



OPEN ACCESS

# Patients' and public views and attitudes towards the sharing of health data for research: a narrative review of the empirical evidence

Shona Kalkman ,<sup>1</sup> Johannes van Delden,<sup>1</sup> Amitava Banerjee,<sup>2</sup> Benoît Tyl,<sup>3</sup> Menno Mostert,<sup>4</sup> Ghislaine van Thiel<sup>1</sup>

► Additional material is published online only. To view please visit the journal online (<http://dx.doi.org/10.1136/medethics-2019-105651>).

<sup>1</sup>Department of Medical Humanities, Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht University, Utrecht, The Netherlands

<sup>2</sup>Farr Institute of Health Informatics Research, University College London, London, United Kingdom

<sup>3</sup>Cardiovascular Center for Therapeutic Innovation, Institut de Recherches Internationales Servier, Suresnes, France

<sup>4</sup>Julius Center for Health Sciences and Primary Care, Medical Humanities, University Medical Center Utrecht, Utrecht, The Netherlands

## Correspondence to

Dr Shona Kalkman, Medical Humanities, University Medical Center Utrecht, 3584 CX Utrecht, The Netherlands; [s.kalkman@umcutrecht.nl](mailto:s.kalkman@umcutrecht.nl)

Received 24 June 2019

Revised 24 October 2019

Accepted 28 October 2019

## ABSTRACT

**Introduction** International sharing of health data opens the door to the study of the so-called 'Big Data', which holds great promise for improving patient-centred care. Failure of recent data sharing initiatives indicates an urgent need to invest in societal trust in researchers and institutions. Key to an informed understanding of such a 'social license' is identifying the views patients and the public may hold with regard to data sharing for health research.

**Methods** We performed a narrative review of the empirical evidence addressing patients' and public views and attitudes towards the use of health data for research purposes. The literature databases PubMed (MEDLINE), Embase, Scopus and Google Scholar were searched in April 2019 to identify relevant publications. Patients' and public attitudes were extracted from selected references and thematically categorised.

**Results** Twenty-seven papers were included for review, including both qualitative and quantitative studies and systematic reviews. Results suggest widespread—though conditional—support among patients and the public for data sharing for health research. Despite the fact that participants recognise actual or potential benefits of data research, they expressed concerns about breaches of confidentiality and potential abuses of the data. Studies showed agreement on the following conditions: value, privacy, risk minimisation, data security, transparency, control, information, trust, responsibility and accountability.

**Conclusions** Our results indicate that a social license for data-intensive health research cannot simply be presumed. To strengthen the social license, identified conditions ought to be operationalised in a governance framework that incorporates the diverse patient and public values, needs and interests.

## INTRODUCTION

Large-scale, international data sharing opens the door to the study of so-called 'Big Data', which holds great promise for improving patient-centred care. Big Data health research is envisioned to take precision medicine to the next level through increased understanding of disease aetiology and phenotypes, treatment effects, disease management and healthcare expenditure.<sup>1</sup> However, lack of public trust is proven to be detrimental to the goals of data sharing.<sup>2</sup> The case of *care.data* in the UK offers a blatant example of a data sharing initiative gone awry. Criticism predominantly focused on limited public awareness and lack of clarity on the

goals of the programme and ways to opt out.<sup>3</sup> Citizens are becoming increasingly aware and critical of data privacy issues, and this warrants renewed investments to maintain public trust in data-intensive health research. Here, we use the term *data-intensive* health research to refer to a practice of grand-scale capture, (re)use and/or linkage of a wide variety of health-related data on individuals.

Within the European Union (EU), the recently adopted General Data Protection Regulation (GDPR) (EU 2016/679) addresses some of the concerns the public may have with respect to privacy and data protection. One of the primary goals of the GDPR is to give individuals control over their personal data, most notably through consent.<sup>4</sup> Other lawful grounds for the processing of personal data are listed, but it is unclear how these would exactly apply to scientific research. Legal norms remain open to interpretation and thus offer limited guidance to researchers.<sup>5,6</sup> In Recital 33, the GDPR actually mentions that additional ethical standards are necessary for the processing of personal data for scientific research. This indicates a recognised need for entities undertaking activities likely to incite public unease to go beyond compliance with legal requirements.<sup>7</sup> Complementary ethical governance then becomes a prerequisite for securing public trust in data-intensive health research.

A concept that could be of use in developing ethical governance is that of a 'social license to operate'.<sup>7</sup> The social license captures the notion of a mandate granted by society to certain occupational groups to determine for themselves what constitutes proper conduct, under the condition that such conduct is in line with society's expectations. The term 'social license' was first used in the 1950s by American sociologist Everett Hughes to address relations between professional occupations and society.<sup>8</sup> The concept has been used since to frame, for example, corporate social responsibility in the mining industry,<sup>9</sup> governance of medical research in general<sup>8</sup> and of data-intensive health research more specifically.<sup>7,10</sup> As such, adequate ethical governance then becomes a precondition for obtaining a social license for data sharing activities.

Key to an informed understanding of the social license is identifying the expectations society may hold with regard to sharing of and access to health data. Here, relevant societal actors are the subjects of Big Data health research, constituting both patients and the general public. Identification of patients' and public views and attitudes allows



© Author(s) (or their employer(s)) 2019. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

**To cite:** Kalkman S, van Delden J, Banerjee A, et al. *J Med Ethics* Epub ahead of print: [please include Day Month Year]. doi:10.1136/medethics-2019-105651

for a better understanding of the elements of a socially sanctioned governance framework. We know of the existence of research papers that have captured these views using quantitative or qualitative methods or a combination of both. So far, systematic reviews of the literature have limited their scope to citizens of specific countries,<sup>11 12</sup> qualitative studies only<sup>13</sup> or the sharing of genomic data.<sup>14</sup> Therefore, we performed an up-to-date narrative review of both quantitative and qualitative studies to explore predominant patient and public views and attitudes towards data sharing for health research.

## METHODS

We searched the literature databases PubMed (MEDLINE), Embase, Scopus and Google Scholar in April 2019 for publications addressing patients' and public views and attitudes towards the use of health data for research purposes. Synonyms of the following terms (connected by 'AND') were used to search titles and/or abstracts of indexed references: *patient* or *public*; *views*; *data sharing*; *research* (See [box 1](#) and online supplementary appendix 1). To merit inclusion, an article had to report results from an original research study (qualitative, quantitative or mixed methods) on attitudes of individuals regarding use of data for health research. We restricted eligibility to records published in English and studies performed between 2009 and 2019. We chose 2009 as a lower limit because we assume that patients' and public perspectives might have changed substantially with increasing awareness and use of digital (health) technologies. Systematic reviews and meta-analyses synthesising the empirical literature on this topic also qualified for review. Reports from stakeholder meet-ups and workshops were eligible as long as they included patients or the public as participants. Since we were only interested in empirical evidence, expert opinion and publications merely advocating for the inclusion of patients' and public views in Big Data health research were excluded. Studies that predominantly reported on views of other stakeholders—such as clinicians, researchers, policy makers or industry—were excluded. Articles reporting on conference proceedings, or views regarding (demographic) data collection in low or middle income countries or for public health and care/quality improvement were not considered relevant to this review. Despite our specific interest in data sharing within the European context, we broadened eligibility criteria to include studies performed in the USA, Canada, Australia and New Zealand. Additional articles were identified through consultation with experts and review of references in the manuscript identified through the literature database searches. Views and attitudes of patients and the public were identified from selected references and reviewed by means of thematic content analysis.

### Box 1 Key search terms

```
(patient* OR public OR citizen*)
AND
(attitude* OR view* OR perspective* OR opinion* OR interview*
OR qualitative* OR questionnaire* OR survey*)
AND
("data sharing" OR "data access" OR "data transfer")
AND
Research
```

Asterisks ("\*") are used as a wildcard to allow any given search terms to be truncated or remain the same.

