npj | digital medicine

Article

Published in partnership with Seoul National University Bundang Hospital

https://doi.org/10.1038/s41746-025-01868-9

Worldwide willingness to share health data high but privacy, consent and transparency paramount, a meta-analysis

Quita Olsen[1], Amalie Dyda[2], Leanna Woods[1], Elton Lobo[1], Rebekah Eden[3], Michelle A. Krahe[4], Bernadette Richards[5], Nalini Pather[5], Lesley McGee, Clair Sullivan[1] & Jason D. Pole[1] ✉

Healthcare delivery is under strain, and the reusing of routinely collected data promises improved outcomes. Still, concerns remain about the public's willingness to share their health data. This study examines worldwide willingness to share health data for secondary purposes. Five electronic databases were searched for eligible studies published since January 2020. Articles were included if they quantitatively examined the primary outcome; the public's willingness to share health data for secondary use, while secondary outcomes included demographic and perception measures associated with willingness to share. Sixty-five articles reported a wide range (24–100%) of public willingness to share resulting in a pooled estimate of 77% (95% CI: 71–82%) among predominantly high-income countries. Participants remain concerned about privacy, consent, and transparency. Future work should consider public education, assessing diverse populations and developing and deploying a validated tool measuring willingness to share data.

Entrusting personal health information to health organisations has broad implications now, more than ever, given health data can be stored digitally, centrally, transferred and repurposed for other uses[1]. The primary purpose of a patient's visit to a clinician or health organisation is to seek improvements to their own health and wellbeing[2]. Clinicians collect relevant information, genomics and biospecimens to aid in delivering health care with this resulting information becoming health data. Beyond its primary function, this health data holds additional value when utilised for secondary purposes[3]. Secondary purposes of health data can be defined as an application beyond the original intent for which it was collected[4], and includes quality improvement and research[5]. These secondary purposes often contribute to healthcare improvements, benefiting both the individual and the overall population[4]. The secondary uses of health data are varied (Table 1).

Willingness to share health data for secondary purposes refers to the attitudes, perceptions and trust that health consumers have towards data sharing[6]. It encompasses concerns, motivations, perceived benefits and, conditions towards sharing[6,7]. Willingness to share health data is a key component of the broader concept of social licence - the dynamic, implicit, and informal set of permissions granted by the public to organisations, reflecting broader social expectations and norms[6,8]. There is an increasing need to understand the publics' willingness to share data as the imperative to conduct large scale analysis for healthcare improvement increases. By developing clear and transparent processes which have social approval, the loss of public trust[6,9–15] and ethical debt; past actions or policies that violate ethical principles and do not consider the longer-term consequences[16], can be avoided. Some consideration exists for the social licence of health data sharing in the literature, but we don't have a collective quantitative understanding of willingness to share or a meta-analysis providing a summary estimate. Therefore, we undertook a systematic review and meta-analysis to address the following question: "What proportion of the global population is willing to share their data for secondary purposes and how does willingness to share differ by population".

Results

Of the 4085 studies identified, 95 met the inclusion criteria, 65 were suitable for quantitative extraction[17–81] and 52 for meta-analysis (provided a suitable quantitative outcome). Figure 1 All 52 studies reported on de-identifiable data except for 1 study that did not make a distinction[30]. Eight studies[28,38,44,46,54,71,77,78] reported on identified data and de-identified data jointly, but only de-identified data was included in willingness to share results. Identified data was broadly defined as data that can reasonably be

[1]Queensland Digital Health Centre, Centre for Health Services Research, The University of Queensland, Brisbane, QLD, Australia. [2]School of Public Health, The University of Queensland, Brisbane, QLD, Australia. [3]School of Business, The University of Queensland, Brisbane, QLD, Australia. [4]College of Medicine and Dentistry, James Cook University, Cairns, QLD, Australia. [5]Academy for Medical Education, Medical School, The University of Queensland, Brisbane, QLD, Australia. Unaffiliated: Lesley McGee ✉e-mail: j.pole@uq.edu.au

Table 1 | Secondary uses of health data

Uses of Health Data	Examples
Clinical settings	Rapid interpretation of complex images, early detection of disease[109] Guide clinical decision making such as machine learning-based predictive models[110]
Safety of medical products after market authorisation	Adverse effects drug monitoring, SENTINEL[111]8, DARWIN[112]
Research	Development of AI and machine learning algorithms[98]
Resource allocation and development	Policy development, health system quality improvement[113]
Public health	Surveillance, evidence-based education[114]
Teaching resource	Fostering evidence-based education[115]
Innovative uses	Combining health data with nanomedicine and targeted drug delivery[116]

linked to or identify a specific individual regardless of the exclusion of personal identifiers like name[82].

Study quality

The MMAT critical appraisal tool for the 65 quantitative studies assessed that 47 studies were of a high methodological quality. They met the screening criteria with clear research questions and appropriate data, as well, as at least three of the five study specific questions. Six studies did not have a clear research question making further assessment difficult[18,24,31,34,48,78]. Most studies had a lower methodological quality when reporting how their sample represented the target population and how they assessed non-responses (Supplementary Table 1).

Study characteristics

The 65 quantitative papers included the perspectives of 141,193 participants from 34 countries[17–81]. Two large studies contributed 46% of participants overall: a UK study of 29,275 participants[44], and a global study on sharing DNA with 36,268 participants[55] (Supplementary Table 2).

Sixty-three studies were cross sectional with one study reporting a RCT[61], and the other reporting pre-post (observational, nonrandomised intervention) analysis[62]. The intervention for both the RCT and pre-post studies consisted of a short education component on willingness to share health data. Of the studies that reported survey dates, data were collected between June 2015 and June 2022, with the exception of one study that gathered data in 2009[80].

Willingness to share

Among 65 studies, 52 (80%) reported the willingness to share in a measure that could be analysed as a proportion. A total of 117,905 participants from 52 studies (including 3 studies that reported two different groups within the same study[19,66,73]) were included in the meta-analysis for willingness to share health data. (Supplementary Table 3 and Supplementary Fig. 1) The pooled random-effects model estimated 77.2% (95% CI: 71–82%) of participants were willing to share their health data. The prediction interval ranged between 27 and 97% reflecting substantial variability in willingness to share across studies. Willingness to share health data in individual studies ranged between 24% in the UK[33], and almost 100% in Norway[30], with 7 articles reporting willingness to share at less than 50%[18,26,33,46,55,68,76]. Most articles (56%) reported over 70% of participants were willing to share their health data for secondary purposes (Table 2).

The studies had very high heterogeneity ($\tau^2 = 1.18$, $I^2 = 99.6\%$, $p < 0.001$) reflecting diverse perspectives on health data sharing across different studies. The LFK index was 3.52, indicating major asymmetry and suggesting publication bias or small-study effects. A sensitivity meta-analysis of willingness to share was performed on studies assessed to have low risk of bias. (Supplementary Table 3) Twenty-four studies, one reporting Swedish and Scottish data separately were included[75]. The random-effects model estimated 78.3% (95% CI: 71–84%) also with high heterogeneity ($\tau^2 = 1.3$, $I^2 = 99.7\%$, $p < 0.001$). The sensitivity analysis aligns closely with the pooled effects from all the random effects models (77.2%, 76.8% and 78.3%) suggesting the overall results are robust.

Unwillingness to share

A total of 33 studies involving 100,911 participants were included in the meta-analysis to assess unwillingness to share health data. (Supplementary Fig. 1) The pooled random-effects model estimated 13.2% of individuals were unwilling to share health data (95% CI: 8.8–19.1%) with high heterogeneity ($\tau^2 = 1.7$, $I^2 = 99.7\%$, $p < 0.001$) and a negative LFK index (−3.43) that skewed towards lower levels of willingness to share. The observed asymmetry was influenced by two outliers, a European discrete choice experiment[25], and white US participants who were mothers with high unwillingness to share (84% and 64%, respectively)[26].

Unwillingness to share ranged between 0.7% for German patients with cancer[51], and 84% across European counties[25]. Most studies (81.3%) reported less than 30% of participants were unwilling to share their health data. One study reported different proportions for populations in Sweden and Scotland and were therefore analysed separately[75], and another reported unwillingness to share for German participants only[66].

The global pooled results of willingness to share (77%) and unwillingness to share (13%) suggest a global population with 10% uncertainty.

Meta-analyses of participant preferences and factors impacting willingness to share

A pooled proportion of participants willing to share their data with different types of organisations and stratified meta-analyses of participants region, country, type of health data, and patient or general public status was undertaken. Significant findings are reported below, with non-significant findings in Supplementary Notes 1 (Region, Country and Type of health data).

