



Real-world learnings for digital health industry–NHS collaboration: Life sciences vision in action[☆]

Rebecca Pope^a, Alexandros Zenonos^a, William Bryant^b, Anastasia Spiridou^b, Daniel Key^b, Shiren Patel^b, Jack Robinson^a, Anna Styles^a, Chris Rockenbach^b, Gina Bicknell^c, Pavithra Rajendran^b, Andrew M. Taylor^b, Neil J. Sebire^b

^a Roche Products Limited, United Kingdom

^b Great Ormond Street Hospital for Children NHS Foundation Trust and NIHR GOSH Biomedical Research Centre, United Kingdom

^c Pinsent Masons LLP, United Kingdom

ARTICLE INFO

Keywords:

Life sciences industry strategy
Informatics
Information technology
Partnership
Collaboration

ABSTRACT

Several publications have indicated potential benefit from collaboration with industry regarding wider use of anonymised routine NHS healthcare data. However, there is limited guidance regarding exactly how such collaborations between NHS hospitals and industry partners should best be carried out, and specific issues that need to be addressed at an individual project or collaboration level to achieve desired benefit. Specifically, routine health data are complex, not collected in a format optimised for secondary use, and often require interpretation based on clinical understanding of the medical conditions or patients.

In order to address these issues, a formal partnership collaboration was established between an NHS organisation (Great Ormond Street Hospital for Children) and a pharmaceutical company (Roche Products Limited), to jointly understand the problems that require solving in order to maximise such use of NHS data to support improved patient outcomes and other patient/NHS benefit in a more sustainable way.

We present the learnings from the first 2 years of the 5-year collaboration addressing aspects such as complexities of NHS Electronic Patient Record (EPR), data engineering and use of modern technology to optimise such data. Plus, the development of appropriate technology and data infrastructure within the NHS to support interoperability and prepare the NHS for wider application of artificial intelligence. We also highlight the staff skills and training needed to support such systems in the NHS, governance structures and processes needed to ensure appropriate use of tools and data and how best to co-design with patients, their families, and clinical teams. It is hoped that this review may provide useful information for both healthcare organisations and industry partners working towards the future of optimal use of data and technology for healthcare benefit.

Introduction

In 2017, as the UK government planned its future outside the European Union, the Life Sciences Industrial Strategy was published,¹ to position the UK to take advantage of future health technology trends. This strategy sought to address challenges under five key themes such as science and growth, but specifically included NHS (National Health Service)–industry collaboration, facilitating better patient care through innovative treatments and technologies, and optimising use of data and digital tools to support research and care. The strategy proposed that the NHS should work in a ‘new collaborative environment where industry and the health system work together, underpinned by the rich datasets that the NHS can now provide’, to support discovery and development

of new medicines, diagnostics and MedTech capabilities, and catalysing medical innovation using data analytics, artificial intelligence (AI) and engineering.

Subsequently, there have been several UK government publications focused on data and digital technology in the NHS, strategic and technical,^{2–10} with common overarching themes including:¹¹ reducing the places that data are collected, stored and disseminated; making data accessible to trusted individuals, teams, organisations and businesses (where appropriate) via platforms (secure data environments, SDEs); upskilling clinical, analytical, academic and senior management for digital and data literacy; establishing governance, standards and mechanisms ensuring that NHS data can be securely accessible for those who need it, while deriving patient benefit; establishing mechanisms for trans-

This article reflects the opinions of the author(s) and should not be taken to represent the policy of the Royal College of Physicians unless specifically stated.

[☆] This project has been completed by Great Ormond Street Hospital NHS Foundation Trust and Roche Products Ltd as part of a collaborative working agreement. Roche Products Ltd had no influence on the results or decision to publish regarding this work. M-GB-00015553 | 2024.

<https://doi.org/10.1016/j.fhj.2024.100168>

Received 18 March 2024; Received in revised form 23 July 2024; Accepted 26 July 2024

2514-6645/© 2024 The Authors. Published by Elsevier Ltd on behalf of Royal College of Physicians. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

parency and accountability over data use and standardising how patients and the public are informed and involved. The central role of data and data-driven technologies such as AI was underscored by the UK government's mandate to NHS England (NHSE), specifically priority three mandate to deliver recovery through the use of data and technology.¹²

While such goals and mandates indicate the strategic direction alongside high-level principles and frameworks,¹³⁻¹⁵ there is limited information regarding specific practical issues to be addressed and pragmatic actions required from NHS organisations and industry partners to achieve these goals. The recent NHSE Transformation Directorate's publication on effective NHS data partnerships provides structure regarding aspects for consideration;¹⁵ however, there remains a paucity of practical guidance regarding real-world application. For example, several fundamental components are lacking in many NHS organisations that are required to meet the guidelines suggested for effective data partnerships, including availability of appropriately skilled staff with expertise in health-data engineering and analysis, resources for dedicated staff, appropriate NHS Information, Communication and Technology (ICT) infrastructures, such as secure, approved and flexible cloud platforms, appropriate governance structures to support innovation and collaboration, and suitable delivery structures within NHS organisations to develop, evaluate and maintain informatics or AI software tools for clinical or operational purposes.