## RESULTS

### Study characteristics

Searches in PubMed (MEDLINE), Embase, Scopus and Google Scholar resulted in a total of 1153 non-unique records (see online supplementary appendix 1). We identified 27 papers for review, including 12 survey or questionnaire studies (quantitative), 8 interview or focus group studies (qualitative), 1 mixed methods study and 6 systematic reviews (see [table 1](#)). Most records were excluded because they were not relevant to our research question or because they did not report on findings from original (empirical) research studies. Ten studies reported on views of patients, 11 on views of the public/citizens and 6 studies combined views of patients, research participants and the public.

### Willingness to share data for health research

Reviewed papers suggest widespread support for the sharing of data for health research.

Four systematic reviews synthesising the views of patients and the public report that willingness for data to be linked and shared for research purposes is high<sup>11–14</sup> and that people are generally open to and understand the benefits of data sharing.<sup>15</sup>

Outpatients from a German university hospital who participated in a questionnaire study (n=503) expressed a strong willingness (93%) to give broad consent for secondary use of data,<sup>16</sup> and 93% of a sample of UK citizens with Parkinson's disease (n=306) were willing to share their data.<sup>17</sup> Wide support for sharing of data internationally<sup>18 19</sup> and in multicentre studies<sup>20</sup> was reported among patient participants. Goodman *et al* found that most participants in a sample of US patients with cancer (n=228) were willing to have their data made available for 'as many research studies as possible'.<sup>21</sup> Regarding the use of anonymised healthcare data for research purposes, a qualitative study found UK rheumatology patients and patient representatives in support of data sharing (n=40).<sup>22</sup>

Public respondents in survey studies recognised the benefits of storing electronic health information,<sup>23</sup> and 78.8% (n=151) of surveyed Canadians felt positive about the use of routinely collected data for health research.<sup>24</sup> The majority (55%) of a sample of older Swiss citizens (n=40) were in favour of placing genetic data at disposal for research.<sup>25</sup> Focus group discussions convened in the UK showed that just over 50% of the members of the Citizens Council of The National Institute for Health and Care Excellence (NICE) said they would have no concerns about NICE using anonymised data derived from personal care records to evaluate treatments,<sup>26</sup> and all participants in one qualitative study were keen to contribute to the National Healthcare Service (NHS)-related research.<sup>27</sup>

### Motivations to share data

Patients and public participants expressed similar reasons and motivations for their willingness to share data for health research, including contributing to advancements in healthcare, returning incurred benefits and the hope of future personal health benefits ([tables 2–4](#)).

In the two systematic reviews that addressed this topic, sharing data for 'the common good' or 'the greater good' was identified as one of the most prevalent motivations.<sup>12 14</sup>

For patients specifically, to help future patients or people with similar health problems was an important reason.<sup>14 16</sup> One survey study conducted among German outpatients found that 72% listed returning their own benefits incurred from research as a driver for sharing clinical data.<sup>16</sup> Patients with rare disease were also motivated by 'great hope and trust' in the development of international databases for health research.<sup>19</sup> Among patients,

Table 1 Study characteristics

| No. | Reference                                   | Perspective         | Study aim   | Date of data collection                  | Setting   | Sample (n, gender, age, etc)  | Method of data collection  |
|-----|---|---------------------|---|--|---|---|--|
| 1   | O'Brien <i>et al</i> , 2019 <sup>28</sup>   | Patients            | To examine patient perspective on the risks and benefits of linking existing data sources for research.   | Between December 2015 and February 2016. | Online patient community PatientsLikeMe (PLM).  | n=3516; female (73.8%); >65 years (14.5%); Caucasian (86.4%); completed college/postgraduate education (44.9%).                       | Questionnaire  |
| 2   | McCormick <i>et al</i> , 2019 <sup>24</sup> | Public              | To benchmark the views of Canadians about the use of administrative/routinely collected data for health research.   | Between January and August 2017.         | Websites, email and social media of three Canadian joint and skin disease patients' organisations.  | n=151; female (77.5%); British Columbians (55.6%); university graduates (57.6%); chronic disease (66.9%).                             | Cross-sectional online survey  |
| 3   | Colombo <i>et al</i> , 2019 <sup>30</sup>   | Patients and public | To gather knowledge on the opinions and attitudes of Italian patient and citizen groups on individual participant data sharing from clinical studies.                     | Between June 2017 and November 2017.     | Contacts of patient and citizen groups in Italy.  | n=280; oncology and palliative care (32.1%); operated locally or regionally (46.2%); involved in clinical research (48.6%).           | Cross-sectional online survey  |
| 4   | Richter <i>et al</i> , 2019 <sup>16</sup>   | Patients            | To examine whether abolishing consent for secondary data use would be acceptable to patients  | Between March 2018 and May 2018.         | Outpatients of a northern German university hospital  | n=503; female (65%); >60 years (≈18%); completed high school (≈21%).  | Questionnaire  |
| 5   | Stockdale <i>et al</i> <sup>12</sup>        | Public              | To systematically review the literature on UK and Irish public views of patient data used in research.  | Studies published between 2006 and 2016. | Studies using a UK or Irish sample.   | 20 UK and Ireland based papers (qualitative, qualitative and mixed methods).  | Systematic review  |
| 6   | Shah <i>et al</i> , 2019 <sup>34</sup>      | Patients            | To investigate research participants' beliefs about the importance of protecting their privacy, advancing research quickly and controlling future data sharing            | Not specified                            | Subset of participants in four European countries enrolled in the DIRECT (Diabetes Research on Patient Stratification) project.   | n=855; >60 years (73%); female (43%); qualifications above secondary school (60%); diabetes type 2 (70%).                             | Survey   |
| 7   | Shah <i>et al</i> , 2018 <sup>36</sup>      | Patients and public | To understand participants' future data governance preferences.   | Between September 2015 and March 2016.   | Patients diagnosed with diabetes type 2 and individuals at high risk of the disease but not receiving treatment for diabetes (participants enrolled in the DIRECT project)  | n=855; >60 years (73%); female (43%); vocational or professional qualifications (41%); degree level (19%); secondary education (37%). | Survey   |
| 8   | Howe <i>et al</i> , 2018 <sup>15</sup>      | Patients and public | To systematically review international evidence of research participants' attitudes towards the sharing of data for secondary research use.                               | Studies published between 2002 and 2017. | Studies originating from Japan, Thailand, India, Kenya, Canada, Vietnam and the USA.  | 9 papers included for review (8/9 qualitative studies)  | Systematic review  |
| 9   | Goytia <i>et al</i> , 2018 <sup>32</sup>    | Patients            | To gain insight from stakeholders into their understanding of Big Data, interest and concerns in contributing to health research.   | Not specified                            | Patients and disease groups (rare and chronic) from free-standing community organisations and disease support groups from various neighbourhoods in New York City (USA).  | n=138 (from eight patient/advocate groups); female (85%); non-white (91%); experience as participants in research studies (33%).      | Qualitative study based on 'opportunistic' listening sessions led by trained facilitators during pre-existing patient, community and clinician group meetings. |
| 10  | Mählmann <i>et al</i> , 2017 <sup>25</sup>  | Public              | To assess the willingness of older Swiss adults to share genetic data for research purposes and to investigate factors that might impact their willingness to share data. | Between December 2013 and April 2014.    | Older Swiss adults attending the Seniorenuniversität Zürich, Switzerland.   | n=40; female (52.5%); respondents aged between 67 and 92 years.   | Semistructured interviews  |
| 11  | Mursaleen <i>et al</i> , 2017 <sup>17</sup> | Patients            | To establish patient attitudes to ownership and sharing of their own medical data.  | Between June 2016 and September 2016.    | People with Parkinson's disease in the UK.  | n=306; female (55%); between 55 and 74 years (68%); mean number of years diagnosed 7.1  | Online survey  |
| 12  | Mazor <i>et al</i> , 2017 <sup>20</sup>     | Patients            | To understand stakeholders' views on data sharing in multicentre comparative effectiveness research studies.  | Between June 2015 and February 2016      | US patients from two existing groups: (1) a bariatric surgery patient advisory panel; and (2) patients who participated in the Arthritis Partnership with Comparative Effectiveness Research, a Patient-Powered Research Network within the National Patient-Centred Clinical Research Network (PCORnet). | n=15 patients   | Qualitative study based on interviews  |
| 13  | Goodman <i>et al</i> , 2017 <sup>21</sup>   | Patients            | To examine participant preferences regarding the use of deidentified data in large research datasets  | 2013                                     | US cancer patients recruited from the Northwest Cancer Genetics Registry.   | n=228; female (63.6%); mean age 64.3 years; white (93.3%); bachelor's degree (55.3%).   | Online survey  |
| 14  | Sanderson <i>et al</i> , 2017 <sup>31</sup> | Public              | To assess willingness to participate in a biobank using different consent and data sharing models.  | Between April and July 2015.             | Participants recruited at multiple healthcare systems participating in the Electronic Medical Records and Genomics (eMERGE) Network (USA).  | n=13 000; female (63%); self-identified white (51%); less than a bachelor's degree (42%); annual household income ≤\$60 000 (44%).    | Survey   |
| 15  | Patil <i>et al</i> , 2016 <sup>23</sup>     | Public              | To assess the public's preferences regarding potential privacy threats from devices or services storing health-related personal data.                                     | Between August and November 2013.        | Respondents from 27 EU member countries.  | n=20 882; female (52.3%); ≥65 years (19.1%)   | Survey   |