Most studies reported willingness to share with research organisations ($n = 38$) with decreased willingness for government ($n = 16$) and particularly for-profit organisations ($n = 16$) using the information for commercial purposes. (defined in Supplementary Table 4) The meta-analyses showed participants were most willing to share for the purposes of research. The pooled random effect model showed the highest proportion for research organisations (80.2%, 95% CI: 74–85%) and the least willingness to share with for-profit organisations who were using the data for commercial purposes (25.4%, 95% CI: 19–33%).

There were a similar number of studies representing patient groups and the general public: in hospital settings or health organisations ($n = 21$), from patient groups ($n = 6$), or the general public ($n = 28$). There were sufficient studies of patients with cancer ($n = 5$) to examine them as a subgroup in the meta-analysis as with patients from non-specific settings and the general public[32,41,42,51,72]. The random effect model estimated patients with cancer had the highest pooled willingness to share (90.9%, 95% CI: 73–97%). Patients in other settings (81.1%, 95% CI: 72–88%) were also higher than the general public (69.7%, 95% CI: 62–77%) with significant difference between the groups ($p = 0.0037$).

Participant characteristics and willingness to share health data

There were significant associations reported between willingness to share and participant characteristics within individual studies (Table 3).

https://doi.org/10.1038/s41746-025-01868-9

Table 2 | Meta-analyses of willingness to share overall and by region, country type of health data and between unwell and health populations

Meta-analysis	N	Proportion	95% CI	95% CI	t^2	*p* value	LFK index	Studies references
Willingness to share (overall)	55	0.77	0.71	0.82	1.18		3.52	17–24,26–35,37,41,42,44–47,49–51,53–55,57,58,60–73,75–77,79–81
Willingness to share (overall using Paule Mandel method)	55	0.77	0.71	0.82	1.20			See above
Willingness to share (sensitivity analysis – high quality)	25	0.78	0.71	0.84	1.34		4.06	17,19,22,27,29,30,32,33,35,37,41,42,44–47,50,51,53–55,57,58,61–65,68,70–73,76,77,79,81
Unwillingness to share (overall)	33	0.13	0.09	0.19	1.67		−3.43	17,24–26,29,31,34,35,37,41,42,44,47,49–51,53–55,57,60–63,65–67,71,72,75,77,79
Willingness to share with type of organisation reusing the data								
Willingness to share for research	38	0.80	0.74	0.85	1.19		6.53	17,20,22,27,28,30–33,37,41,42,44–46,51,53–55,57,58,60–72,76,77,79–81
Willingness to share with government	16	0.70	0.57	0.80	1.12		−5.54	17,20,31–33,41,42,44–46,53,61,69,71,72,77
Willingness to share with for profit organisation for health purposes	18	0.56	0.45	0.66	0.79		4.49	20,27,30,32,41,42,44,45,51,53,55,69–72,74,79,80
Willingness to share with for-profit for commercial purposes	16	0.25	0.19	0.33	0.51		1.62	17,20,30,33,42,44,45,53,61,63,65,66,71,72,74,77
Willingness to share stratified by region								
East Asia and Pacific	15	0.76	0.65	0.84	0.77	0.61	0.46	27–29,35,41,45,46,53,60,67,68,76,77,79
Europe and Central Asia	23	0.81	0.71	0.88	1.65			17,19–22,30,33,44,50,51,57,58,62,63,65,66,69,71,75
North America	13	0.76	0.64	0.84	0.80			26,31,32,34,37,42,47,54,61,64,70,72,81
Willingness to share stratified by country								
Australia	8	0.78	0.61	0.89	0.91	0.56	−0.33	27,28,41,56,67,76,77,79
Canada	4	0.79	0.78	0.81	0.00			31,47,56,72
Germany	5	0.83	0.61	0.94	0.85			51,56,65,66,69
UK	9	0.71	0.54	0.84	0.94			17,20,33,44,49,50,56,71,75
USA	11	0.75	0.61	0.86	0.95			26,32,34,37,42,54,56,61,64,70,81
Willingness to share stratified by type of health data								
Health data	38	0.75	0.67	0.82	1.40	0.09	3.10	17,18,20,24,26–30,32–35,37,41,45–47,49,53,54,57,58,60,63–71,75,76,80
Genomic	9	0.70	0.57	0.80	0.48			19,22,31,55,61,62,77,79
Mental health data	4	0.83	0.72	0.90	0.15			21,44,50,81
Oncological data	3	0.82	0.32	0.98	0.83			42,51,72
Willingness to share stratified by patients or the general public								
Patients with cancer	5	0.91	0.73	0.97	1.03	0.0037	3.46	32,41,42,51,72
Patients in a hospital setting or in treatment	22	0.81	0.72	0.88	1.19			17,19–21,29,30,34,35,37,46,54,57,58,61,63,65–68,71,80,81
The general public	28	0.70	0.62	0.77	0.83			18,19,22,24,26–28,31,33,44,45,47,49,50,53,55,60,62–64,66,69,70,75–77,79

N = number of studies, t^2 = tau squared, LFK index[108].

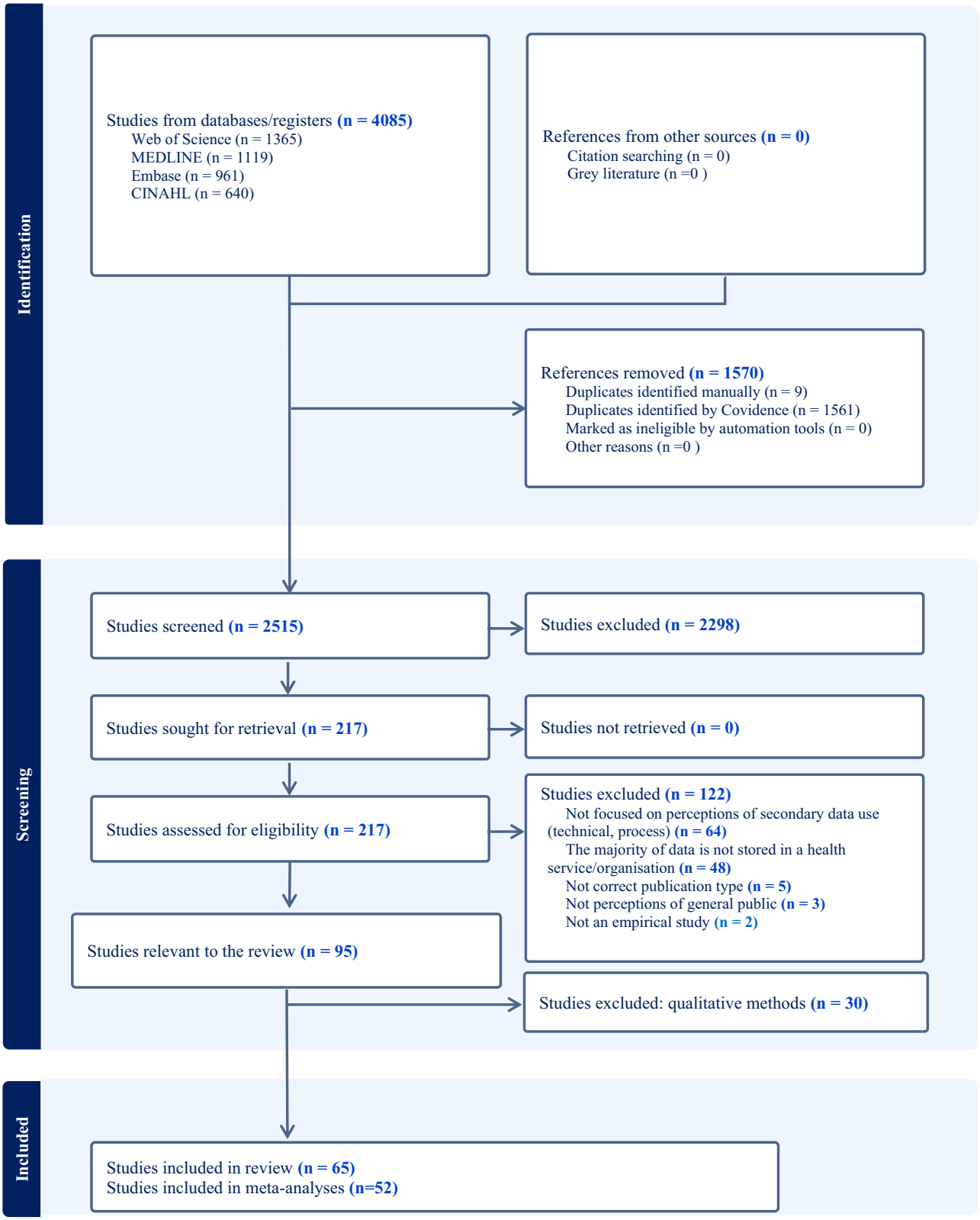

Fig. 1 | PRISMA flow diagram. Included and excluded studies extracted from Web of Science, Medline, Embase and CINAHL on willingness to share with reasons for exclusion. Studies that met the inclusion criteria but were qualitative were excluded.