We present learnings regarding the commercial and operational practicalities of an NHS trust (Great Ormond Street Hospital for Children, 'GOSH') working in partnership with industry (Roche Products Limited, 'Roche UK') towards the aims of the Life Sciences Industrial Strategy to deliver benefit through better use of data and digital tools. We provide candid, constructive reflections on the issues, barriers and potential mitigations, to moving from strategy to sustainable impact, which can be scalable across the NHS and beyond.

Partnership process

In December 2021, Roche UK and GOSH established a formal collaboration towards personalised healthcare by jointly addressing data engineering and analysis issues required to optimise use of routine anonymised health data for patient benefit. GOSH and Roche UK have been working collaboratively in partnership for 24 months, strategically focusing on areas including establishing a 'Clinical Intelligence Unit', improving clinical/operational decision support, enriching datasets with complex genomic and free-text derived data using Natural Language Processing (NLP), optimising real-world data for paediatric personalised healthcare, and improving insights through data from devices and wearables. The partnership was highlighted as an exemplar collaborative approach by NHSE.¹⁵

GOSH is an international paediatric centre of excellence, digitally mature¹⁶ and hosts the NHSE North Thames Genomics Laboratory Hub,¹⁷ one of seven in England.¹⁸ GOSH's Data, Research, Innovation and Virtual Environments Unit ('GOSH DRIVE') is a leading informatics centre,¹⁹ within which the partnership has developed a Clinical Intelligence Unit (CIU) to allow technologies such as AI to move through the software development lifecycle (SDLC) into real-world clinical practice/deployment, including validation and maintenance. Roche UK is a global pioneer in personalised healthcare, including use of data and technology, working towards improving patient care through collaboration with the NHS.

As identified by the Topol Review,²⁰ there is a need to attract particular skills into the NHS as data scientists, engineers and technical specialists to create technological solutions to improve care and productivity. Roche UK funding and seconded staff, under its collaboration with GOSH, catalysed this pipeline for benefit of the NHS and the wider Life Sciences ecosystem. The aim was to co-develop processes, products and tools using AI (such as NLP) to enable automated extraction of unstructured information (eg genomics, images, text) into clinical workflows, to augment operational and clinical decision-making at scale. The

partnership was supported by a project steering group made up of representatives of GOSH and Roche UK and is expected to last for 5 years. GOSH provides access to the GOSH DRIVE infrastructure for the purpose of the collaboration across several workstreams (Fig. 1).

Governance framework

To support cross-sector collaboration, the Association of the British Pharmaceutical Industry (ABPI) published 'routes' to cross-sector working (Fig. 2)²¹. Roche UK and GOSH specifically aimed to pool skills, resources and capabilities for the benefits of patients and the NHS, therefore a 'Collaborative Working Agreement' was used, a relatively new approach for NHS organisations, and aspects required establishment of appropriate trust governance. In the absence of an appropriate body at the time to negotiate commercial agreements (as recommended by the Life Sciences Industrial Strategy),¹ this challenge to GOSH was mitigated by support by an external legal provider (Pinsent Masons), who worked with Roche UK's legal team to navigate commercial arrangements, through the Collaborative Working Agreement.

Legal negotiations focused on practical solutions. It was agreed that Intellectual Property Rights (IPRs) should be solely owned by GOSH (save for those related only to Roche UK). This unblocked an area that often involves protracted negotiations in industry-sponsored collaborations and created an atmosphere of trust, emphasising that promoting benefit, rather than seeking competitive advantage, was the driver for collaboration. With IP ownership resolved, the legal teams focused on the parties' overall relationship, and means by which data could be used. These issues were governed by the Collaboration Agreement, which set out the aims using real-world anonymised data to improve paediatric healthcare, including roles and responsibilities (Roche UK providing data science and analytical expertise and funding, and GOSH DRIVE facilitating access to physical premises, digital infrastructure, clinical expertise and matched funding). The parties made an equal contribution to project governance, each contributing to the Project Steering Group. It was agreed that no patient data would be shared or transferred from GOSH to Roche UK. While GOSH had well-established processes for research governance, a specific challenge was the establishment of internal approval processes for use of deidentified data for innovation. Thus, a Data Partnerships Committee (DPC) was established to deal with issues relating to collaboration for innovation, and all activity was approved by the DPC.