Continued

Table 1 Continued

| No. | Reference                                   | Perspective         | Study aim  | Date of data collection                  | Setting  | Sample (n, gender, age, etc)   | Method of data collection  |
|-----|---|---------------------|--|--|--|--|--|
| 16  | Aitken <i>et al</i> , 2016 <sup>13</sup>    | Public              | To systematically review the literature examining public attitudes towards the sharing or linkage of health data for research purposes.  | Studies conducted between 1999 and 2013. | Studies primarily originating from the UK and USA.   | 25 studies included for review (focus groups, interviews, deliberative events, dialogue workshops)   | Systematic review  |
| 17  | Spencer <i>et al</i> , 2016 <sup>22</sup>   | Patients            | To explore patient perspectives on the use of anonymised healthcare data for research purposes.  | Not specified                            | Patients recruited from a rheumatology outpatient clinic and from a patient and public involvement health research network (UK).   | n=40; female (58%); ages ranged from 23 to 88 years (mean 61); self-identified white British (97.5%); chronic rheumatic disease (100%).  | Qualitative study based on 26 interviews and three focus groups.   |
| 18  | McCormack <i>et al</i> , 2016 <sup>18</sup> | Patients            | To document rare disease patients' attitudes to participation in genomics research, particularly around large-scale, international data and biosample sharing.                                   | 2014                                     | Rare disease patients recruited during the EURORDIS Membership Meeting at the European Conference on Rare Diseases 2014 in Berlin and the EURORDIS Summer School for Expert Patients 2014 in Barcelona.  | n=52; female (61.5%); from 16 countries.   | Qualitative study based on focus group discussions   |
| 19  | NICE Citizens Council, 2015 <sup>26</sup>   | Public              | To explore citizens' views regarding the ethical and practical issues that need to be considered in the use of anonymised information derived from personal care records to evaluate treatments. | 2015                                     | The NICE Citizens Council is a panel of 30 members of the public that provides a public perspective on challenging social and moral issues that NICE needs to take into account when producing guidance. | n=30   | Qualitative study based on facilitated discussions at the annual 2 day meeting of the NICE Citizens Council. |
| 20  | Garrison <i>et al</i> , 2016 <sup>11</sup>  | Patients and public | To systematically review attitudes towards biobanking, broad consent and data sharing  | Studies conducted between 2001 and 2015. | Studies conducted in the USA.  | 48 papers including a total of 35 969 individuals; female (54.2%); self-identified white (51.3%).  | Systematic review  |
| 21  | Joly <i>et al</i> , 2015 <sup>33</sup>      | Public              | To examine public views about governance structure, consent and data sharing in biobanking.  | Between February 2013 and July 2014.     | Canadian adults who self-identified as being a past or potential future donor of tissue samples or genetic data to a biobank or genetic database.  | n=114; female (46%); ≥50 years (32%); did not attend university (50%).   | Survey   |
| 22  | Darquy <i>et al</i> , 2016 <sup>19</sup>    | Patients            | To explore patient views on the sharing of their medical data in the context of compiling a European rare disease database.  | 2012                                     | Participants recruited from 5 European countries through the European Leukodystrophies Association and LeukoTreat partners.  | n=46   | Questionnaire  |
| 23  | Taylor and Taylor, 2014 <sup>27</sup>       | Public              | To investigate public views about preferable/acceptable consent models for use of personal confidential data in health research.   | Not specified                            | People with different levels and kinds of involvement in the National Health Service and/or health research.   | n=28   | Mixed methods incorporating a structured questionnaire and in-depth focus group discussions.                 |
| 24  | Shabani <i>et al</i> , 2014 <sup>14</sup>   | Patients and public | To solicit public and research participants' attitudes with respect to genomic data sharing.   | Studies published between 2008 and 2013. | –  | 15 papers included for review (quantitative and qualitative).  | Systematic review  |
| 25  | Hill <i>et al</i> , 2013 <sup>27</sup>      | Public              | To determine the range of public opinion about the use of existing medical data for research and to explore views about consent to a secondary review of medical records for research.           | Not specified                            | Reviewed studies conducted in the USA, UK, Ireland, Canada and New Zealand. Older men recruited from rural and suburban primary care practices in the UK.  | 27 papers included for review (quantitative, qualitative, systematic reviews). n=19; female (0%); ≥50 years (100%); mean age 61 years.   | Systematic review and qualitative study (focus group).   |
| 26  | Haga & O'Daniel, 2011 <sup>29</sup>         | Public              | To explore public attitudes regarding data sharing practices in genomics research.   | Between 2008 and Between, 2009.          | Focus groups convened in Durham (North Carolina), USA.   | n=100; female (73%); African–American (76%), median age 40–49 years.   | Qualitative study based on 10 focus group discussions.   |
| 27  | Lemke <i>et al</i> , 2010 <sup>35</sup>     | Patients and public | To assess public and biorepository participant attitudes towards research participation and sharing of genetic research data.  | May 2008                                 | 49 individuals recruited from diverse Chicago (USA) neighbourhoods, of whom 28 in 3 public focus groups and 21 in 3 NUGene biorepository participant focus groups.                                       | n=28 public respondents; female (75%); some college education or more (75%); African–American (46%)<br>n=21 participant respondents; female (67%); some college education or more (95%); Caucasian (76%) . | Qualitative study based on six focus group discussions.  |

NICE, National Institute for Health and Care Excellence.

support of research in general,<sup>16</sup> the value attached to answering 'important' research questions,<sup>20</sup> and a desire to contribute to advancements in medicine<sup>14</sup> were prevalent reasons in favour of data sharing. Ultimately, the belief that data sharing could lead to improvements in health outcome and care was reported.<sup>20</sup>

Only one original study research paper addressed public motivations. This study found that older citizens mentioned altruistic reasons and the greater good in a series of interviews as reasons

to share genetic data for research.<sup>25</sup> In these interviews, citizens expressed no expectations of an immediate impact or beneficial return but ultimately wanted to help the next generation.