Twenty-eight studies reported participant ethnicity, and fifteen articles included marginalised populations such as racial and indigenous groups[17,36,42,47,61,74,76,78,79,81], individuals with disabilities and health conditions[19,33,44], sexual minorities[44], and those affected by socio-economic factors[77]. Seven studies reported the percentage of disability and disease in their study participants, ranging from 10 to 69%[18–20,50,52,63,69,79].

Six studies[36,38,42,47,61,76], found a correlation between ethnicity and willingness to share health data with four studies conducted in the USA[36,38,42,61].

Table 3 | Participant characteristics

Participant characteristics with significant associations for willingness to share				
Participant characteristics	**Study location (References)**	***N***	**Sample size for participants *n* (%)**	**Proportion (reported first)[a] Significant Association (with association referent)[b]**
Ethnicity				
Non-Hispanic White	USA[42] USA[61]	2	554 (84.4) 66 (82.5)	95%[42 a] Mean (SD) 2.88 (0.92) (Other 2.15 (0.98))[61 b]
Other/Mixed Race/Ethnicity	USA[42] Canada[47]	2	19 (2.9) 441 (8.0)	88%[42 a] OR 1.57 (North American)[47 b]
Hispanic	USA[36]	1	833 (23.3)	*β* Co-efficient 0.12 (non-Hispanic)[36 b]
Asian	USA[38]	1	62 (9.3)	*β* Co-efficient 3.12 (White 2.78)[38 b]
East and Southeast Asian	Canada[47]	1	361 (7.3)	OR 0.69 (North American)[47 b]
South Asian	Canada[47]	1	151 (3.0)	OR 0.66 (North American)[47 b]
Aboriginal	Australia[76]	1	30 (4.3)	86%[76 a]
Sex				
Females	UK[20] USA[32] Jordan[18] Canada[31] USA[38] Canada[72]	6	180 (58.4) 546 (81.9) 521 (50.4) 517 (74.5) 323 (49.7) 92 (50)	91.6%[20 a] 76.5%[32 a] *β* Co-efficient 1.43 more hesitant (than males)[18 b] mean = 3.6, SD = 1.2 (Male mean = 3.3, SD = 1.3)[31 b] *β* Co-efficient 2.66 (Male 2.98)[38 b] 20% share specific information (Males 7%)[72 a]
Males	UK[20] USA[32]	2	121 (39.3) 121 (18.1)	97.5%[20 a] 69.5%[32 a]
Age				
Older cohort	Singapore[53] Jordan[18] Canada[31] USA[36] UK[44] Canada[47] Canada[72]	7	269 (26.9) 160 (15.5) Age mean 45 1001 (28.3) 1274 (36) – 4% 1584 (31.8) 162 (89)	60.3%[53 a] *β* Co-efficient > 35 years 0.69 less hesitant (than < 24yo)[18 b] Less comfortable with release of data to databases (r(678) = 0.47, *p* < 0.01)[31 b] *β* Co-efficient −0.16 (45–59 years), −0.17 (≥60 years), (ref 18–29 years)[36 b] Increased +75[44] OR 1.46–3.36 in 60 to ≥80 years (ref 18–29 years)[47 b] 2% (uncomfortable)(>50 years)[72 a]
Younger cohort	Singapore[53] UK[44] Canada[72]	3	337 (33.7) – 24% 19 (10)	71.8%[53 a] Increased 25–44[44] 13% (uncomfortable)(≤49 years)[72 a]
Private companies and age	Australia[27]	2	460 (18.1) 552 (21.80)	Older > 65 (60.2–70.1%)[27 a] 49.2–56.4% (<29 years)[27 a]
	USA[74]		554 (30.9) 483 (26.2)	*β* Co-efficient −0.104 (30–44 years) −0.154 (45–59 years)(ref 18–29 years)[74 b] commercial
Education				
Higher education	Switzerland[63] Australia[76] Singapore[53] USA[38] Australia[41] Jordan[49] EU[58] USA[74]	8	543 (54) – 51% 364 (36.4) 127 (20) 41 (31.1) 1004 (84.1) Unclear 610 (33.1)	77%[63 a] 76%[76 a] 69.9%[53 a] 57%[38 a b] Graduate school *β* Co-efficient 0.51 bachelor's degree (ref less than bachelor's degree)[41 b] OR 0.299 (High school or lower)[49 b] Spearman rank (*ρ* = 0.096, *n* = 962, *P* = 0.003)[58 b] *β* Co-efficient Bachelors or above 0.197, (less than high school)[74 b] commercial
Lower education	Australia[76] Singapore[53] Switzerland[63] USA[38]	4	463 (46) Unclear 45% 129 (20)	71%[76 a] 60.5%[53 a] 53%[63 a] 41%[38 a] High school
Motivation	Portugal[19]	1	148 (31.2)	OR 2.2; Carers with more than 12 years education select discovery of a cure[19 b]
Governance	Australia[27]	1	1708 (100)	82%[27 a] Ethics Committee oversight (81.2–87.5%)[27]
Existing illness				
Chronic disease	Switzerland[63] Switzerland[22]	2	225 (18) 110 (26.4)	81%[63 a] 51.7%[22 a]
Cancer	Germany[69]	1	791 (68.7)	β Co-efficient 0.271, *p* = 0.019 (no cancer exposure)[69 b]
Health status	USA[26]	1	594 (95.5)	OR 0.38 Child is excellent to good (Poor to Fair)[26 b]

Participant characteristics of studies reporting proportions and significant associations with willingness to share: ethnicity, marginalisation, sex, age, education, employment, existing illness and location. Higher number of studies and proportions reported first.

N number of studies, *OR* odds ratio, *β Co-efficient* beta co-efficient, *r* correlation coefficient.

[a]Proportions of participants characteristic who were willing to share (reported first).

[b]Significant associations ($p < 0.05$) with participant characteristics that were not reported as a proportion or able to be converted to a proportion ie. OR, β with the characteristic comparison in brackets

The highest proportion of participants willing to share were non-Hispanic white participants at 95% in the USA[42]. In an Australian wide study, the Aboriginal and Torres Strait Islanders sub-population were the most willing to share their health data (86%), however, there was limited sample ($n = 30$) and further reporting in the same study provided conflicting results[76]. There was limited ethnic comparisons reported.

Sex and age characteristics were included in 59 studies, with different age categories collected for most, the youngest participants were aged 12 years and the oldest over 90 years.

Females' willingness to share ranged between 77 and 92% and males' willingness ranged between 70 and 98%[20,32]. Comparisons between females and males regarding their willingness to share health data revealed mixed results across several studies. Two studies reported significance ($p < 0.05$) with males more likely to share[18,38], and two where females were more likely[31,72].

There was variability in the willingness to share health data among different age groups, with inconsistent age categories reported across the studies. Nine studies found significant differences between older and younger participants[18,27,31,32,44,47,53,72,74]. Greater willingness to share among younger participants was reported in some studies[18,32,53,74], and other studies found that older participants were more open to sharing their data[27,47].

Ten studies found a correlation between education levels and willingness to share health data. Most studies reported a positive relationship between higher education and greater willingness to share, but some studies reported the opposite[41,49]. Two studies reported higher education, motivated willingness to share with finding a cure[19], and increased governance expectations[27]. The governance expectations were ethics committee oversight, (82%) transparency about how their data was being used, (81%) knowledge about who was using the data, (81%) and publication of research findings (82%)[27]. In contrast, participants with an education level of year 10 or less had lower preferences for these conditions (ranging from 42% to 54%)[27].

Key factors for sharing health data

Studies identified several key factors influencing participants' willingness to share their health data. These were grouped into Preferences, Conditions and Attributes for sharing. Studies reported participants had preferences for consent ($n = 23$), privacy ($n = 25$), knowledge ($n = 4$) and regional sharing ($n = 9$). (Fig. 2) (Supplementary Table 5) Preferences are the choices that participants wish to be respected. When reporting on preferences for consent, privacy, knowledge and regional sharing, we report the proportion (or significant association) of agreement with willingness to share. Consent refers to permission participants give to sharing their data and knowledge relates to how informed participants were about sharing. Conditions ($n = 24$) report the preferences that must be implemented before participants are willing to share their health data, for instance, data security. (Supplementary Table 6) Attributes included, concerns ($n = 25$), motivations ($n = 30$) and trust ($n = 13$) and report how important these attributes are when participants are considering sharing their health data. (Supplementary Table 7) As an example, data security would have a level of importance attributed to it by participants under privacy, would have to be implemented under conditions and would be a source of anxiety for that proportion of participants under concerns. These findings adhere to how the studies have reported participant preferences for willingness to share health data as measures of importance, conditions or attributes.