A Data Protection Impact Assessment (DPIA) set out steps to ensure that data used through the collaboration are appropriate, safe and compliant with data protection legislation, namely the General Data Protection Regulation (GDPR). GOSH and Roche jointly developed the DPIA through an iterative process, resulting in a DPIA that was legally compliant, practical and jointly implementable since the purposes and procedures were agreed together. As well as these specific legal requirements, our partnership approach included further ethical consideration and data privacy with the parties entering a Data Agreement. Providing further safeguards if, in exceptional circumstances, there was any requirement for data to be shared (such as nature and/or rarity of a condition, or datasets comprising small numbers), the data and purposes for which it would be used, would be determined by the Project Steering Group and Trust Data Partnership Committee in accordance with the DPIA, which includes establishing and complying with the legal basis.

Roche UK staff working on GOSH premises through the partnership were issued with GOSH honorary contracts, which provided additional governance, as staff were accountable in accordance with GOSH policies, procedures and clinical governance frameworks. This also benefited Roche UK, since such staff could contribute to activities within an approved secure NHS environment under the supervision of GOSH employees. A service level agreement was used to manage the arrangements, clarify the terms of the contracts and the status of Roche UK staff members.

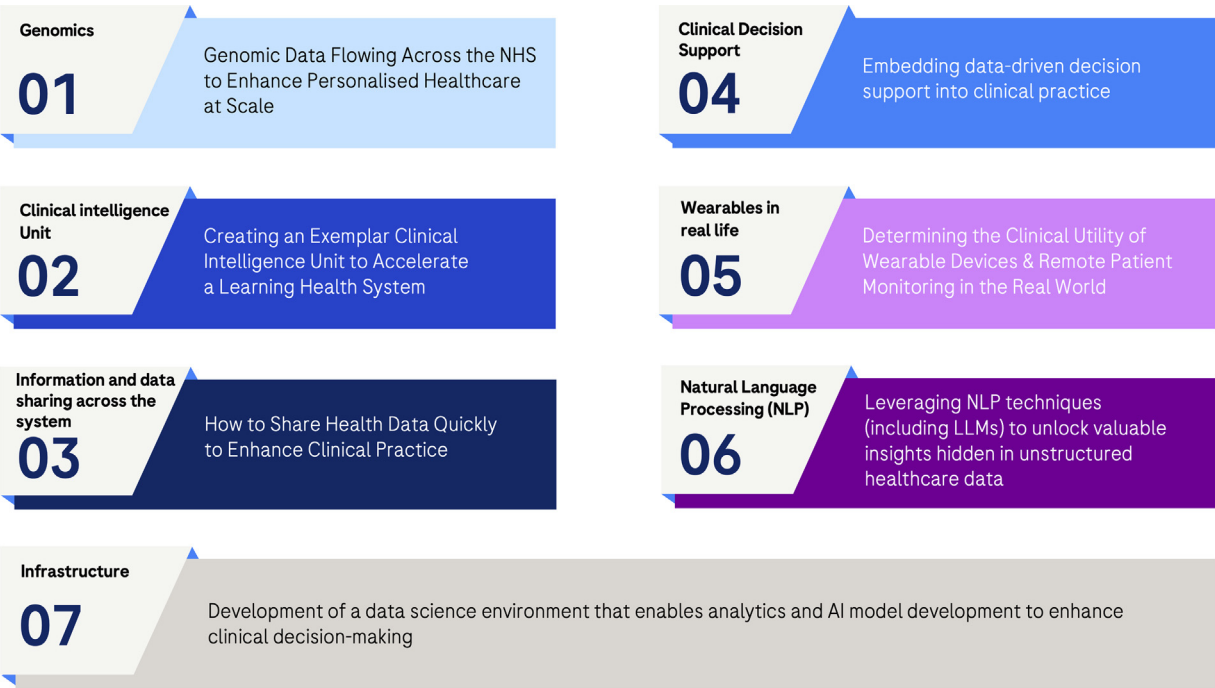


Fig. 1. Schematic illustrating main activity workstreams within the NHS-Industry partnership.