### Perceived benefits of data sharing

Patients and the public perceive that data sharing could lead to better patient care through improved diagnosis and treatment



**Table 2** Patients' views and attitudes towards the sharing of health data for research

| Overall willingness to share data   | Motivations to share data   | Perceived benefits of data sharing   | Perceived risks of data sharing   | Barriers to share data  | Factors affecting willingness to share data   | Conditions for sharing  |
|---|---|--|---|---|---|---|
| Strong willingness to give broad consent for secondary data use (93%) <sup>16</sup>   | 98% considered that the altruistic benefits of sharing healthcare data outweighed the risks <sup>14</sup> | 'Helping my doctor make better decisions about my health' (94%) <sup>28</sup>                        | Concerns that data could not truly be deidentified <sup>32</sup>  | If data sharing is motivated by financial gain or profit <sup>20</sup>  | Significant association between country and attitudes towards data sharing <sup>34</sup>  | Protection of privacy <sup>34</sup>   |
| Strong approval of abolishing patient consent (76%) <sup>16</sup>   | Helping future patients (67%) <sup>16</sup>   | Improving patient care and advancing understanding of treatment risks and side effects <sup>20</sup> | Data security risks <sup>16 18 20</sup>   | If an entity might profit by selling their data <sup>20</sup>   | Significant association between age and data sharing <sup>17</sup>  | Studies that offer value and minimise security risks <sup>20</sup>  |
| 93% of respondents were willing to share data <sup>17</sup>   | Returning own benefits from research (72%) <sup>16</sup>  | Increased chance of receiving personal health information <sup>21</sup>                              | Concerns about misuse of data <sup>18</sup>   | Lack of transparency and awareness around the use of data, making it difficult to secure public trust <sup>22</sup> | No clear relationship was found between data sharing and the number of years diagnosed, sex, medication class or health confidence <sup>17</sup>  | Researchers protect patients' privacy and information <sup>21</sup>   |
| Stakeholders were open to data sharing in multicentre studies <sup>20</sup>   | Improving health outcomes or care <sup>20</sup>   |  | Health data being 'stolen by hackers' (87%) <sup>28</sup>   |   | Expressions of trust and attitudes to risk are often affected by the nature of the rare disease a patient has, as well as regulatory and cultural practices in their home country <sup>18</sup> | 20% found that participants and their de-identified data may remain linked to allow for return of individual health results and to support further research <sup>21</sup> |
| Most participants expressed a desire that their data should be available for as many research studies as possible <sup>21</sup>       | Answering important research questions <sup>20</sup>  |  | Detrimental consequences of data falling into the wrong hands, such as insurance companies <sup>22</sup>  |   | Perceptions of the benefits and value of research increased willingness <sup>20</sup>   | No consensus among patients on ownership of, access to and usage of their research data <sup>17</sup>   |
| Patients were supportive of sharing their anonymised electronic patient record for research <sup>22</sup>                             | Great hope and trust in the development of this type of research <sup>19</sup>                            |  | Perceived possibility that open knowledge could lead to discrimination as having a rare disease was recognised as an inherent vulnerability <sup>18</sup> |   | Electronic interface system as a means of enabling feedback regarding data recipients and associated research results increased willingness <sup>22</sup>                                       | Need for information and transparency on database governance <sup>19</sup>  |
| Participants positively disposed towards research and towards allowing data and biosamples to be shared internationally <sup>18</sup> | Support of research in general (86%) <sup>16</sup>  |  |   |   | Willingness increased if social security number (90%) and insurance ID (82%) were removed from the data for linkage and research use <sup>28</sup>  | Transparency in conditions framing access to data, all research conducted, partnerships with the pharmaceutical industry <sup>19</sup>                                    |
| Patients were strongly in favour of sharing data in the context of compiling a European leukodystrophies database <sup>19</sup>       |   |  |   |   | Privacy-protecting methods that share summary-level data (though concerns about increased cost and loss of validity) increased willingness <sup>20</sup>  | Use of an electronic interface to enable greater control over consent choices <sup>22</sup>   |
|   |   |  |   |   | Having control over what data are shared and with whom increased willingness <sup>34</sup>  | Wanting to know who does research (34%) <sup>16</sup>   |
|   |   |  |   |   | Willingness increased if patients were able to learn how their data was protected (84%) <sup>28</sup>   | Wanting to know type of research (37%) <sup>16</sup>  |
|   |   |  |   |   |   | Patients need access to research results <sup>19</sup>  |
|   |   |  |   |   |   | Need for public campaigns to inform stakeholders about Big Data <sup>32</sup>   |
|   |   |  |   |   |   | Central role of clinicians in introducing patients to research <sup>32</sup>  |

options and more efficient use of resources. Patients seem to also value the potential of (direct) personal health benefits.

Two systematic reviews reported on perceived benefits of data sharing for health research purposes. Howe *et al* mentioned perceived benefits to research participants or the immediate community, benefits to the public and benefits to research and science.<sup>15</sup> Shabani *et al* also listed accelerating research advancement and maximising the value of resources as perceived benefits.<sup>14</sup>

Surveyed patients perceived that data sharing could help their doctor 'make better decisions' about their health (94%, n=3516)<sup>28</sup> or result in an increased chance of receiving personalised health information (n=228).<sup>21</sup>

In the original studies reviewed, advantages and potential benefits of data sharing were generally recognised by public and patient participants.<sup>22 29</sup> Data sharing was believed to enable the study of long-term treatment effects and rare events, as well as the study of large numbers of people,<sup>24</sup> to improve diagnosis<sup>25</sup>