Participant preferences for consent with willingness to share health data differed between and within countries. Participants preferred data use with consent[24,25,59], than without consent[24,40,51,80], where only one study had more than 50% of participants preferring data use by default[51]. Broad consent[30,37,51,57,72], was preferred by more than 50% of participants in comparison to specific consent[31,37,66]. Broad consent is where permission to share data could be granted for multiple studies and specific consent is where each research project requiring participant data would need to have permission every time. Another two studies reported specific consent was not important[58,66]. Opt-in preferences were higher (36–91%)[27,37,67], than opt-out preferences[27,37,43,56,57,67,80,81], with only one study reporting participant opt-out preferences over 50%[37]. Opt-in consent requires participants to actively give permission, while opt-out includes participants unless they explicitly decline[12]. Different studies were interested in different aspects of consent with mixed results, for instance, within Australia participants expressed different preferences across three studies[27,56,67].

Two studies reported preferences (20–73%) for a waiver of consent with Human Research Ethics Committees either notifying participants[67,79], or broad consent applying[67]. One study reported 54% of participants wished to be able to withdraw consent[56]. Another study of the general public in Sweden reported waiver of consent for faster medical progress in a risk benefit analysis[24].

Of the fifteen studies reporting de-identified or anonymised data sharing preferences, most studies ($n = 14$) reported more than 50% of participants preferred their data to be de-identified or anonymous. All studies preferred de-identification over identification. Deidentified data was preferred in Korea (50.17%)[46], Australia (78%)[41], Hong Kong (72.6%)[60], Japan (59%)[80], UK (66.7%)[74], the USA[78], and Japan 59%[80]. Six studies[46,71,78] reported preferences for sharing identified data that was context specific, with three studies reporting participants were more than 50% willing to share identified data[28,44,77]. Two studies showed a significant increase in willingness to share as privacy protections increased[28,39]. Eight European countries willingness to share were also significantly associated with privacy: preferring anonymity (32.9%) over pseudonymisation (28.2%)[58]. German participants with higher privacy concerns were less willing to share[69] and another German study found that 69.9% of participants would share data if social security and telephone details were removed[51]. Nearly half of UK participants were willing to share often if data was anonymised (49.2%)[71].

Existing knowledge of health data sharing for research was reported in three studies. They found that 36% of mental health patients[21], 37% cancer patients[42], and 52% emergency presentations did not know that the data may be used for research[67]. Where participants had a higher understanding about the use of their genetic data they perceived greater benefits to sharing[22].

The relationship between geographical distance from participants' region and willingness to share health data were explored in nine studies. Most articles reported a preference for sharing health data locally and found that willingness to share decreased as geographical distance increased[29,44,65,72,77]. There was wide variation in the level of willingness to share locally ($n = 6$, 18–84%), nationally ($n = 2$, 65–81%) and internationally ($n = 5$, 19–96%).

Twenty studies specified conditions for the sharing of health data by participants. Conditions of sharing that were reported in the studies were access to information[30,32,42,43,51,54,56–58,72], penalties for misuse[45,56,57,60,80], privacy protections[36,42,63,78], ethical oversight[31,42,54], control[40,45,58], security[42,51,63,65], and the types of researchers who could access the data[77]. (Supplementary Table 6) Participants conditions for information and transparency on data sharing ranged between 25 and 84% with most studies reporting more than 50% of participants required information[30,42,57,58,72]. There was strong support (40–93%) for issuing sanctions or penalties for data misuse[56,57,60,80]. Participants wanted conditions for oversight and governance (70–94%)[31,42,54], such as a Human Research Ethics Board and a review process (82–83%)[80]. There were high expectations for control of health data as a condition for participants who wanted to decide and control access to their data (51–77%)[40,45,58]. Security and legal protection conditions ranged between 8 and 58%[51,63], with one study[65] reporting 72% of participants required a nationwide data base, and another study where 88% required a privacy officer[31]. In a 22-country study 77.5% of participants also wanted to be able to delete their data[58], access their own DNA and medical data (38%) and to be able to withdraw (54%)[56].

Study participants held concerns over many aspects of sharing health data. The most reported concern was re-identification with 31–83% of participants concerned there may be enough information that a person can be re-identified after sharing[17,31,33,41,42,45,49,52,63]. When data could be used to identify participants 20–70% of participants were concerned about

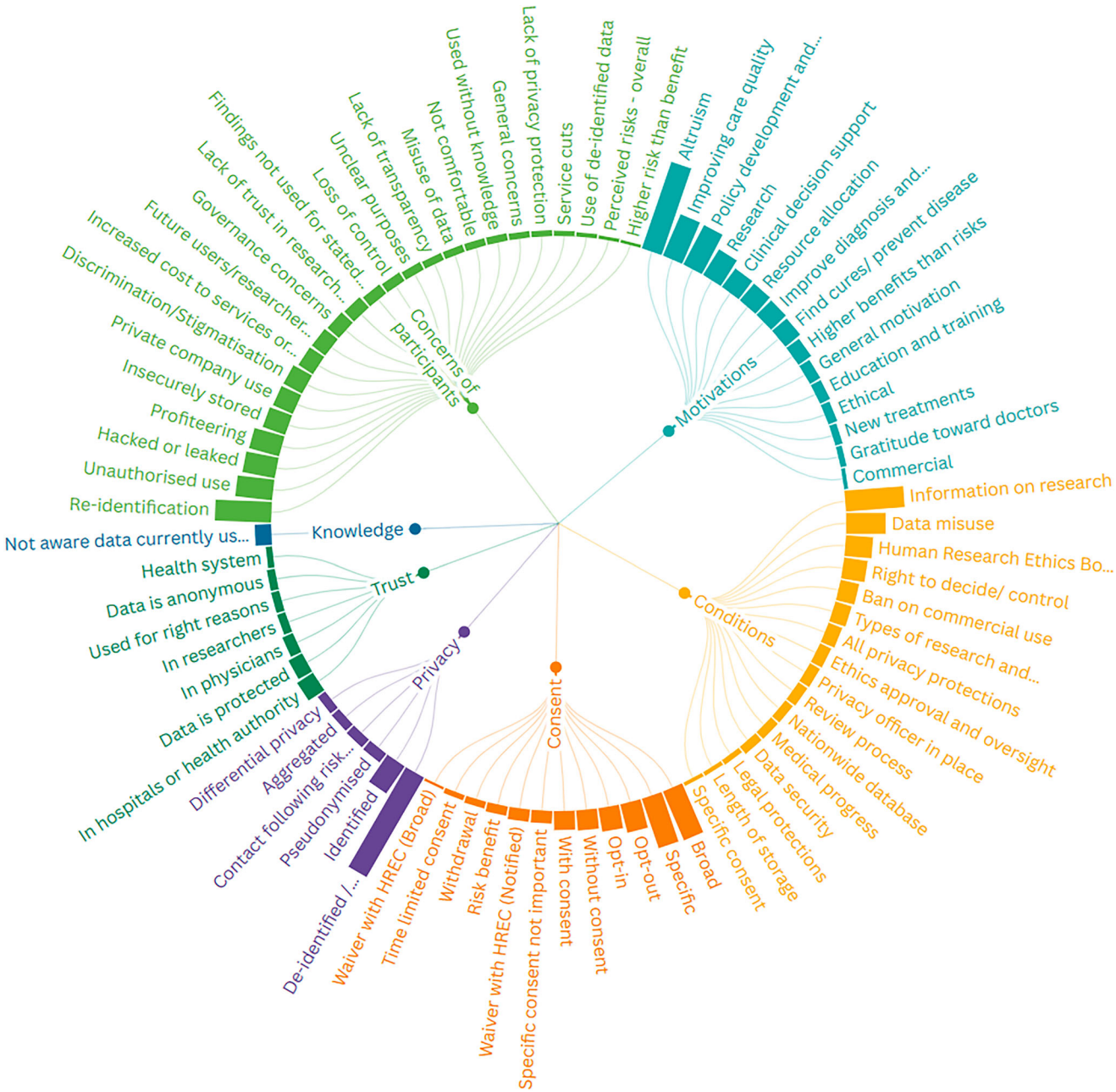

Fig. 2 | Radial tree of participant preferences, conditions and attitudes. Participants proportional preferences for sharing health data mapped to Concerns, Motivations, Conditions, Consent, Privacy, Trust and Knowledge. For an interactive visual map https://public.flourish.studio/visualisation/21200433/.

discrimination or stigmatisation[19,41,49,77]. Participants were highly concerned about aspects of data security, with unauthorised use (27–64%)[22,23,41,42,52], being hacked or leaked (51–75%)[41,45,49,52], insecurely stored (53–69%)[33,49,65], and misuse of data (58%) reported[49].