Routes to Cross-Sector Working
Choosing the most appropriate model

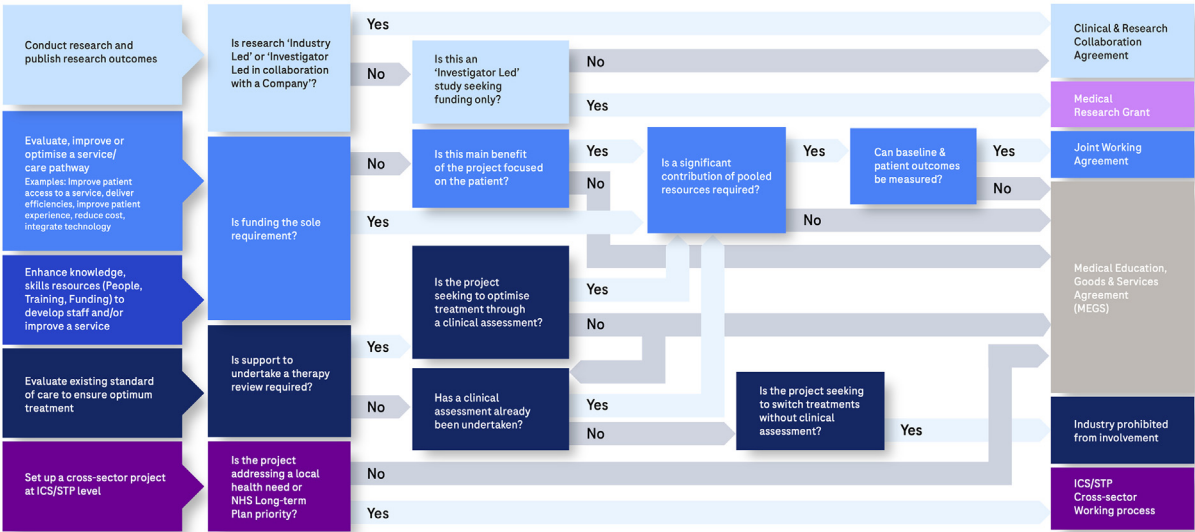


Fig. 2. Key aspects of the ABPI Code of Conduct guidance on routes to cross sector working relevant to establishment of the NHS-Industry partnership. (Modified from ABPI guidance²¹).

Learnings

The partnership aims to foster collaboration between Roche UK and GOSH’s legal teams and no barriers arose. However, we foresee potential barriers more generally around NHS access to appropriate legal advisers, both monetary and in the expertise needed when dealing with complex data issues. GOSH, in negotiating directly with Roche UK, incurred sunk legal costs to enter the partnership. This may hamper NHS organisations from contracting with industry partners, creating barriers to uptake and co-creation of technology across the NHS.

Recommendations

Currently, the NHS uses collective bargaining power for procurement decisions and we recommend that a similar national-level, single appraisal process is established for data partnership and collaboration agreements with industry, or, as a minimum, guidance and templates, similar to the Lambert Toolkit for collaborative research projects between academia and industry, or model clinical trial and related agreements available via the Integrated Research Application System (IRAS).

We support use of anonymised data (defined according to ICO guidance) for industry–NHS partnerships wherever possible, as well as a secure data environment meeting appropriate security standards and data sharing agreements to further safeguard data privacy. In joint working to date, such data have been sufficient. We advocate establishment of a formal process that includes appropriate senior information governance staff responsible for compliance with data regulations. We suggest that there is little benefit in industry partners looking to commercialise intellectual property when working with the NHS, but rather that code/tooling developed should be shared under open licences for reuse across the sector. This bolsters public trust and ensures that use of NHS data is fair and equitable while supporting open coding in the NHS.²²

Infrastructure

In 2019, it was estimated that curated NHS data could generate £9.6 billion per annum of value.²³ There has been £2.1 billion investment for NHS IT upgrades,²⁴ £260 million for Secure Data Environments (SDEs), £60 million ring-fenced to expand the manufacturing arm of the life sciences sector²⁵ and recently awarded £480 million for an NHS Federated Data Platform,²⁶ to support the Life Sciences Vision.¹ The proposed NHS SDEs are intended to support NHS data use for research and analysis. However, such approaches are inadequate for developing and implementing analytical tools and algorithms into clinical practice, specifically through ideation, planning, design, implementation, testing, deployment and maintenance. However, NHS organisations have limited in-house infrastructure to support such ambitions, making it difficult to move digital interventions to scale.²⁷ NHS organisations will require appropriate clinical-grade SDEs and cloud infrastructures to support development and deployment pipelines outside of proprietary EHR-associated tools.

We identified several learnings and specific challenges regarding infrastructure limitations using a real-world use-case of NLP to extract structured information from genomic reports, which are almost universally generated and stored as document files, usually PDF format, which allows ease of use for clinicians managing an individual patient, but does not support secondary use and analysis or development of clinical decision support tools.

Learnings

Although GOSH already had a cloud-based Digital Research Environment (DRE/SDE), this was primarily designed to support secure provisioning and analysis of structured data for research and operational purposes rather than tooling across the software development life-cycle. Therefore, GOSH data scientists were initially required to run pre-trained NLP models,^{28,29} on NHS-provided laptops, which required extended run times (>2 days or crashing) and inability to fine-tune models due to lack of compute. This specific challenge was further amplified by the well-documented obstacle that NHS IT policies actively prohibit installation of otherwise standard software tools required for this type of work (eg installing Docker or PaaS applications).²² An interim solution was provision of a higher-specification desktop than normal NHS use and set-up of an on-premises Linux Sandbox to run Docker/PaaS for containerised NLP models, consisting of a staging server for processing potentially identifiable information, with no internet access, and a development server with restricted network access. However, this approach is not sustainable or scalable and precludes the NHS incorporating industry best practices of analytic, data science and software development communities. For example, this led to a specific challenge in the development of a version-controlled NLP codebase on GitLab, as GOSH's DevOps platform instance was several versions behind the current stable version, since this application was not deemed part of the role of NHS ICT. This specific challenge culminated in the manual loading

