Ford, Elizabeth, Stanley, Katryn, Rees-Roberts, Melanie, Giles, Sarah, Goddard, Katie and Armes, Jo (2022) *Understanding Public Priorities and Perceptions of the Use of Linked Healthcare Data in South East England.* In: Challenges of Trustable AI and Added-Value on Health. Studies in Health Technology and Informatics . IOS Press, pp. 594-598. ISBN 978-1-64368-284-6.

Understanding Public Priorities and Perceptions of the Use of Linked Healthcare Data in South East England

Elizabeth FORD [a,1], Kathryn STANLEY [a] Melanie REES-ROBERTS [b], Sarah GILES [c], Katie GODDARD [a], Jo ARMES [d]

[a] *Brighton and Sussex Medical School, Brighton, UK*
[b] *University of Kent, Canterbury, UK*
[c] *Public Advisor, Kent, UK*
[d] *University of Surrey, Guildford, UK*

Abstract. The counties of Kent, Surrey and Sussex (KSS) in South East England are creating anonymized, linked databases of healthcare records for audit, service planning and research for the first time. We consulted with 79 citizens from KSS in 5 deliberative focus groups, asking about perceived benefits and concerns regarding these new data assets. Participants hoped the linked datasets could be used for joining up care and information, improving efficiency, and improving healthcare provision, but were concerned about missing and inaccurate data, data breaches and hacking, use of data by profit-making organisations, and stigma and discrimination. Findings will be used to underpin governance and engagement strategies for integrated datasets in KSS.

Keywords. Linked data, patient data, public engagement, data governance, privacy

1. Introduction

In England, the government has given national funding to each region to set up digital transformation programmes in the National Health Service (NHS). Each area, generally at a county level, now has an integrated care system (ICS), bringing together previously separated healthcare providers under one umbrella organization [1]. The transformation includes plans to create an extracted, anonymized and linked patient record database for audit, service planning, commissioning and research. This is the first time that the counties of Kent, Sussex, and Surrey (KSS) in South East England have had such health data infrastructure. The NIHR-funded "Unlocking Data to Inform Public Health Policy and Practice" project (NIHR133761) has brought together universities, local government, and NHS to use these emerging datasets for public health planning and policy.

One of the key issues to consider when linking and curating data from patients' health records is how the public view this activity. We saw an opportunity to engage with citizens of Kent, Surrey and Sussex to create data governance structures from the start which were inclusive, transparent, and which have citizen support. There was very little

---

[1] Corresponding Author, Dr E Ford, Department of Primary Care and Public Health, Brighton and Sussex Medical School, Watson Building, Village Way, Falmer, BN1 9RH e.m.ford@bsms.ac.uk.

local evidence of public awareness, or "reasonable expectations" of patients and public around population data linkage for analytics. A publicly-supported and informed data governance and access structure in KSS would ensure uses of patient data are appropriate, ethical, and can be prioritized against public benefit and need. According to social licence theory, the public expect that organisations who, for example, hold and use the public's data, will go beyond the requirements of formal regulation, and adhere to voluntary codes of trustworthy behaviour and transparency [2]. Where the public are satisfied that the motivations and competency of the organisation are trustworthy, they may confer a "social licence" to operate. Additionally, the public consider core research ethics principles when weighing up approval of data-sharing and linkage schemes [3-5].

In this study we aimed to engage with citizens of Kent, Sussex and Surrey, to understand views around linked health dataset use, particularly, what benefits they see will accrue for the public from these assets and what concerns they have about their use.

2. Methods

2.1. Study Design

This project was approved by Brighton and Sussex Medical School Research Governance & Ethics Committee (ref: ER/BSMS2730/7). We conducted deliberative discussion focus groups according to the methodology of Rothwell et al [6]. Deliberative research is an approach for gathering wider views about health issues when there are many complex issues to weigh against each other [4]. Deliberative discussion focus groups include a range of informative presentations, interspersed with facilitated plenary and breakout discussions.

2.2. Recruitment and Procedure

We advertised our study through a large range of health, social care and support organisations in Kent, Surrey and Sussex via an online flyer, giving a link to a website at which potential participants could express an interest. This was followed by sending participants a full information sheet and consent form. Once consented, participants were booked into a focus group and were sent further online reading about issues to be discussed. Participants were sent a £50 voucher after completing the focus group.