Participants across seven studies expressed concerns for profiteering or sharing of health data with private companies with concern levels ranging between 27% to 90%[27,33,63,65,75,77,79]. Participants were concerned that sharing data could lead to increased costs or service cuts[22,33,63]. Studies reported concerns in trusting researchers and research processes[22,41,45,49,52,53,63,71,74,75]. The highest level of concern (65%) was for governance and legislating against data misuse[71], with more than 50% of participants concerned about privacy protection[63], transparency[49], use without knowledge[52], unclear research purposes[49,75], and findings from secondary studies not being used for the stated research purposes[22,41]. Participants were concerned about a lack of control and in one study 60% of the participants wanted control over their health data all the time[57].

Motivations for sharing health data were high with more than 58% of participants in studies sharing for altruistic purposes[33,35,41,49,60,63,65,66,71], or were generally motivated[22,40,48,64,68]. Participants were strongly motivated by improvements to the health system (48–93%)[41,74], but less motivated for commercial organisations (15–16%)[33,46]. Some studies reported self-benefit as a motivation for sharing health data[64], such as Finnish participants being highly motivated to gain benefits following consent to use health data for sharing (89%)[62].

Trust was reported in 12 studies. Trust in researchers[51,59,69], was lower in one study (65%) than trust in hospitals and health authorities (71–72%)[21,60,69] but higher than trust in private companies (54%)[60,69]. Participants in three studies were more willing to share their records when they trusted researchers ($p < 0.001$)[40], health authorities ($p < 0.001$)[21], and university researchers (63%)[22]. A loss of trust in research institutions was reported in three studies[49,53,74], for instance, with 35% of Singaporean participants not trusting research processes and protections[53].

Discussion

The advent of big data, including AI, in healthcare is forcing a rapid assessment of our social licence to repurpose routinely collected healthcare data for secondary purposes. This study provides a quantitative and granular view of willingness to share health data. This is the first known review and meta-analysis for willingness to share health data for secondary purposes providing a proportional assessment of global willingness and unwillingness to share.

There is a willingness to share health data, in predominately high-income countries. Data sharing is dependent on context and in these populations motivations for sharing data were primarily for the public good and allowing access to health data to facilitate research for disease or for health organisation benefits. Conversely, only 13% of participants were unwilling to share health data.

Ten percent of participants were unaccounted for in willingness to share or unwillingness to share in the pooled results. While 22 studies explored concerns, barriers and risks of participants, only one article specifically reported unwillingness to share[25]. Understanding public perspectives may shed light on potential barriers to health data sharing and help identify ways to address misinformation.

The meta-analysis revealed a clear preference for sharing health data with research organisations over for-profit organisations, even when the purposes of sharing are for health outcomes. This reflects the general public's concerns about exploitation of health data, loss of control, potential misuse and increased costs associated with for-profit organisations. The literature reflects these findings, recognising there is an uneasy tension between organisations with the resources to develop shared data, and the responsibility to return those benefits to the public sphere[13,14,83–85]. To address the public's reservation with sharing health data with for-profit organisations, robust frameworks and governance aligned with the publics expectations is essential. The risk of not clearly managing commercial sharing is the loss of trust in sharing with publicly acceptable entities, such as research organisations.

The health status of patients also impacts willingness to share, with those who are unwell, particularly those with cancer, more likely to share their data. Stratifying willingness to share by health status showed significant differences between patients with cancer, patients and the general public. However, few individual studies compared unwell and healthy populations ($n = 3$) and given the large number of studies from unwell populations ($n = 27$) there may be bias towards the positive given unwell populations are more willing to provide their data[63]. Kalkman et al.[6] provided a narrative review that explored the difference between patient's and the general public willingness to share, and more recent reviews have conflicting results or focus on one group only[83,84].

Stratification by type of health data or country did not yield significant differences in willingness to share, although further studies may be required to confirm these findings and improve heterogeneity. Within individual studies, the type of health data significantly influenced willingness to share with mental health and genetic data perceived as more sensitive than general health data and biospecimens. These results align with Cascini et al., who reported similar differences in health, biobank and genomic data[13]. Other reviews highlight the sensitivity of different types of health data, although, none as a quantitative comparison.

Although willingness to share health data is generally high, participants' concerns and conditions for sharing health data have remained consistently elevated across various studies. This indicates a potential disconnect between the public's knowledge or perception of data sharing principles and their practical application, likely because consent for sharing de-identified data often occurs at a system level. Very few studies ($n = 3$) examined participants' existing understanding and knowledge of data health sharing, with findings indicating a substantial proportion of participants (36–52%) were unaware of their jurisdictions' current data sharing practices. Limited research has explored the relationship between education and willingness to share with one showing increased willingness[62], and the other with unclear results[61]. Other disconnects between sharing practices[86] and public perceptions are the publics strong preferences for privacy protections and consent, suggesting that organisations may not be effectively communicating sharing practices to address public apprehension[13,83,84].

High levels of concern impact the conditions participants require to share their health data. Most participants required more information and control. In a US study, respondents were vocal about sanctions being applied to misuse of health data and wanted to know about the type of research and researchers accessing the data[87]. This lack of public knowledge is also evident in the broader literature. Botkin et al. found almost all participants thought biospecimens were destroyed and electronic medical records were not accessed by researchers[88]. Two recent reviews reported low public awareness across the studies they reviewed[1,83]. A lack of understanding about health data sharing practice may erode trust in health systems and organisations, exemplified in the care.data controversy in the UK, which was disruptive to patients, providers and resulted in legislative change. Care.data was a UK government initiative in 2013 that intended to upload general practitioner health data to a central database with opt-out consent only. The data could be shared with government, academic research institutions and for-profit organisations. The intentions of the initiative were poorly communicated, resulting in doctors and patients threatening non-participation, ultimately resulting in a significant loss of trust between clinicians, the public and the NHS[89]. Given there is limited understanding and reporting of existing health data sharing in the countries included in this analysis, it remains imperative to improve education and ensure transparency in how health data is reused.

Many studies ($n = 43\%$) captured ethnicity as a demographic characteristic but few studies ($n = 6$) compared willingness to share within their ethnic groups and most ($n = 66\%$) were in the USA. This was particularly evident with indigenous groups who were not reported in the US studies and severely underreported in the Australian and Canadian studies. Although only one study[44], captured 6% of LGBTQI+ participants without reporting their views, it is essential to represent their perspectives on health data sharing, given their specific health needs and concerns, such as those related to AIDS, mpox, and transgender health[90]. Other marginalised groups such as culturally and linguistically diverse[91], homeless[92], migrant populations[93], and domestic violence victims[93], were not represented. Given these groups likely have unmet health and security needs around sharing health data, representing their voice in the literature is important[94].

Cumyn et al. provided responses from indigenous communities about transparency and information for the secondary use of health data but were unable to say they were representative of indigenous perspectives and indigenous data governance[85]. Marginalised populations have high levels of concern about sharing health data and are often underrepresented in studies. (90) This issue can be addressed by fostering partnerships between researchers, policy makers, and community leaders to better represent these populations and develop more inclusive data sharing policies.

The exploration of different types of consent and privacy within the studies represents variations to informed consent, anonymity and the right to withdrawal enshrined in ethics, legislation and enacted by ethical review boards[86]. Participants across studies were content with informed consent encompassing broader interpretations, with the potential of health data sharing applied to multiple research projects over varying timelines. Other studies have found similar findings[95,96]. Hutchings et al. recommended consent could be dynamic, with a single consent model not suitable for all data sharing studies[14]. Participants' preferences for anonymity as reflected in the privacy results were less flexible with most studies preferring de-identified data to be shared. In the studies exploring identified data[28,38,44,45,54,71,77,78], reasons for removing privacy protections were not always clear or how sharing identified data would be of benefit[54].

