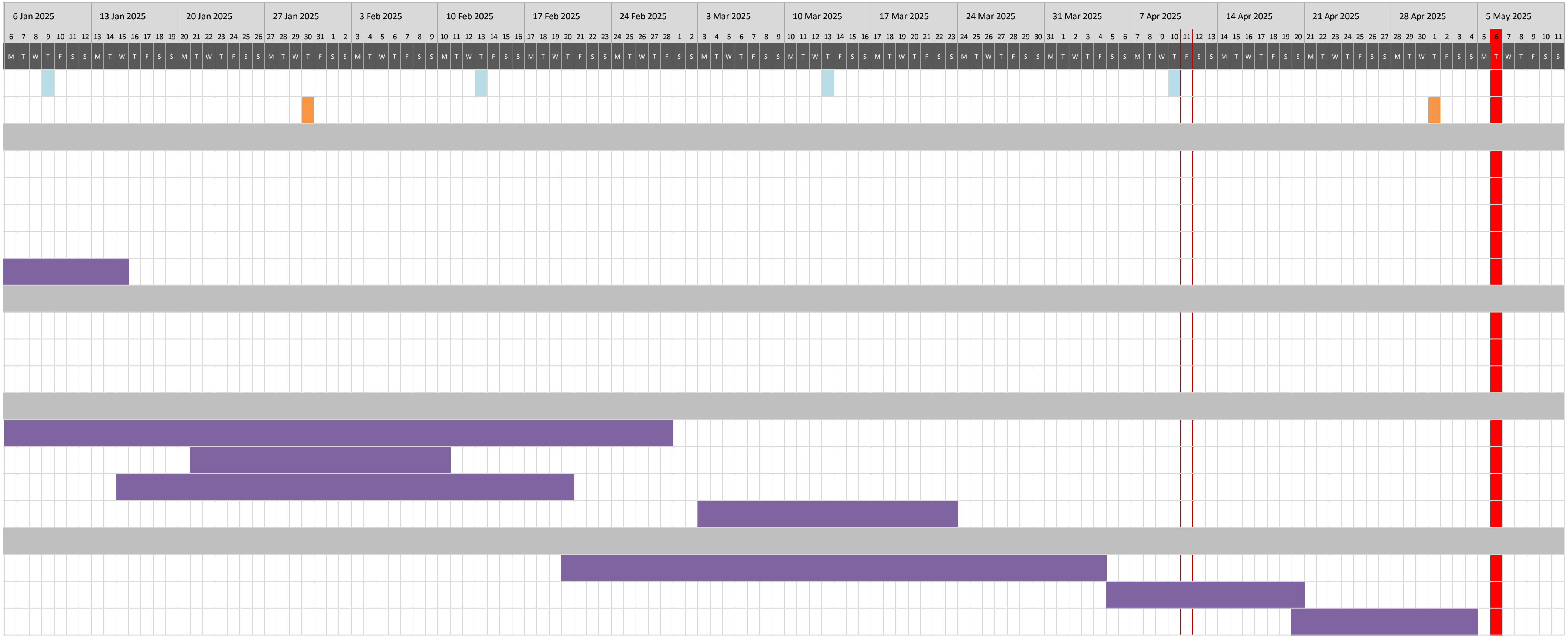
Appendix 15. MSc Project Plan Gantt Chart

Mon, 07/10/2024

1

TASK	START	END
Supervisions/Meetings		
Progress Meeting/Presentation	TBC	
Introduction		
Project brief	7/10/24	17/10/24
Start Literature review	28/10/24	16/11/24
Ethics application - Approval	7/10/24	17/10/24
Draft Introduction Chapter	16/11/24	22/12/24
Update Literature Review Chapter	19/11/24	15/1/25
Methods		
Methods selection and justification	14/10/24	24/10/24
Complete draft Methods	24/10/24	8/11/24
Update Methods Chapter	20/1/24	25/2/24
Data Collection & Analysis		
Primary Data Collection	6/1/25	28/2/25
Start primary data analysis	21/1/25	10/2/25
Draft Results chapter	15/1/25	20/2/25
Update Results Chapter	3/3/25	23/3/25
Discussion & Conclusion		
Draft Discusison Chapter	20/2/25	4/4/25
Outline Conclusions	5/4/25	20/4/25
Complete dissertation and final revision	20/4/25	4/5/25




6 Jan 2025
13 Jan 2025
20 Jan 2025
27 Jan 2025
3 Feb 2025
10 Feb 2025
17 Feb 2025
24 Feb 2025
3 Mar 2025
10 Mar 2025
17 Mar 2025
24 Mar 2025
31 Mar 2025
7 Apr 2025
14 Apr 2025
21 Apr 2025
28 Apr 2025
5 May 2025