Focus groups were conducted remotely using Zoom video-conferencing software and were recorded. Each group consisted of two sessions one week apart totalling 3 hours; at the beginning, participants were informed both benefits and risks would be explored. All participants completed an online pre-group demographics questionnaire and post-group feedback questionnaire; both included a question on willingness to share their medical records to ensure a diverse range of views were represented.

2.3. Focus group content and discussion questions.

Session 1: What integrated datasets of health and social care data could be used for in public health (with case studies on joining up care for diabetes and multi-morbidity).
Q1: What could be the benefits of using these data for research and commissioning for the public?

Q2: What can be achieved with these datasets that otherwise would not happen?

Q3: What are the most important types of projects that should be done with these data? What do you feel is less important or should not be done?

Session 2: Governance and privacy issues, concerns and solutions for linked data.

Q1: Do you have any concerns or suggestions around the current framework for protecting health and social care data in integrated datasets?

Q2: What are your views on data being linked together, particularly health and non-health data (e.g. housing benefits, crime reporting, or educational attainment data)? Do you have any particular worries about this?

Q3: Who should have access to these datasets for analysis, and who should we be cautious about giving access to? What vetting procedure should they go through?

2.4. Data Analysis

Questionnaire responses were analysed descriptively. Zoom discussion groups were audio-recorded and transcribed verbatim. Written transcripts were uploaded to NVivo (version 1.4.1) for thematic analysis according to the 6-steps of Braun and Clarke [7]. Firstly, transcripts were read repeatedly and initial ideas noted down, secondly interesting features of the data were coded. Codes were then collated into potential themes and a thematic map was produced. Then authors reviewed the specifics of each theme against the study aims, the overall story, and the thematic map, and selected extracts which best represented each theme.

3. Results

Of 152 people who expressed an interest to take part, 79 completed one of 5 focus groups.

Table 1. Demographics of the sample (ND: Not disclosed)

Characteristic	N (Total = 79)
Gender	Female: 52 Male 26 Non-Binary: 1
Ethnicity	White 63, Asian 10, Black 5, Mixed 1
Educational Attainment	School leaver 23; University education 44; ND: 12
County of Residence	Kent 11, Sussex 49 Surrey 19
Disability Status	Yes: 17 No: 47; ND: 15
Willingness to share medical data (pre)	Willing 62 Unwilling 15 Don't know 2
Willingness to share medical data (post)	Willing 65 Unwilling 7 Don't know 7

The key themes expressed by participants were around (1) the potential **<u>benefits</u>** of using integrated data, and (2) the potential **<u>risks or concerns</u>**.

3.1. Benefits

The benefits expected from linking data for service planning and research included **joining-up care and information, improving efficiency, and improving healthcare provision.**

Many participants referred to prior experiences of fragmented care provision which focused on single issues or where different parts of the health system were not aware of what other providers were doing. Participants hoped that the new linked datasets would help planners to commission more joined-up services.

"All this dataset analysis and helping to link them together will enable each department to do their job better. All of them are wasting so much time in doing the same thing again or doing the wrong thing because they haven't got the information. This will make everything so much more efficient for everybody."

Participants also hoped that better information flows would reduce the need to repeat their stories to multiple clinicians, wasting time and effort, and risking key clinical information being lost or missed:

"The efficiency for the NHS and all the other players will be in the non-duplication of all these visits ...I explained a problem I have to my GP; I then go along to the hospital and then I explain it to a nurse, I then go and see a consultant and rather strangely he didn't know anything about it, 'can you tell me what your problem is?'"

Participants also hoped that the system overview gained by analysing linked health service data would lead to more efficient use of resources, especially targeting care to the appropriate places or the appropriate people:

"If they're wasting less because they're able to stop duplicating things and stuff like that, then everyone's going to get better care overall because the funding won't be used inappropriately and wasted."

Lastly, participants hoped that the joining-up and efficient use of resources would result in improved healthcare, and improved health in KSS, including mental health.

"I think with physical disabilities, with other areas, ...sometimes the doctor will say… we need evidence-based therapies. The only way we're going to have this evidence base is if we start linking data and collecting data in an appropriate way so it can be properly analysed and then used… The outcomes will be so, so much better than what we have today."