There are some limitations to the generalisability of our findings. There is a lack of representation for non-Western countries and regions which is reflective of the broader literature, Kalkman et al. observed global disease burden was probably not captured in "big data" or large data sets[6]. Nearly all included studies were cross-sectional, making them suitable for assessing perceptions, although this study design is susceptible to bias[97]. Additionally,

Table 4 | Table of inclusion and exclusion criteria

Inclusion criteria	Exclusion criteria
The perceptions of the general public or health consumers only	Health care professionals, commercial health organisation representatives
All ages and demographics	
Health data stored in a health organisation with an implied data custodian and generated by healthcare providers (e.g. personal health data, biospecimen or genomic data)	Commercial organisation collecting health data such as wearables (e.g. step counter)
Health data that is used for a secondary purpose (e.g. research, data mining, quality improvement or AI training and development)	Health data used for purpose with which it was collected (e.g. patient treatment)
For multiple health data repositories, the majority of data had to be held in a health facility	If the majority of data is stored outside a health organisation (e.g. social media or with wearables)
Primary research studies with qualitative, quantitative or mixed methods	Clinical trials as consent to share had previously been obtained
Full text studies in English, published in peer reviewed journals Jan 2020 to Dec 2023	Reviews, non-empirical studies, editorials, commentaries, grey literature, theses, conference abstracts, protocols and posters

high heterogeneity was observed, which was expected given the diverse populations and conditions represented in the studies. However, some heterogeneity might also stem from differences in study methodology and lack of a standardised measure of willingness to share.

Our review encompassed the time period of a global pandemic[34,38,44,47,49,62,69–71,73] and the public's growing awareness of AI[17,64,70,81,98], which may have had significant impacts on public perceptions and consequently, individuals' willingness to share health data. These exogenous events both required extensive data sharing and collaborative research and was an important period to measure attitudes towards data health sharing[99,100].

Reassuringly, the proportion of the global population that is willing to share their data for secondary purposes in jurisdictions that are grappling with big data and AI implications is high. Despite highly publicised data breaches such as Cambridge Analytica[101], health data for secondary purposes seems to enjoy a protected status[102], perhaps, in part, driven by the public's collective goodwill and altruism. When health data decision makers are considering sharing, our review shows the public is generally positive, wish to stay informed, and appear to be supportive of existing protections if they are transparent. When researchers are considering using participants health data they should design studies with participant preferences at the forefront that are transparent, and consider open public communication of results. Co-design with consumer collaboration, that may include pre-survey study preferences such as opt-in preferences could be some of the strategies considered[103].

This review synthesises diverse data on willingness to share health data, particularly during this current time of heightened public awareness. Findings underscore the need for public education on data-sharing benefits to build trust and suggest that future research focus on developing standardised measures and addressing underrepresented populations. Expanding studies to low- and middle-income countries, as well as incorporating marginalised perspectives, will support more inclusive data-sharing policies.

Methods

The systematic review was undertaken following the PRISMA for systematic reviews statement reporting guidelines (2020)[104] and developed under the PRISMA 2020 checklist (Supplementary Table 8). The protocol was registered in PROSPERO (CRD42024504135). The quality of the quantitative papers were such that a meta-analysis was able to be undertaken and this review concentrated on the quantitative papers only.

Information sources

Four databases were searched: Medline (Ovid), Cumulative Index to Nursing and Allied Health Literature (CINAHL), Web of Science and Scopus on 29th Jan 2024. Where article information was unclear, or incomplete study authors were contacted by email (to date no author provided further information when contacted).

Search strategy

The search terms were developed collaboratively by the authors and an academic librarian. Keywords were developed around the concepts of health or medical records, secondary use, attitudes and perceptions, with the full search strategy for all databases provided (Supplementary Notes 2). Social licence concepts included trust, perception, attitudes and their synonyms. Truncation and proximity operators were used to ensure sufficient article coverage. Database filters applied to the search terms included publication dates from Jan 2020 onward, with exclusions for conference abstracts, reviews, case reports, protocols or proceedings.

Eligibility criteria

Studies included in the review focused on the perceptions from the general public about health data repurposed for secondary use and stored within health organisations. (Table 4) Health organisations were defined as organisations that deliver, manage and support the provision of health care collecting patient data for that purpose. The time period was selected to provide a contemporary assessment and account for possible exogenous influences on willingness to share health data.

This review was scoped to quantitative outcomes only with quantitative studies or mixed method studies reporting a proportional, granular perspective of the global publics' willingness to share. Qualitative studies providing rich, nuanced public perspectives were reported elsewhere. The primary outcome was the willingness to share health data for secondary purposes. This outcome was measured in quantitative terms and could include either proportional or custom measures, such as beta-coefficients. If studies provided secondary quantitative associations to the primary outcome, but not a primary quantitative measurement of willingness to share, they were included because they met the inclusion criteria for the review and reported a willingness to share quantitatively. Secondary outcomes included perspectives, opinions and attitudes of the general public for sharing health data. If a comparison group was reported in a primary study, it was noted.

Selection process

Identified articles were uploaded into Covidence[105] a scientific literature tool. Evaluation of papers against the inclusion and exclusion criteria was performed by two independent reviewers in a pool of seven reviewers at selection of title/abstract and again at full text screening. Reasons for exclusion at full text screening were recorded.

Data collection process

Data extraction was undertaken by one primary reviewer using a standardised form (Supplementary Table 4) that had been developed and pilot tested by the research team. A second reviewer then verified and corrected the extracted data for accuracy. Mixed method extraction for quantitative data was the same. Quantitative articles that reported the primary outcome numerically were selected for meta-analysis. Where articles contained unclear information such as demographic data, the authors were contacted

for clarification. Regardless of the author's responses, all articles were included in the review.

Data items

The primary outcome was the proportion of the general public's willingness to share health data. Data extraction for the primary outcome aligned with the closest ethical benchmarks, such as consent, and privacy. Parameters to the reported willingness to share, such as if the data was de-identified, were captured, when reported. Significant associations with willingness to share, like ethnicity, were captured with measure of effect and significance included. Strong proportional associations were also reported. Unwillingness to share was collected to delineate participants who were willing to share from those who were undecided. Likert scales responses indicating agreement or comfort with data sharing were combined for an overall willingness proportion (Supplementary Table 3).

Secondary associations linked to willingness to share, such as demographic factors, perceptions, organisations receiving data, data type and barriers or motivations were documented. Health data type was categorised into four types: general health data, genomic, oncological, and mental health data. Organisations receiving the data for secondary purposes were organised to research, government, for-profit with health purposes and for-profit for commercial purposes, if reported. Demographic factors, perceptions, organisation type, type of data, and barriers and motivations associated with willingness to share were documented, if reported. Article detail (funding, country, setting, study type) and study methods were extracted along with demographic information (age, gender, education, employment, ethnicity, patient and marginalised group status). Country income levels were assessed using World Bank gross national income (GNI) per capital categories[106]. Study dates and cultural/historical context were captured when reported.

Study risk of bias assessment

All studies were assessed for risk of bias using the Mixed Methods Appraisal Tools (MMAT) by two independent reviewers (Supplementary Table 1)[107]. Differences were resolved by author consensus. The MMAT was scored a 1 if both reviewers reached consensus for each question. If consensus was not reached a response of "No" or "Can't tell" contributed a score of 0 for that question. The maximum score for the first two screening questions was 2 and the maximum score for the study design specific scores was 5.

Effect measures

The primary outcome was reported as a percentage of those willing to share, hence a proportion meta-analysis was undertaken on studies that reported such measures or could be converted to a proportion. Secondary outcomes of unwillingness to share were reported in a proportional meta-analysis and some study characteristics such as type of health data was reported as a stratified meta-analysis.

Synthesis methods

The included articles were tabulated and included willingness to share, sample size and details specific to the participants, and significant secondary outcomes. (Supplementary Tables 2 and 3) The preferences for data sharing were visualised to a radial tree. The primary outcome, willingness to share was fitted to a random effects model in R for windows (Version 4.4.1 and RStudio 2024.09.0 + 375 "Cranberry Hibiscus" 2024-09-16) and heterogeneity reported. The proportions were logit transformed (some proportions were close to 1) to stabilise variance, with a maximum likelihood (ML) estimate of between study variance and Hartung-Knapp adjustments. A further meta-analysis with Paule-Mandel estimator was performed to test robustness of the model. A sensitivity meta-analysis with lower risk of bias studies was also undertaken. It was expected that there was high heterogeneity as perceptions of willingness to share differed according to the type of data being shared and the governance of the participants locality. Heterogeneity was reported as a I^2 statistic and tested with Cochran's Q test, to quantify variability. Missing outcomes, both secondary and primary were excluded from the meta-analysis but included in the synthesis where they reported an association with willingness to share.

Reporting bias assessment

The publication bias in a prevalence meta-analysis requires specific statistical models. The DOI plot and LFK Index were used as measures for small study bias and the LFK index reported[108].

Data availability

The data relied on for the meta-analyses is in Supplementary Table 3 and in Supplementary Tables 7–9 for participant preferences, conditions and concerns. All code utilised in this study is available on request from the authors. Any further data utilised in this study is available on request from the authors.

Code availability

The underlying code in this study is available on request from the authors. Software utilised for this study was R for Windows (Version 4.4.1 and RStudio 2024.09.0 + 375 "Cranberry Hibiscus" 2024-09-16).