of genomic reports to a local desktop folder and outputs stored locally for interpretation. Ideally, code in GitLab should connect to a server, triggered to run when a new batch of reports is automatically loaded into a source directory or run at specific intervals and the output written to an agreed output directory that enables other applications and services to use the data. Reproducible Analytical Pipelines (RAPs) are the gold standard that partnerships should be working towards, but in addition to NHS infrastructure issues, there is no clear NHS workforce responsible and accountable for deployment, maintenance and governance of such tools, processes and infrastructure. Specific technology roles are needed that liaise with existing NHS ICT teams, but also understand NHS regulations and approvals needed to deliver the vision supported by cloud computing. We propose this as another important, often-overlooked dimension towards 'bridging the chasm between AI and clinical implementation'.³⁰

On-premises vs cloud computing

Cloud computing enables ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (eg networks, servers, storage, applications and services) that can be rapidly provisioned and released with minimal management effort, benefits recognised strategically by the NHS.³¹ NHS Digital stated that cloud services are a safe location to store patient information, provided that certain data sovereignty conditions are met, with further endorsement via the National Information Board's Personalised Health and Care Framework (2020).³² However, discussions regarding implementation of cloud computing that enable RAPs within GOSH was met with concerns, which may be common across NHS ICT Departments. To mitigate this, there is a need to decouple analytic team RAP requirements from clinical services. Importantly, NHS ICT teams have limited resources and cloud expertise, making implementation difficult. There may also be concern regarding cost since most NHS trust budgets are driven by capital expenditure rather than the revenue model commonplace for industry. More broadly, questions and processes on mitigating and assessing risks associated with cloud services (eg cybersecurity) need to be worked through with the SRIO and relevant stakeholders.

While we are currently using on-premise MLOps environment with functionalities such as MLFlow Tracking server for tracking and monitoring ML models, and enabling container registry within GitLab for continuous integration and development, the partnership is continuing in-process establishment of suitable cloud infrastructure, which continues to have significant barriers and challenges to full implementation. However, even this staged approach requires additional resource from the NHS ICT team that is challenging to secure with competing demands. Part of the new approach involves upskilling NHS ICT and DRIVE teams in cloud infrastructure, including continuous integration and development and DevOps skills for sustainable production-level support, as the partnership matures.

Recommendations

Initially, both parties underestimated the amount of organisational challenges that would be encountered within an NHS organisation to implement modern, open working and computational methods in a hospital; in addition to cultural change and upskilling required to support infrastructure and AI builds at scale, these include NHS ICT resource availability, skill sets and technical infrastructure issues. It was necessary to reallocate funding to support a previously unplanned technical ICT lead to facilitate these changes and work with NHS ICT stakeholders, which continues to prove essential since requirements are quite different from NHS ICT business as usual. Despite cloud guidance frameworks,³¹ we recommend significant additional support and resourcing for NHS ICT teams to implement the government's 'Cloud First' policy.³³

Table 1
Summary of major issues identified for NHS-Industry collaboration and suggested mitigations.

Issue	Challenge	Recommendation
Industry requirements and codes of practice	Need for ABPI and regulatory requirements, which are not standard practice in NHS organisations	- Education for NHS partner regarding requirements - Joint material tracker - Steering committee oversight
Internal approval processes for use of deidentified data for innovation	Established processes for research governance but NHS less mature regarding data for innovation	- Establishment of specific data partnership committee for all governance and issues relating to use of data for innovation, including with commercial partners
Legal contractual arrangements for industry partnerships	NHS organisations less able to access specialised legal advice both in terms of monetary investment and level of sophistication needed	- Establish central NHS data partnership/collaboration agreements with industry, and/or guidance and templates for such
Oversight and governance of industry partners working in NHS organisations	Collaborative working requires ‘one team’ approach, hence all team members bound by similar governance processes	- Ensure that industry partner staff have honorary NHS contracts with all associated training and onboarding
Trust and security for secondary use of data	Ensure data security and governance for collaborative working	- Use anonymised data (defined according to ICO guidance) for industry–NHS partnerships as default
Intellectual property (IP) generation	Ensure fair use for all parties and wider NHS for any generated IP	- Code and tooling developed from NHS data should be shared under open licences for review and reuse across the sector
NHS data availability for secondary analysis	Large volumes of routine data may be available, but much is as unstructured or proprietary data formats	- Significant technical and NHS organisation-specific resource is required to generate high-quality, analysis-ready healthcare data for further use, including expertise in areas such as natural language processing
NHS infrastructure for tool development	Few NHS organisations possess suitable cloud-enabled secure environments and ICT skills to support tool development and deployment	- Significant effort and resources are required to establish appropriate NHS ICT infrastructure and skills to support collaborative working using modern methods
Restrictive NHS ICT practices	NHS organisations maximise security and minimise system risks, often resulting in barriers to use of otherwise industry standard tools (eg Gitlab, docker, etc)	- Early engagement with NHS ICT to understand the need for modern coding practices and associated tools - Industry partnerships require dedicated ICT support functions in addition to standard NHS ICT staffing
Clinical engagement	Ongoing clinical engagement from healthcare professionals required, but few have dedicated time for innovation in job plans	- CRIO role and associated innovation officers representing all hospital directorates required to ensure engagement
Project management	Overarching ‘one team’ project management required, but existing styles and methods likely differ between industry and NHS teams	- Dedicated project management tooling used by both parties required - Compromise from NHS and industry to achieve mutually acceptable project management approach - Significant human factor focus required to establish trust and embrace different ways of working