**Table 3** Public views and attitudes towards the sharing of health data for research

| Overall willingness to share   | Motivations to share  | Perceived benefits of data sharing  | Perceived risks of data sharing   | Barriers to share data  | Factors affecting willingness to share data   | Conditions for sharing  |
|--|---|---|---|---|---|---|
| 78.8% felt positively about the use of routinely collected data for health research <sup>24</sup>  | Altruistic reasons and contributing to the greater good <sup>25</sup>   | 'Discussions recognised the benefits of data-sharing' <sup>29</sup>   | 42.4% lacked confidence in data security and privacy <sup>24</sup>  | Data sharing with private companies <sup>25</sup>   | Findings suggested that males and older people were more likely to consent to a review of their medical data (systematic review) <sup>27</sup>  | De-identification of personal information as a top privacy measure (89.4%) <sup>24</sup>  |
| Widespread willingness to share patient data for research <sup>12</sup>  | Sharing for the common good <sup>12</sup>   | Finding new treatments and improving diagnosis <sup>25</sup>  | 51% would worry about their privacy <sup>31</sup>   | Respondents were strongly averse to health insurance companies, private sector pharmaceutical companies and academic researchers viewing their data <sup>23</sup> | Generational differences impacted willingness <sup>25</sup>   | Consent procedures should be audited and an ombudsman should oversee the governance of the use of personal care information for research <sup>26</sup>  |
| The majority of participants were in favour of placing genetic data to research's disposal <sup>25</sup>   | No expectation of an immediate impact or beneficial return but ultimately wanting to help the next generation <sup>25</sup> | Respondents agreed that storage was important for improving treatment quality (75.5%), preventing epidemics (63.9%) and reducing delays (58.9%) <sup>23</sup> | Concerns about privacy and confidentiality <sup>13,29</sup>   |   | Willingness to participate was associated with self-identified white race <sup>31</sup>   | Acceptance of alternative consent models conditional on a number of factors, including: security and confidentiality, no inappropriate commercialisation or detrimental use, transparency, independent overview, the ability to object to any processing considered to be inappropriate or particularly sensitive <sup>37</sup> |
| 66% stated they would be willing to participate in a biobank <sup>31</sup>   |   | Ability to study long-term treatment effects and rare events (75.5%) <sup>24</sup>  | Concerns about a party's competence in keeping data secure <sup>12,25,26</sup>  |   | Willingness to participate was associated with higher educational attainment <sup>31</sup>  | Important to inform research participants of a study's data-sharing plans during the informed consent process <sup>29</sup>   |
| Respondents recognised the benefits of storing electronic health information <sup>23</sup>   |   | Ability to study large numbers of people (72.8%) <sup>24</sup>  | Concerns about different levels of access by third parties were expressed by 48.9%–60.6% <sup>23</sup>  |   | Willingness to participate was associated with lower religiosity <sup>31</sup>  | NICE should hold open days and provide information resources designed to ensure people understand what data are being used for, precisely how it will be used and providing reassurance that personal care data will not be passed on or sold to other organisations <sup>26</sup>  |
| Widespread general—though conditional—support for data linkage and data sharing for research purposes <sup>13</sup>  |   |   | Concerns about potential for data to be sold on to other organisations and used for profit and for purposes other than research <sup>26</sup> |   | Willingness to participate was associated with perceiving more research benefits, fewer concerns and fewer information needs <sup>31</sup>  | Information provision to participants about identified biobank objectives, governance structure and accountability <sup>33</sup>  |
| Just over 50% of the members of the Council said they would have no concerns about NICE using anonymised data derived from personal care records <sup>26</sup>         |   |   | Concerns about data sharing for commercial gain and the potential misuse of information (focus groups) <sup>27</sup>                          |   | Willingness increased if there was perceived actual or potential public benefits from the research <sup>13</sup>  | Appropriate systems and good working practices should be put in place to ensure a consistent approach to research planning, data capture and analysis <sup>26</sup>   |
| Most expressed willingness for their data to be shared with the international scientific community rather than used by one or more Canadian institutions <sup>33</sup> |   |   | Concerns about potential misuse by insurers, the government and other third parties <sup>33</sup>   |   | Willingness increased if there was trust in the individuals or organisations conducting and/or overseeing data linkage/sharing <sup>13</sup>  | Most (86%) participants would want to know what would happen if a researcher misused their health information <sup>31</sup>   |
| Over half the respondents preferred to give a one-time general consent for the future sharing of their samples among researchers <sup>33</sup>                         |   |   | Misuse and abuse of data <sup>13,25</sup> and potential harms arising <sup>13</sup>   |   | Sharing due to financial incentives impacted willingness <sup>25</sup>  |   |
| People are typically willing to accept models of consent other than that which they would prefer <sup>37</sup>   |   |   | Concerns relating to individuals' control over their data <sup>13</sup>   |   | Differences between sharing genetic data or health data impacted willingness <sup>25</sup>  |   |
| All participants were keen to contribute to NHS-related research (focus groups) <sup>27</sup>  |   |   | Concerns about control and ownership of biological samples and data <sup>33</sup>   |   | Sharing data with private companies impacted willingness <sup>25</sup>  |   |
|  |   |   | Concerns centred on transparency about how data are used and how it might be used in the future <sup>26</sup>                                 |   | Participants became more accepting of the use of precollected medical data without consent after being given information about selection bias and research processes (focus groups) <sup>27</sup> |   |

Continued

Table 3 Continued

| Overall willingness to share | Motivations to share | Perceived benefits of data sharing | Perceived risks of data sharing  | Barriers to share data | Factors affecting willingness to share data  | Conditions for sharing |
|------------------------------|----------------------|------------------------------------|--|------------------------|--|------------------------|
|                              |                      |                                    | Concerns about ensuring research is conducted according to good scientific practice and data are used to benefit society <sup>26</sup> |                        | 66.9% wanted to learn more about data stewards granting access to data <sup>24</sup>   |                        |
|                              |                      |                                    | Fear of becoming a transparent citizen <sup>25</sup>   |                        | Discussions were significantly more likely to participate in a study that planned to deposit data in a restricted access online database compared with an open access database <sup>29</sup> |                        |

NHS, National Healthcare Service; NICE, National Institute for Health and Care Excellence.

and treatment quality,<sup>20 23</sup> as well as to stimulate innovation<sup>30</sup> and identify new treatment options.<sup>25</sup> A cross-sectional online survey among patient and citizen groups in Italy (n=280) also identified the perception that data sharing could reduce waste in research.<sup>30</sup>

### Perceived risks of data sharing

The most significant risks of data sharing were perceived to results from breaches of confidentiality, commercial use and potential abuse of the data.

Systematic reviews report on patients' and public concerns about confidentiality in general,<sup>13 15</sup> sometimes linked to the risk of reidentification,<sup>14</sup> concerns about a party's competence in keeping data secure,<sup>12</sup> and concerns that personal information could be mined from genomic data.<sup>14</sup> A systematic review by Stockdale *et al* identified concerns among the public (UK and Ireland) about the motivation a party might have to use the data.<sup>14</sup>

Patients in a UK qualitative study (n=40) perceived 'detrimental' consequences of data 'falling into the wrong hands', such as insurance companies.<sup>22</sup> Respondents from the online patient community PatientsLikeMe were fearful of health data being 'stolen by hackers' (87%, n=3516).<sup>28</sup>

Original research studies flagged data security and privacy as major public concerns.<sup>16 18 20 25 26 29-32</sup> More specifically, many studies found that participants worried about who would have access to the data and about risk of misuses or abuses.<sup>13 15 18 25 27 33</sup> A large pan-European survey among respondents from 27 EU member states revealed public concerns about different levels of access by third parties (48.9%–60.6%, n=20 882).<sup>23</sup> Overall, reviewed papers suggest that patients and the public are concerned about the use of their data for commercial purposes.<sup>14 27</sup> For example, the NICE Citizens Council expressed concerns about the potential for data to be sold to other organisations and used for profit and for purposes other than research.<sup>26</sup> The Citizens Council also highlighted the need for transparency about how data are used and how it might be used in the future and for ensuring the research is conducted according to good scientific practice and that data are used to benefit society. Concerns about control and ownership of data were identified<sup>13 33</sup> and about re-use of data for purposes that participants do not agree on.<sup>30</sup> Fear of discrimination, stigmatisation, exploitation or other repercussions as a consequence of data being shared was widely cited by individuals.<sup>14 15 18</sup>

### Barriers to share data

Studies showed that patients and the public rarely mention barriers to data sharing in absolute terms. Rather, acceptance seemed to decrease if data sharing was financially motivated, and if people did not know how and with whom their data would be shared.

First, individuals often opposed data sharing if it was motivated by financial gain or profit<sup>20</sup> or if the data were shared with commercial/private companies.<sup>14 15</sup> In one large pan-European survey (n=20 882), respondents were found to be strongly averse to health insurance companies and private sector pharmaceutical companies viewing their data.<sup>23</sup> Second, lack of understanding and awareness around the use of data was viewed as a barrier to data sharing.<sup>15 22</sup> Third, lack of transparency and controllability in releasing data were mentioned as factors compromising public trust in data sharing activities.<sup>14 22</sup>

### Factors affecting willingness to share data

A wide range of factors were identified from the literature that impacted individuals' willingness to share data for health research, including geographical factors, age, individual-specific and research-specific characteristics.

#### Geographical factors

McCormack *et al* found that European patients' expressions of trust and attitudes to risk were often affected by the regulatory and cultural practices in their home countries, as well as by the nature of the (rare) disease the patient participant had.<sup>18</sup> Shah *et al* conducted a survey among patients in four Northern European countries (n=855) and found a significant association between country and attitudes towards sharing of deidentified data.<sup>34</sup> Interestingly, Dutch respondents were less likely to support sharing of their deidentified data compared with UK citizens.