Received: 5 February 2025; Accepted: 6 July 2025;
Published online: 23 August 2025

References

1. Skovgaard, L. L., Wadmann, S. & Hoeyer, K. A review of attitudes towards the reuse of health data among people in the European Union: The primacy of purpose and the common good. *Health Policy* **123**, 564–571 (2019).
2. National Academies of Sciences, Engineering, and Medicine; Health and Medicine Division; Board on Health Care Services; Committee on Health Care Utilization and Adults with Disabilities. *Factors That Affect Health-Care Utilization: Health-Care Utilization as a Proxy in Disability Determination* (National Academies Press (US), 2018).
3. Mascalzoni, D. et al. International Charter of principles for sharing bio-specimens and data. *Eur. J. Hum. Genet.* **23**, 721–728 (2015).
4. Safran, C. et al. Toward a national framework for the secondary use of health data: an American Medical Informatics Association White Paper. *J. Am. Med. Inf. Assoc.* **14**, 1–9 (2007).
5. Murdoch, T. B. & Detsky, A. S. The inevitable application of big data to health care. *JAMA* **309**, 1351–1352 (2013).
6. Kalkman, S. et al. Patients' and public views and attitudes towards the sharing of health data for research: a narrative review of the empirical evidence. *J. Med. Ethics* **48**, 3–13 (2022).
7. van Panhuis, W. G. et al. A systematic review of barriers to data sharing in public health. *BMC Public Health* **14**, 1144 (2014).
8. Muller, S. H. A., Kalkman, S., van Thiel, G. J. M. W., Mostert, M. & van Delden, J. J. M. The social licence for data-intensive health research: towards co-creation, public value and trust. *BMC Med. Ethics* **22**, 110 (2021).
9. Nijhawan, L. P. et al. Informed consent: issues and challenges. *J. Adv. Pharm. Technol. Res.* **4**, 134–140 (2013).
10. Lowrance, W. Learning from experience: privacy and the secondary use of data in health research. *J. Health Serv. Res. Policy* **8**, 2–7 (2003).
11. Mikkelsen, R. B., Gjerris, M., Waldemar, G. & Sandøe, P. Broad consent for biobanks is best—provided it is also deep. *BMC Med. Ethics* **20**, 71 (2019).
12. de Man, Y. et al. Opt-in and opt-out consent procedures for the reuse of routinely recorded health data in scientific research and their consequences for consent rate and consent bias: systematic review. *J. Med. Internet Res.* **25**, e42131 (2023).
13. Cascini, F. et al. Health data sharing attitudes towards primary and secondary use of data: a systematic review. *EClinicalMedicine* **71**, 102551 (2024).

14. Hutchings, E., Loomes, M., Butow, P. & Boyle, F. M. A systematic literature review of attitudes towards secondary use and sharing of health administrative and clinical trial data: a focus on consent. *Syst. Rev.* **10**, 132 (2021).
15. Beets, B., Newman, T. P., Howell, E. L., Bao, L. & Yang, S. Surveying public perceptions of artificial intelligence in health care in the United States: systematic review. *J. Med. Internet Res.* **25**, e40337 (2023).
16. Richards, B. & Scheibner, J. Health technology and big data: social licence, trust and the law. *J. Law Med.* **29**, 388–399 (2022).
17. Aggarwal, R., Farag, S., Martin, G., Ashrafian, H. & Darzi, A. Patient perceptions on data sharing and applying artificial intelligence to health care data: cross-sectional survey. *J. Med. Internet Res.* **23**, e26162 (2021).
18. Al-Shami, K. M., Ahmed, W. S. & Alzoubi, K. H. Attitudes toward providing open access for use of biospecimens and health records: a cross-sectional study from Jordan. *Patient Prefer. Adherence* **17**, 895–903 (2023).
19. Amorim, M. et al. Benefits and risks of sharing genomic data for research: comparing the views of rare disease patients, informal carers and healthcare professionals. *Int. J. Environ. Res. Public Health* **19**, 8788 (2022).
20. Atkin, C. et al. Perceptions of anonymised data use and awareness of the NHS data opt-out amongst patients, carers and healthcare staff. *Res. Involv. Engagem.* **7**, 40 (2021).
21. Bakken, V. et al. Attitudes of mental health service users toward storage and use of electronic health records. *Psychiatr. Serv.* **73**, 1013–1018 (2022).
22. Bearth, A. & Siegrist, M. Psychological factors that determine people's willingness-to-share genetic data for research. *Clin. Genet.* **97**, 483–491 (2020).
23. Belfrage, S., Helgesson, G. & Lynoe, N. Trust and digital privacy in healthcare: a cross-sectional descriptive study of trust and attitudes towards uses of electronic health data among the general public in Sweden. *BioMed Cent. Med. Ethics* **23**, 19 (2022).
24. Belfrage, S., Lynöe, N. & Helgesson, G. Willingness to share yet maintain influence: a cross-sectional study on attitudes in Sweden to the use of electronic health data. *Public Health Ethics* **14**, 23–34 (2021).
25. Biasiotto, R. et al. Public preferences for digital health data sharing: discrete choice experiment study in 12 European countries. *J. Med. Internet Res.* **25**, e47066 (2023).
26. Bouras, A. et al. Non-hispanic white mothers' willingness to share personal health data with researchers: survey results from an opt-in panel. *J. Particip. Med.* **12**, e14062 (2020).
27. Braunack-Mayer, A. et al. Sharing government health data with the private sector: community attitudes survey. *J. Med. Internet Res.* **23**, e24200 (2021).
28. Corman, A., Canaway, R., Culnane, C. & Teague, V. Public comprehension of privacy protections applied to health data shared for research: an Australian cross-sectional study. *Int. J. Med. Inform.* **167**, 104859 (2022).
29. Dobson, R. et al. Patient perspectives on the use of health information. *N. Z. Med. J.* **134**, 66 (2021).
30. Eikemo, H., Roten, L. T. & Vaaler, A. E. Research based on existing clinical data and biospecimens: a systematic study of patients' opinions. *BioMed Cent. Med. Ethics* **23**, 60 (2022).
31. Etchegary, H., Darmonkov, G., Simmonds, C., Pullman, D. & Rahman, P. Public attitudes towards genomic data sharing: results from a provincial online survey in Canada. *BioMed Cent. Med. Ethics* **24**, 81 (2023).
32. Franklin, E. F., Nichols, H. M., House, L., Buzaglo, J. & Thiboldeaux, K. Cancer patient perspectives on sharing of medical records and mobile device data for research purposes. *J. Patient Exp.* **7**, 1115–1121 (2020).
33. Fylan, F. & Fylan, B. Co-creating social licence for sharing health and care data. *Int. J. Med. Inform.* **149**, 104439 (2021).
34. Garett, R. & Young, S. D. Ethical views on sharing digital data for public health surveillance: analysis of survey data among patients. *Front. Big Data* **5**, 871236 (2022).
35. Gonzalez-Prieto, C. et al. Exploring older peoples' attitudes and preferences around the use of their healthcare information. *N. Z. Med. J.* **136**, 64–86 (2023).
36. Gupta, R. et al. Consumer views on privacy protections and sharing of personal digital health information. *J. Am. Med. Assoc. Netw. Open* **6**, e231305 (2023).
37. Hammack-Aviran, C. M. et al. Research Use of Electronic Health Records: Patients' Views on Alternative Approaches to Permission. *Am. J. Bioeth. Empir. Bioeth.* **11**, 172–186 (2020).
38. Hanson, C. K. & Liu, K. Think about your friends and family: the disparate impacts of relationship-centered messages on privacy concerns, protective health behavior, and vaccination against Covid-19. *Public Libr. Sci. One* **17**, e0270279 (2022).
39. Helou, S., Abou-Khalil, V., El Helou, E. & Kiyono, K. Factors related to personal health data sharing: data usefulness, sensitivity and anonymity. *Stud. Health Technol. Inf.* **281**, 1051–1055 (2021).
40. Holm, S., Kristiansen, T. B. & Ploug, T. Control, trust and the sharing of health information: the limits of trust. *J. Med. Ethics* **47**, medethics-2019–105887 (2021).
41. Hutchings, E., Butcher, B. E., Butow, P. & Boyle, F. M. Attitudes of Australian breast cancer patients toward the secondary use of administrative and clinical trial data. *Asia-Pac. J. Clin. Oncol.* **19**, e12–e26 (2023).
42. Jagsi, R. et al. Health data sharing perspectives of patients receiving care in cancerlinq-participating oncology practices. *J. Am. Med. Inform. Assoc.* **19**, 626 (2023).
43. Johansson, J. V. et al. Preferences of the public for sharing health data: discrete choice experiment. *J. Med. Inform. Res. Med. Inform.* **9**, e29614 (2021).
44. Jones, L. A. et al. Public opinion on sharing data from health services for clinical and research purposes without explicit consent: an anonymous online survey in the UK. *Br. Med. J. Open* **12**, e057579 (2022).
45. Jung, J. et al. Survey of public attitudes toward the secondary use of public healthcare data in Korea. *Healthc. Inform. Res.* **29**, 377–385 (2023).
46. Jung, Y., Choi, H. & Shim, H. Individual willingness to share personal health information with secondary information users in South Korea. *Health Commun.* **35**, 659–666 (2020).
47. Kallesoe, S. A. S. et al. Canadians' opinions towards COVID-19 data-sharing: a national cross-sectional survey. *Br. Med. J. Open* **13**, e066418 (2023).
48. Kharlamov, A., Hohmann, R. & Parry, G. Data sharing decisions: perceptions and intentions in healthcare. *Strateg. Chang.* **32**, 223–237 (2023).
49. Khatatbeh, M., Gharaibeh, L. F., Khabour, O. F., Abu-Farha, R. K. & Alzoubi, K. H. Jordanian views regarding sharing of medical data for research: a cross-sectional study during COVID-19 pandemic. *Public Libr. Sci. One* **17**, e0265695 (2022).
50. Kirkham, E. J. et al. Experience of clinical services shapes attitudes to mental health data sharing: findings from a UK-wide survey. *BioMed Cent. Public Health* **22**, 357 (2022).
51. Köngeter, A. et al. Patients' willingness to provide their clinical data for research purposes and acceptance of different consent models: findings from a representative survey of patients with cancer. *J. Med. Internet Res.* **24**, e37665 (2022).
52. Kusunose, M. & Muto, K. Public attitudes toward cloud computing and willingness to share personal health records (PHRs) and genome data for health care research in Japan. *Hum. Genome Var.* **10**, 11 (2023).