Clinical engagement for innovation

Clinical engagement is essential for innovation in healthcare. Prior to the partnership, GOSH had a clinical chief research information officer (CRIO) and clinical director of innovation, who were involved with establishing the collaboration. However, projects require additional engagement from clinical teams.

Learnings

A notable challenge included that, while many staff were enthusiastic to participate, a practical barrier is availability of clinical staff time, since few NHS staff have ‘innovation’ within their job plans. In addition, collaborative working with industry is uncommon outside clinical trials and there was initially no single process for engaging staff regard-

ing innovation and managing the relationship with data and partnership teams. Therefore, we established a deputy CRIO and clinical innovation officers from all directorates, whose remit was liaison between GOSH DRIVE and clinical services, to identify and prioritise projects and provide practical input through dedicated time. In addition, a central innovation project management process was established to capture, triage and plan actions, linked to a wider engagement programme across the organisation.

Evaluation metrics for success of this process include the number of innovation projects across trust directorates submitted to this partnership and the degree to which such projects drive improved patient experience, productivity or cost savings. For example, one project identified potential overuse of a specific laboratory test with significant potential cost savings (>£100K) and improved patient experience due to reduced sampling. However, reasons for clinical testing are complex,

and behavioural change by clinicians requires additional analysis and intervention, such as through the EPR at point of ordering

Recommendations

Dedicated clinical staff time is required to work with technology teams and for wider organisational engagement. A CRIO and associated innovation officer roles representing all directorates with dedicated funded time are required. An overarching project management structure is needed encompassing the entire process, including clinical engagement linked to business results in the form of productivity, efficiency and patient benefit gains through a business plan.

Project management and ways of working

Collaboration requires NHS and industry employees to work together to deliver the vision. Creating this 'team interoperability', where people have different but complementary ways of working, highlighted several challenges.

Learnings

Some existing NHS staff were uncertain regarding coworking with industry colleagues in an NHS organisation, and the partnership was initially viewed as additive workload, rather than complementary to existing hospital goals. Mitigating this required the CRIO and innovation director to emphasise that collaboration was aligned with existing strategic goals and wider NHS to use health data for patient benefit, as well as adding clarity on roles and responsibilities. Similarly, industry staff were familiar with agile working, but this is challenging when staff do not share infrastructures and resources. Project management tooling was required since agile working software was not routinely used in the NHS organisation.

Outside clinical trials, it is not common for NHS organisations to collaborate with pharmaceutical companies, and establishing common ways of working as 'one team' with a culture of trust and psychological safety was challenging. For example, it took time to 'speak the same language' ('sprints', 'backlog') and for the team to be comfortable discussing blockers. We experimented with meeting cadences and settled on weekly workstream stand-ups and monthly review meetings as a compromise between full agile working and competing NHS demands. In addition, there were specific meetings for key project milestones, with impact retrospectives and planning discussions.

Recommendations

Dedicated project managers are required to manage the partnership, including new ways of working in the NHS, aided by appropriate project management tooling. We recommend regular reviews and agile delivery,³⁴ with recognition that NHS resourcing may mean that staff have competing demands that compromise their ability to support dedicated sprints. It is important to invest time to agree expectations and processes, such as agile practices and deliverables, with agreed escalation routes for blocked activities. We advocate the use of project management software that has specific functionalities, like assigning tasks within cards that contain the deadline for completion and the person accountable for delivery and highlights interdependencies between workstreams.

Conclusion

We present practical experience with real-world NHS-industry collaboration focused on using data and technology to improve clinical care and research, highlight issues that parties should be aware of and make recommendations for future collaborations (Table 1). We highlight the benefits of such an approach, which has resulted in outputs

including partnership creation with agreed trust and industry partner processes, commercial and intellectual property agreements, staff contractual arrangements and workstreams established; trust governance structures created to ensure compliance; patient engagement through communication with governors and advisory groups; staff engagement, including innovation officers; development of NLP pipelines not previously possible, including to extract structured information from genomic reports taking a fraction of the time required to extract manually; developed schema and implementation document for structured genomic data in Fast Healthcare Interoperability Resources (FHIR)-compliant format (shared with NHSE regarding national genomic data standards); developed a scalable tool for rapid and reproducible research using real-world data; established NHS clinical intelligence unit (CIU) to improve hospital operational efficiency.