#### Age

Among a sample of surveyed patients with Parkinson's disease (UK), a significant association was found between higher age and increased support for data sharing.<sup>17</sup> According to a study based on semistructured interviews with older Swiss citizens, generational differences impacted willingness to share.<sup>25</sup> With respect to public attitudes towards data sharing, findings of one systematic review suggest that males and older people are more likely to consent to sharing their medical data.<sup>27</sup> A systematic review by Shabani *et al* suggests that patient and public participants with

**Table 4** Patients' and public views and attitudes towards the sharing of health data for research

| Overall willingness to share   | Motivations to share  | Perceived benefits of data sharing                           | Perceived risks of data sharing  | Barriers to share data  | Factors affecting willingness to share data   | Conditions for sharing   |
|--|---|--|--|---|---|--|
| 39% approved broad access by researchers and other professions <sup>30</sup>   | 'To help' people who have similar health problems <sup>14</sup> | Benefits to the public <sup>15</sup>                         | Privacy, confidentiality and risk of data reidentification <sup>14 15 30</sup>   | Lack of understanding and awareness <sup>15</sup>   | Participants with higher mean age were substantially less worried about privacy and confidentiality than other groups <sup>14</sup>   | The research being congruent with the participants' values <sup>15</sup>   |
| 97% were supportive of sharing data postproject <sup>36</sup>  | 'To contribute' to advancements in medicine <sup>14</sup>       | Benefit to participants or immediate community <sup>15</sup> | Concerns about misuse of data <sup>15</sup>                                      | Lack of controllability in releasing data <sup>14</sup>   | Willingness lower among individuals from under-represented minorities <sup>11</sup>   | The research being in the public's interest <sup>15</sup>  |
| Participants are open to and understand the advantages of data sharing <sup>15</sup>   | To serve the greater good <sup>14</sup>                         | Benefits to science or research <sup>15</sup>                | Concern that personal information could be mined from genomic data <sup>14</sup> | Sharing with commercial companies <sup>14</sup>   | Willingness lower among individuals with privacy and confidentiality concerns <sup>11</sup>   | Samples are de-identified <sup>11 30 36</sup>  |
| Willingness for data to be shared was high <sup>11</sup>   |   | Accelerating research advancement <sup>14</sup>              | Concerns about potential use of their data by for-profit entities <sup>14</sup>  | Sharing genomic data with a broader group of researchers and for a variety of research purposes <sup>14</sup> | Willingness to endorse data sharing increases with trust in institutions and researchers <sup>14</sup>  | Respecting privacy of data <sup>11 14</sup>  |
| Broad consent was often preferred over tiered or study-specific consent, particularly when broad consent was the only option <sup>11</sup> |   | Advancement of innovation <sup>30</sup>                      | Re-use of data for purposes that participants do not agree on <sup>30</sup>      |   | Distrust of the government as an oversight body for genetic research data <sup>35</sup>   | Risks are mitigated <sup>15</sup>  |
| 90% were supportive to share data with universities postproject <sup>36</sup>  |   | Maximising the value of resources <sup>14</sup>              | Fear of using data for discriminatory purposes <sup>14</sup>                     |   | Participants with some college or a college degree were more likely to choose restricted data release <sup>14</sup>   | Highly secure database <sup>30 36</sup>  |
| 56% were supportive to share data with commercial companies postproject <sup>36</sup>  |   | Reducing waste in research <sup>30</sup>                     | Stigmatisation or repercussions <sup>15</sup>                                    |   | Personal perceptions of sensitivity of genomic data were influenced by elements such as race, gender, age, marital status and/or educational level <sup>14</sup>              | Postproject Data Access Committee should involve a researcher from the original research project, a clinician, patient representative, and a participant in the original study <sup>36</sup> |
|  |   |  | Fear of exploitation <sup>15</sup>   |   | Reputation of and trust in research organisation <sup>35</sup>  | Data access agreements <sup>30</sup>   |
|  |   |  |  |   | Willingness lower when pharmaceutical companies had access to data <sup>11</sup>  | Researchers of original study to monitor data used by other researchers <sup>35</sup>  |
|  |   |  |  |   | Being asked for consent for each study would make participants (81%) feel 'respected and involved', and 74% agreed that they would feel that they 'had control' <sup>14</sup> | Need for transparency in data sharing and monitoring policies <sup>35</sup>  |
|  |   |  |  |   | Participants desire to be notified when their data are (re)used and to be informed of the results of studies using their data <sup>15</sup>                                   | Participants having understood that their data could be shared (transparency) <sup>15</sup>  |
|  |   |  |  |   | Participants desire to be involved in the data sharing process <sup>15</sup>  | Information on consequences of a breach of protection and penalties <sup>35</sup>  |
|  |   |  |  |   | To know what organisation/agency has oversight responsibilities for genetic research data as <sup>35</sup>  | Participants want to know more about how the data will be shared and with whom <sup>35</sup>   |
|  |   |  |  |   |   | Information to participants <sup>30</sup>  |
|  |   |  |  |   |   | Logistics of biobanks are communicated <sup>11</sup>   |
|  |   |  |  |   |   | Knowing more about how the data will be shared and with whom <sup>14</sup>   |
|  |   |  |  |   |   | Trust in the ability of the original institution to carry out the oversight tasks <sup>14</sup>  |
|  |   |  |  |   |   | Sanctions for misuse <sup>30</sup>   |

higher mean age are substantially less worried about privacy and confidentiality than other groups.<sup>14</sup>

**Individual-specific characteristics**

A systematic review into patients' and public perspectives on data sharing in the USA suggests that individuals from under-represented minorities are less willing to share data.<sup>11</sup> A large multisite survey (n=13 000) among the US public found that willingness to share was associated with self-identified white

race, higher educational attainment and lower religiosity.<sup>31</sup> In another systematic review, race, gender, age, marital status and/or educational level all seemed to influence how people perceived sensitivity of genomic data and the sharing thereof.<sup>14</sup> However, a UK study among patients with Parkinson's disease found no clear relationship between data sharing and the number of years diagnosed, sex, medication class or health confidence.<sup>17</sup>



Factors that clearly positively affected attitudes towards data sharing were perceptions of the (public) benefits and value of the research,<sup>13 20</sup> fewer concerns and fewer information needs,<sup>31</sup> and higher trust in and reputation of individuals or organisations conducting and/or overseeing data sharing.<sup>12–14 35</sup> Conversely, willingness decreased with higher privacy and confidentiality concerns<sup>11</sup> and higher distrust of the government as an oversight body for (genetic) research data.<sup>35</sup>

#### Research-specific characteristics

Privacy measures increased people's willingness to share their data for health research, such as removal of social security numbers (90%, n=3516) and insurance ID (82%, n=3516), the sharing of only summary-level or aggregate data<sup>20</sup> and deposition of data in a restricted access online database.<sup>29</sup> Expressions of having control over what data are shared and with whom positively affected attitudes towards data sharing.<sup>34</sup> In one study, being asked for consent for each study made participants (81%) feel 'respected and involved', and 74% agreed that they would feel that they 'had control'.<sup>14</sup> With respect to data sharing without prospective consent, participants became more accepting after being given information about the research processes and selection bias.<sup>27</sup> Less support was observed for data sharing due to financial incentives<sup>25</sup> and, more specifically, if data would be shared with private companies, such as insurance or pharmaceutical companies.<sup>11 25</sup>

#### Conditions for sharing

Widespread willingness to share data for health research very rarely led to participants' unconditional support. Studies showed agreement on the following conditions for responsible data sharing: value, privacy, minimising risks, data security, transparency, control, information, trust, responsibility and accountability.