3.2. Perceived risks and concerns

Patients noted a number of concerns about the uses of linked datasets, which centred on **missing and inaccurate data, data breaches and hacking, use of data by profit-making organisations, stigma and discrimination**. Participants were concerned that the data held in linked healthcare records would not be "clean" or accurate, and therefore, any planning or research conducted using the data, would also not be reliable.

"How do we ensure the cleanliness of the data and make sure that the data is accurate enough to build the right picture?"

Furthermore, participants were worried that the data could be hacked, or that anonymization was not a secure way to protect privacy. They were additionally concerned with who might get access to the data and suggested users of data should leave an audit trail *"a bit like you swipe a key card in a building."*

"I don't trust personally that all of this sharing of information isn't going to get hacked. I don't trust the fact that it's all going to be anonymised, there's human error that's going to come in here."

Part of these concerns focused on the issue that participants did not want their data to be used for a profit motive. They felt that research paid for by drug or insurance companies would be less transparent, and less in the public interest.

"Researchers are funded by people, organisations. Who they are funded by and what is their ultimate goal? For example, drug companies fund a huge amount of research but they will not publish those research results if it's not advantageous to them."

Lastly, participants were worried about stigma; both in terms of feeling uncomfortable with data on sensitive issues being shared for planning and research, and also worrying about stigma or discrimination which could arise from misuse of the data.

> *"I think people may feel there's some stigmas to things like [welfare] benefits and I think the stigmas and mental issues, and also sexual issues, is why the health clinics are always stringent about being confidential because people find these issues really difficult."*

4. Discussion and Conclusion

We identified core issues of both support and concern for the linking of anonymised NHS patient records in South East England to facilitate healthcare audit, planning and research in the region. These findings are similar to concerns expressed by patients in previous literature [3,4,8], although previous research has not reported on the public identifying increased efficiency of health services as a benefit of data-linkage. Concerns about poor data quality were also not prominent themes in prior research. We identified that citizens of KSS are keen to see data used to improve healthcare delivery, efficiency and connection of services, but also seek reassurance that data is appropriately safeguarded, and will only be used for the public good. We note limitations of our study: starting with presenting on and discussing the benefits of the datasets may have primed the participants to feel more positive about uses of data, although we did inform participants at the start that we would discuss both benefits and risks. Our findings may not generalize outside KSS, although congruence with national research is largely good.

Findings will enable the ICSs in KSS to create publicly supported governance, engagement and communication strategies, securing a social license for the use of linked health data in the region.

References

[1]   The King's Fund. Integrated care systems explained: making sense of systems, places and neighbourhoods. 2021 Available from: https://www.kingsfund.org.uk/publications/integrated-care-systems-explained.

[2]   Gunningham N, Kagan RA, Thornton D. Social License and Environmental Protection: Why Businesses Go Beyond Compliance. Law & Social Inquiry. 2004;29(2):307-41.

[3]   Stockdale J, Cassell J, Ford E. "Giving something back": A systematic review and ethical enquiry into public views on the use of patient data for research in the United Kingdom and the Republic of Ireland. Wellcome Open Res. 2019;3:6-.

[4]   Ford E, Oswald M, Hassan L, Bozentko K, Nenadic G, Cassell J. Should free-text data in electronic medical records be shared for research? A citizens' jury study in the UK. Journal of Medical Ethics. 2020;46(6):367-77.

[5]   Ford E, Boyd A, Bowles JKF, Havard A, Aldridge RW, Curcin V, et al. Our data, our society, our health: A vision for inclusive and transparent health data science in the United Kingdom and beyond. Learning Health Systems. 2019;3(3):e10191.

[6]   Rothwell E, Anderson R, Botkin JR. Deliberative Discussion Focus Groups. Qual Health Res. 2016;26(6):734-40.

[7]   Braun V & Clarke V. Using thematic analysis in psychology. Qualitative Research in Psychology, 2006;3(2), 77-101.

[8]   Aitken M, de St Jorre J, Pagliari C, Jepson R, Cunningham-Burley S. Public responses to the sharing and linkage of health data for research purposes: a systematic review and thematic synthesis of qualitative studies. BMC medical ethics. 2016 Dec;17(1):1-24.