Given that the vast majority of clinical information is recorded as unstructured data through free-text clinical notes, discharge summaries, histopathology, genomic, radiology reports and referral letters, next steps of the partnership are to build scalable information retrieval technologies that enable the ability to filter and query large quantities of both structured and unstructured clinical data, with the ultimate intention to provide clinicians with an 'Informatics Consult' tool³⁵ to improve research, operational efficiencies and clinical decision making, at scale.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: GOSH organisation reports financial support was provided by Roche. GOSH organisation reports a relationship with Roche that includes: funding grants. Article coauthored by GOSH and Roche together as part of a 5-year collaboration agreement between NHS and industry as part of UK Life Sciences Industry Strategy reporting our joint learnings for others. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

CRedit authorship contribution statement

Rebecca Pope: Writing – review & editing, Writing – original draft, Supervision, Resources, Project administration, Conceptualization. **Alexandros Zenonos:** Writing – review & editing, Writing – original draft, Supervision, Project administration. **William Bryant:** Writing – review & editing, Writing – original draft, Supervision, Resources, Project administration. **Anastasia Spiridou:** Writing – review & editing, Writing – original draft, Supervision, Project administration. **Daniel Key:** Writing – review & editing, Writing – original draft, Supervision. **Shiren Patel:** Writing – review & editing, Writing – original draft, Resources, Project administration. **Jack Robinson:** Writing – review & editing, Writing – original draft, Resources. **Anna Styles:** Writing – review & editing, Writing – original draft. **Chris Rockenbach:** Writing – review & editing, Writing – original draft, Project administration. **Gina Bicknell:** Writing – review & editing, Writing – original draft. **Pavithra Rajendran:** Writing – review & editing, Writing – original draft. **Andrew M. Taylor:** Writing – review & editing, Writing – original draft, Supervision, Resources, Conceptualization. **Neil J. Sebire:** Writing – review & editing, Writing – original draft, Supervision, Resources, Conceptualization.

References

- Life sciences: industrial strategy - GOV.UK. <https://www.gov.uk/government/publications/life-sciences-industrial-strategy> (accessed Dec 29, 2017).
- Better, broader, safer: using health data for research and analysis - GOV.UK. <https://www.gov.uk/government/publications/better-broader-safer-using-health-data-for-research-and-analysis> (accessed Sept 16, 2022).
- Data saves lives: reshaping health and social care with data. GOV.UK. 2022; published online June 15. <https://www.gov.uk/government/publications/data-saves-lives-reshaping-health-and-social-care-with-data> (accessed Oct 24, 2023).