#### Value

One systematic review found that participants found it important that the research as a result of data sharing should be in the public's interest and should reflect participants' values.<sup>15</sup> The NICE Citizens Council advocated for appropriate systems and good working practices to ensure a consistent approach to research planning, data capture and analysis.<sup>26</sup>

#### Privacy, risks and data security

The need to protect individuals' privacy was considered paramount<sup>11 14 21 34</sup> and participants often viewed deidentification of personal data as a top privacy measure.<sup>11 24 30 36</sup> One survey among US patients with cancer found that only 20% (n=228) of participants found linkage of individuals with their deidentified data acceptable for return of individual health results and to support further research.<sup>21</sup> Secured access to databases was considered an important measure to ensure data security in data sharing activities.<sup>30 34</sup> A systematic review of participants' attitudes towards data sharing showed that people established risk minimisation as another condition for data sharing.<sup>15</sup> Findings by Mazor *et al* suggest that patients only support studies that offer value and minimise security risks.<sup>20</sup>

#### Transparency and control

Conditions regarding transparency were information about how data will be shared and with whom,<sup>14 35</sup> the type of research that is to be performed, by whom the research will be performed,<sup>16</sup> information on data sharing and monitoring policies and database governance,<sup>35</sup> conditions framing access to data and data

access agreements,<sup>24 28 30</sup> and any partnerships with the pharmaceutical industry.<sup>19</sup> More generally, participants expressed the desire to be involved in the data sharing process,<sup>35</sup> to be notified when their data are (re)used and to be informed of the results of studies using their data.<sup>15</sup> Spencer *et al* identified use of an electronic interface as a highly valued means to enable greater control over consent choices.<sup>22</sup> When asked about the use of personal data for health research by the NHS, UK citizens were typically willing to accept models of consent other than the ones they would prefer.<sup>37</sup> Acceptance of consent models with lower levels of individual control was found to be dependent on a number of factors, including adequate transparency, control over detrimental use and commercialisation, and the ability to object, particularly to any processing considered to be inappropriate or particularly sensitive.<sup>37</sup>

#### Information and trust

One systematic review identified trust in the ability of the original institution to carry out the oversight tasks as a major condition for responsible data sharing.<sup>14</sup> Appropriate education and information about data sharing was thought to include public campaigns to inform stakeholders about Big Data<sup>32</sup> and information communicated at open days of research institutions (such as NICE) to ensure people understand what their data are being used for and to reassure them that personal data will not be passed on or sold to other organisations.<sup>26</sup> The informed consent process for study participation was believed to include information about the fact that individuals' data could potentially be shared,<sup>15 30</sup> the objectives of data sharing and (biobank) research, the study's data sharing plans,<sup>29</sup> governance structure, logistics and accountability.<sup>33</sup>

#### Responsibility and accountability

Participants often placed the responsibility for data sharing practices on the shoulders of researchers. Secondary use of data collected earlier for scientific research was viewed to require a data access committee that involves a researcher from the original research project, a clinician, patient representative and a participant in the original study.<sup>36</sup> Researchers of the original study were required to monitor data used by other researchers.<sup>36</sup> In terms of accountability, patient and public groups in Italy (n=280) placed high value on sanctions for misuse of data.<sup>30</sup> Information on penalties or other consequences of a breach of protection or misuse was considered important by many.<sup>31 35</sup>

## DISCUSSION

In this study, we narratively reviewed 27 papers on patients' and public views on and attitudes towards the use of health data for scientific research. Studies reported a widespread—though conditional—support for the linkage and sharing of data for health research. The only outlier seems to be the finding that just over half (n=25) of the NICE Citizens Council answered 'no' to the question whether they had any concerns if NICE used anonymised data to fill in the gaps if NICE was not getting enough evidence in 'the usual ways'.<sup>26</sup> However, we hasten to point out that the question about willingness to share is different from the question whether people have concerns or not. In addition, after a 2-day discussion meeting Council members were perhaps more sensitised to the potential concerns regarding data sharing. Therefore, we suggest that the way and context within which questions are phrased may influence the answers people give.

Overall, people expressed similar motivations to share their data, perceived similar benefits (despite some variation between patients and citizens), yet at the same time displayed a range of concerns, predominantly relating to confidentiality and data security, awareness about access and control, and potential harms resulting from these risks. Both patient and public participants conveyed that certain factors would increase or reduce their willingness to have their data shared. For example, the presence of privacy-protecting measures (eg, data deidentification and the use of secured databases) seemed to increase willingness to share, as well as transparency and information about data sharing processes and responsibilities. The identified views and attitudes appeared to come together in the conditions stipulated by participants: value, privacy and confidentiality, minimising risks, data security, transparency, control, information, trust, responsibility and accountability.

In our Introduction, we mentioned that identifying patients' and public views and attitudes allows for a better understanding of the elements of a socially sanctioned governance framework. In other words, what work should our governance framework be doing in order to obtain a social license? This review urges researchers and institutions to address people's diverse concerns and to make an effort to meet the conditions identified. Without these conditions, institutions lack trustworthiness, which is vital for the proceedings of medicine and biomedical science. As such, a social license is not a 'nice to have' but a 'need to have'. Our results also confirm that patients and the public indeed care about more than legal compliance alone, and wish to be engaged through information, transparency and control. This work supports the findings of a recent systematic review into ethical principles of data sharing as specified in various international ethical guidelines and literature.<sup>38</sup> What this body of research implies is considerable diversity of values and beliefs both between and within countries.

The goal of this narrative review was to identify the most internationally dominant, aggregated patient and public views about the broad topic of data sharing for health research. We deliberately opted for the methodology of a *narrative review* rather than a *systematic review*. Most narrative reviews deal with a broad range of issues to a given topic rather than addressing a particular topic in depth.<sup>39</sup> This means narrative reviews may be most useful for obtaining a broad perspective on a topic, and that they often are less useful in generating quantitative answers to specific clinical questions. However, because narrative reviews do not require specification of the search and selection strategy and the way of critically appraising literature can be variable, the connection between evidence generated by narrative reviews and (clinical) recommendations is less rigorous and risk of bias exists. This is something to take into account in this study. A risk of bias assessment was not possible due to the heterogeneity of the findings. We acknowledge that our methodological choices may have affected the discriminative power or granularity of our findings. For example, there is a difference between sharing of routinely collected health data versus secondary use of health data collected for research purposes. And we can only make loose assumptions about potential differences between patient and public views.

In addition, we should mention that this work is centred around studies conducted in Western countries as the whole Big Data space and literature is dominated by Western countries, higher socioeconomic status and Caucasians. However, most of the disease burden globally and within countries is most probably not represented in the 'Big Data' and so we have to stress the lack of generalisability to large parts of the world.

Nevertheless, we believe our findings point towards essential elements of a governance framework for data sharing for health research purposes. If we are to conclude that the identified conditions ought to act as the pillars of a governance framework, the next step is to identify how these conditions could be practically operationalised. For example, if people value information, transparency and control, what type of consent is most likely to valorise these conditions? And what policy for returning research results would be desirable? Once we know what to value, we can start thinking about the ways to acknowledge that value. A new challenge arising here, however, is what to do when people hold different or even conflicting values or preferences. Discrete choice experiments could help to test people's preferences regarding specific topics, such as preferred modes of informed consent. Apart from empirical work, conceptual analysis is needed to clarify how public trust, trustworthiness of institutions and accountability are interconnected.

## CONCLUSION

This narrative review suggests widespread—though conditional—support among patients and the public for data sharing for health research. Despite the fact that participants recognise actual or potential benefits of health research, they report a number of significant concerns and related conditions. We believe identified conditions (eg, social value, data security, transparency and accountability) ought to be operationalised in a value-based governance framework that incorporates the diverse patient and public values, needs and interests, and which reflects the way these same conditions are met, to strengthen the social license for Big Data health research.

**Acknowledgements** We thank Susanne Løgstrup (European Heart Network) and Evert-Ben van Veen (Medlaw) for their valuable feedback during various stages in drafting the manuscript.

**Contributors** JvD and GvT conceived the idea for this work. SK designed the study and performed acquisition of the data. SK and GvT analysed the data. All authors contributed substantially to interpretation of the data. SK drafted the manuscript, and JvD, AB, BT, MM and GvT substantively revised it. All authors approved the submitted version of the manuscript.

**Funding** The results presented in this short report were part of Work Package 7 of the BigData@Heart consortium, which received funding from the Innovative Medicines Initiative 2 Joint Undertaking under grant agreement No. 116055. This Joint Undertaking receives support from the European Horizon 2020 research and innovation programme and European Federation of Pharmaceutical Industries and Associations.

**Competing interests** None declared.

**Patient consent for publication** Not required.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Open access** This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

## ORCID iD

Shona Kalkman <http://orcid.org/0000-0002-9051-0156>

## REFERENCES

- 1 Hemingway H, Asselbergs FW, Danesh J, *et al*. Big data from electronic health records for early and late translational cardiovascular research: challenges and potential. *Eur Heart J* 2018;39(16):1481–95.
- 2 Aitken M, Cunningham-Burley S, Pagliari C. Moving from trust to trustworthiness: experiences of public engagement in the Scottish health informatics programme. *Sci Public Policy* 2016;43(5):713–23.

- 3 Taylor M. Information governance as a force for good? lessons to be learnt from Care data. *Scripted* 2014;11(1):1–8.
- 4 General Data Protection Regulation. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing directive 95/46/EC. Available: <https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32016R0679&from=EN#d1e3265-1-1> [Accessed 8 May 2019].
- 5 Phillips M. International data-sharing norms: from the OECD to the general data protection regulation (GDPR). *Hum Genet* 2018;137(8):575–82.
- 6 van Veen E-B. Observational health research in Europe: understanding the general data protection regulation and underlying debate. *Eur J Cancer* 2018;104:70–80.
- 7 Carter P, Laurie GT, Dixon-Woods M. The social licence for research: why care data ran into trouble. *J Med Ethics* 2015;41(5):404–9.
- 8 Dixon-Woods M, Ashcroft RE. Regulation and the social licence for medical research. *Med Health Care Philos* 2008;11(4):381–91.
- 9 Boutilier RG. Frequently asked questions about the social licence to operate. *Impact Assessment and Project Appraisal* 2014;32(4):263–72.
- 10 Allen J, Adams C, Flack F. The role of data custodians in establishing and maintaining social licence for health research. *Bioethics* 2019;33(4):502–10.
- 11 Garrison Nanibaa' A, Sathe NA, Antommara AHM, et al. A systematic literature review of individuals' perspectives on broad consent and data sharing in the United States. *Genet Med* 2016;18(7):663–71.
- 12 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* 2018;3.
- 13 Aitken M, de St Jorre J, Pagliari C, et al. Public responses to the sharing and linkage of health data for research purposes: a systematic review and thematic synthesis of qualitative studies. *BMC Med Ethics* 2016;17(1):73.
- 14 Shabani M, Bezuidenhout L, Borry P. Attitudes of research participants and the general public towards genomic data sharing: a systematic literature review. *Expert Rev Mol Diagn* 2014;14(8):1053–65.
- 15 Howe N, Giles E, Newbury-Birch D, et al. Systematic review of participants' attitudes towards data sharing: a thematic synthesis. *J Health Serv Res Policy* 2018;23(2):123–33.
- 16 Richter G, Borzikowsky C, Lieb W, et al. Patient views on research use of clinical data without consent: legal, but also acceptable? *Eur J Hum Genet* 2019;27(6):841–7.
- 17 Mursaleen LR, Stamford JA, Jones DA, et al. Attitudes towards data collection, ownership and sharing among patients with Parkinson's disease. *J Parkinsons Dis* 2017;7(3):523–31.
- 18 McCormack P, Kole A, Gainotti S, et al. 'You should at least ask'. The expectations, hopes and fears of rare disease patients on large-scale data and biomaterial sharing for genomics research. *Eur J Hum Genet* 2016;24(10):1403–8.
- 19 Darquy S, Moutel G, Lapointe A-S, et al. Patient/family views on data sharing in rare diseases: study in the European LeukoTreat project. *Eur J Hum Genet* 2016;24(3):338–43.
- 20 Mazor KM, Richards A, Gallagher M, et al. Stakeholders' views on data sharing in multicenter studies. *J Comp Eff Res* 2017;6(6):537–47.
- 21 Goodman D, Johnson CO, Bowen D, et al. De-identified genomic data sharing: the research participant perspective. *J Community Genet* 2017;8(3):173–81.
- 22 Spencer K, Sanders C, Whitley EA, et al. Patient perspectives on sharing Anonymized personal health data using a digital system for dynamic consent and research feedback: a qualitative study. *J Med Internet Res* 2016;18(4):e66.
- 23 Patil S, Lu H, Saunders CL, et al. Public preferences for electronic health data storage, access, and sharing - evidence from a pan-European survey. *J Am Med Inform Assoc* 2016;23(6):1096–106.
- 24 McCormick N, Hamilton CB, Koehn CL, et al. Canadians' views on the use of routinely collected data in health research: a patient-oriented cross-sectional survey. *CMAJ Open* 2019;7(2):E203–9.
- 25 Mählmann L, Schee Gen Halfmann S, von Wyl A, et al. Attitudes towards personal genomics and sharing of genetic data among older Swiss adults: a qualitative study. *Public Health Genomics* 2017;20(5):293–306.
- 26 NICE Citizens Council. *What ethical and practical issues need to be considered in the use of Anonymised information derived from personal care records as part of the evaluation of treatments and delivery of care? citizens Council reports No. 18. NICE citizens Council reports*. London: National Institute for Health and Care Excellence (NICE), 2015.
- 27 Hill EM, Turner EL, Martin RM, et al. "Let's get the best quality research we can": public awareness and acceptance of consent to use existing data in health research: a systematic review and qualitative study. *BMC Med Res Methodol* 2013;13(1):72.
- 28 O'Brien EC, Rodriguez AM, Kum H-C, et al. Patient perspectives on the linkage of health data for research: insights from an online patient community questionnaire. *Int J Med Inform* 2019;127:9–17.
- 29 Haga SB, O'Daniel J. Public perspectives regarding data-sharing practices in genomics research. *Public Health Genomics* 2011;14(6):319–24.
- 30 Colombo C, Roberto A, Krleza-Jeric K, et al. Sharing individual participant data from clinical studies: a cross-sectional online survey among Italian patient and citizen groups. *BMJ Open* 2019;9(2):e024863.
- 31 Sanderson SC, Brothers KB, Mercaldo ND, et al. Public attitudes toward consent and data sharing in Biobank research: a large multi-site experimental survey in the US. *Am J Hum Genet* 2017;100(3):414–27.
- 32 Goytia CN, Kastenbaum I, Shelley D, et al. A tale of 2 constituencies: exploring patient and clinician perspectives in the age of big data. *Med Care* 2018;56(Suppl 10 Suppl 1):S64–9.
- 33 Joly Y, Dalpé G, So D, et al. Fair shares and sharing fairly: a survey of public views on open science, informed consent and participatory research in biobanking. *PLoS One* 2015;10(7):e0129893.
- 34 Shah N, Coathup V, Teare H, et al. Motivations for data sharing-views of research participants from four European countries: a direct study. *Eur J Hum Genet* 2019;27(5):721–9.
- 35 Lemke AA, Wolf WA, Hebert-Beirne J, et al. Public and Biobank participant attitudes toward genetic research participation and data sharing. *Public Health Genomics* 2010;13(6):368–77.
- 36 Shah N, Coathup V, Teare H, et al. Sharing data for future research-engaging participants' views about data governance beyond the original project: a direct study. *Genet Med* 2019;21(5):1131–8.
- 37 Taylor MJ, Taylor N. Health research access to personal Confidential data in England and Wales: assessing any gap in public attitude between preferable and acceptable models of consent. *Life Sci Soc Policy* 2014;10(1).
- 38 Kalkman S, Mostert M, Gerlinger C, et al. Responsible data sharing in international health research: a systematic review of principles and norms. *BMC Med Ethics* 2019;20(1):21.
- 39 Cook DJ, Mulrow CD, Haynes RB. Systematic reviews: synthesis of best evidence for clinical decisions. *Ann Intern Med* 1997;126(5):376–80.