4. Transforming for a digital future: 2022 to 2025 roadmap for digital and data - updated September 2023. GOV.UK. <https://www.gov.uk/government/publications/roadmap-for-digital-and-data-2022-to-2025/transforming-for-a-digital-future-2022-to-2025-roadmap-for-digital-and-data> (accessed Oct 24, 2023).
5. Reproducible analytical pipelines – GSS. <https://gss.civilservice.gov.uk/reproducible-analytical-pipelines/> (accessed Sept 16, 2022).
6. Data: a new direction. GOV.UK. 2021; published online Sept 10. <https://www.gov.uk/government/consultations/data-a-new-direction> (accessed Oct 24, 2023).
7. Consultation on the future regulation of medical devices in the United Kingdom. GOV.UK. 2021; published online Sept 16. <https://www.gov.uk/government/consultations/consultation-on-the-future-regulation-of-medical-devices-in-the-united-kingdom> (accessed Oct 24, 2023).
8. A plan for digital health and social care. GOV.UK. <https://www.gov.uk/government/publications/a-plan-for-digital-health-and-social-care/a-plan-for-digital-health-and-social-care> (accessed Oct 24, 2023).
9. UK Digital Strategy. GOV.UK. <https://www.gov.uk/government/publications/uks-digital-strategy/uk-digital-strategy> (accessed Oct 24, 2023).
10. Data protection and digital information bill - parliamentary bills - UK Parliament. <https://bills.parliament.uk/bills/3322> (accessed Oct 24, 2023).
11. Bennett insights: an overview of UK Data Policy Developments | Bennett Institute for Applied Data Science. 2022; published online July 27. <https://www.bennett.ox.ac.uk/blog/2022/07/bennett-insights-an-overview-of-uk-data-policy-developments/> (accessed Oct 24, 2023).
12. The government's 2023 mandate to NHS England. GOV.UK. <https://www.gov.uk/government/publications/nhs-mandate-2023/the-governments-2023-mandate-to-nhs-england> (accessed Oct 24, 2023).
13. Value sharing framework for NHS data partnerships. NHS Transformation Directorate. <https://transform.england.nhs.uk/key-tools-and-info/centre-improving-data-collaboration/value-sharing-framework-for-nhs-data-partnerships/> (accessed Oct 24, 2023).
14. NHS data: maximising its impact for all. Health Management and Policy Alert. https://kingsfund.blogs.com/health_management/2023/04/nhs-data-maximising-its-impact-for-all.html (accessed Oct 24, 2023).
15. A guide to effective NHS data partnerships. NHS Transformation Directorate. <https://transform.england.nhs.uk/key-tools-and-info/centre-improving-data-collaboration/guide-to-effective-nhs-data-partnerships/> (accessed Oct 24, 2023).
16. GOSH becomes first UK hospital to achieve HIMSS stage 7. *Digital Health*. 2020. published online Sept 17. <https://www.digitalhealth.net/2020/09/gosh-himss-stage-7/>. (accessed Oct 24, 2023).
17. We are the North Thames Genomic Laboratory Hub. <https://www.norththamesglh.nhs.uk/> (accessed Oct 24, 2023).
18. NHS England » Genomic Laboratory Hubs. <https://www.england.nhs.uk/genomics/genomic-laboratory-hubs/> (accessed Oct 24, 2023).
19. Great Ormond Street Hospital Digital Research Innovation and Virtual Environment. GOSH DRIVE. <https://www.goshdrive.com> (accessed Oct 24, 2023).
20. Topol review | Health Education England. <https://www.hee.nhs.uk/our-work/topol-review> (accessed March 25, 2019).
21. Routes to cross-sector working. <https://www.abpi.org.uk/partnerships/working-with-the-nhs/joint-working-a-toolkit-for-industry-and-the-nhs/routes-to-cross-sector-working/> (accessed Oct 27, 2023).
22. Goldacre review. <https://www.goldacrerreview.org> (accessed Dec 5, 2023).
23. ey-value-of-health-care-data-v20-final.pdf. https://assets.ey.com/content/dam/ey-sites/ey-com/en_gl/topics/life-sciences/life-sciences-pdfs/ey-value-of-health-care-data-v20-final.pdf (accessed Dec 5, 2023).
24. Sunak confirms £2.1bn for NHS technology for more 'connected' hospitals. *Digital Health*. 2021. published online Oct 27. <https://www.digitalhealth.net/2021/10/sunak-confirms-2-1bn-for-nhs-technology-for-more-connected-hospitals/>. (accessed Dec 5, 2023).
25. £260 million to boost healthcare research and manufacturing. GOV.UK. <https://www.gov.uk/government/news/260-million-to-boost-healthcare-research-and-manufacturing> (accessed Dec 5, 2023).
26. NHS Federated Data Platform - contracts finder. <https://www.contractsfinder.service.gov.uk/notice/b44330d4-69a4-4b81-ab77-174c1dd8a27d?origin=SearchResults&p=1> (accessed Dec 5, 2023).
27. NHS Data - Maximising impact for All.pdf. <https://spiral.imperial.ac.uk/bitstream/10044/1/103404/8/NHS%20Data%20-%20Maximising%20Impact%20for%20All.pdf> (accessed Oct 24, 2023).
28. Lee J, Yoon W, Kim S, et al. BioBERT: a pre-trained biomedical language representation model for biomedical text mining. *Bioinformatics*. 2020;36:1234–1240.
29. Weber L, Sanger M, Münchmeyer J, Habibi M, Leser U, Akbik A. HunFlair: an easy-to-use tool for state-of-the-art biomedical named entity recognition. 2020; published online Aug 18. DOI:10.48550/arXiv.2008.07347.
30. Aristidou A, Jena R, Topol EJ. Bridging the chasm between AI and clinical implementation. *Digital Med*. 2022;399:620.
31. NHS and social care data: off-shoring and the use of public cloud services Guidance. NHS Digital. <https://digital.nhs.uk/data-and-information/looking-after-information/data-security-and-information-governance/nhs-and-social-care-data-off-shoring-and-the-use-of-public-cloud-services/guidance> (accessed Dec 6, 2023).
32. Cloud security – good practice guide. NHS Digital. <https://digital.nhs.uk/services/cloud-centre-of-excellence/cloud-security-good-practice-guide> (accessed Dec 6, 2023).
33. Government Cloud First policy. GOV.UK. 2023; published online June 19. <https://www.gov.uk/guidance/government-cloud-first-policy> (accessed Dec 6, 2023).
34. Agile delivery - service manual - GOV.UK. <https://www.gov.uk/service-manual/agile-delivery> (accessed Dec 6, 2023).
35. Lai A, Chang WH, Parisinos CA, et al. An informatics consult approach for generating clinical evidence for treatment decisions. *BMC Med Informatics*. 2021;21:281.