53. Lysaght, T. et al. Trust and trade-offs in sharing data for precision medicine: a national survey of Singapore. *J. Personal. Med.* **11**, 921 (2021).
54. McCormick, J. B., Hopkins, M., Lehman, E. B. & Green, M. J. Mining the data: exploring rural patients' attitudes about the use of their personal information in research. *Am. J. Bioeth. Empir. Bioeth.* **13**, 89–106 (2022).
55. Middleton, A. et al. Global public perceptions of genomic data sharing: what shapes the willingness to donate DNA and health data?. *Am. J. Hum. Genet.* **107**, 743–752 (2020).
56. Milne, R. et al. Demonstrating trustworthiness when collecting and sharing genomic data: public views across 22 countries. *Genome Med* **13**, 92 (2021).
57. Muller, S. H. A., van Thiel, G., Mostert, M. & van Delden, J. J. M. Dynamic consent, communication and return of results in large-scale health data reuse: survey of public preferences. *Digit. Health* **9**, 20552076231190997 (2023).
58. Muller, S. H. A., van Thiel, G., Vrana, M., Mostert, M. & van Delden, J. J. M. Patients' and publics' preferences for data-intensive health research governance: survey study. *J. Med. Inform. Res. Human Factors* **9**, e36797 (2022).
59. Nwebonyi, N., Silva, S. & De Freitas, C. Public Views about involvement in decision-making on health data sharing, access, use and reuse: the importance of trust in science and other institutions. *Front. Public Health* **10**, 852971 (2022).
60. Ong, R. Factors affecting patient and public perceptions of the adoption of electronic health record sharing: a Hong Kong study. *Int. J. Med. Inform.* **178**, 105193 (2023).
61. Parobek, C. M. et al. Video education about genetic privacy and patient perspectives about sharing prenatal genetic data: a randomized trial. *Am. J. Obstet. Gynecol.* **227**, e1–87.e13 (2022).
62. Parvinen, L., Alamäki, A., Hallikainen, H. & Mäki, M. Exploring the challenges of and solutions to sharing personal genomic data for use in healthcare. *Health Inform. J*. **29**, https://doi.org/10.1177/14604582231152185 (2023).
63. Pletscher, F., Lerch, K. M. & Glinz, D. Willingness to share anonymised routinely collected clinical health data in Switzerland: a cross-sectional survey. *Swiss Med. Wkly* **152**, w30182 (2022).
64. Raddatz, N. I., Kettinger, W. J. & Coyne, J. Giving to get well: patients' willingness to manage and share health information on ai-driven platforms. *Commun. Assoc. Inf. Syst.* **52**, 1017–1049 (2023).
65. Richter, G., Borzikowsky, C., Hoyer, B. F., Laudes, M. & Krawczak, M. Secondary research use of personal medical data: patient attitudes towards data donation. *BioMed Cent. Med. Ethics* **22**, 495–502 (2021).
66. Richter, G. et al. Secondary research use of personal medical data: attitudes from patient and population surveys in The Netherlands and Germany. *Eur. J. Hum. Genet.* **29**, 495–502 (2021).
67. Schultz-Swarthfigure, C., Kelly, A. M. & Zion, D. Emergency department patients' attitudes towards the use of data in their clinical record for research without their consent. *J. Med. Ethics* **49**, 75–78 (2023).
68. Shi, J. et al. Factors influencing the sharing of personal health data based on the integrated theory of privacy calculus and theory of planned behaviors framework: results of a cross-sectional study of Chinese Patients in the Yangtze River Delta. *J. Med. Internet Res.* **25**, e46562 (2023).
69. Silber, H. et al. A preregistered vignette experiment on determinants of health data sharing behavior: Willingness to donate sensor data, medical records, and biomarkers. *Polit. Life Sci.* **41**, 161–181 (2022).
70. Soellner, M. & Koenigstorfer, J. Motive perception pathways to the release of personal information to healthcare organizations. *BioMed Cent. Inform. Decis. Mak.* **22**, 240 (2022).
71. Summers, C., Griffiths, F., Cave, J. & Panesar, A. Understanding the security and privacy concerns about the use of identifiable health data in the context of the COVID-19 pandemic: survey study of public attitudes toward COVID-19 and data-sharing. *J. Med. Inform. Res. Form. Res.* **6**, e29337 (2022).
72. Tosoni, S. et al. The use of personal health information outside the circle of care: consent preferences of patients from an academic health care institution. *BioMed. Cent. Med. Ethics* **22**, 1–14 (2021).
73. Tosoni, S. et al. Patient consent preferences on sharing personal health information during the COVID-19 pandemic: "the more informed we are, the more likely we are to help". *BioMed Cent. Med. Ethics* **23**, 53 (2022).
74. Trinidad, M. G., Platt, J. & Kardia, S. L. R. The public's comfort with sharing health data with third-party commercial companies. *Humanit. Soc. Sci. Commun.* **7**, 149 (2020).
75. Tully, M. P., Bernsten, C., Aitken, M. & Vass, C. Public preferences regarding data linkage for research: a discrete choice experiment comparing Scotland and Sweden. *BioMed Cent. Med. Inform. Decis. Mak.* **20**, 109 (2020).
76. Varhol, R. J. et al. Public preference on sharing health data to inform research, health policy and clinical practice in Australia: a stated preference experiment. *Public Libr. Sci. One* **18**, e0290528 (2023).
77. Vidgen, M. E., Kaladharan, S., Malacova, E., Hurst, C. & Waddell, N. Sharing genomic data from clinical testing with researchers: public survey of expectations of clinical genomic data management in Queensland, Australia. *BioMed Cent. Med. Ethics* **21**, 119 (2020).
78. Wagner, A. L. et al. US residents' preferences for sharing of electronic health record and genetic information: a discrete choice experiment. *Value Health* **26**, 1301–1307 (2023).
79. Warren, V., Critchley, C., McWhirter, R., Walshe, J. & Nicol, D. Context matters in genomic data sharing: a qualitative investigation into responses from the Australian public. *BioMed Cent. Med. Genom.* **15**, 275 (2022).
80. Yamanaka, M., Suzuki, M. & Sato, K. Patient perspectives on research use of residual biospecimens and health information: On the necessity of obtaining societal consent by creating a governance structure based on value-sharing. *Res. Ethics* **17**, 103–119 (2021).
81. Yarborough, B. J. H. & Stumbo, S. P. Patient perspectives on acceptability of, and implementation preferences for, use of electronic health records and machine learning to identify suicide risk. *Gen. Hosp. Psychiatry* **70**, 31–37 (2021).
82. Yiannakoulias, N. Understanding identifiability in secondary health data. *Can. J. Public Health* **102**, 291–293 (2011).
83. Baines, R. et al. Patient and public willingness to share personal health data for third-party or secondary uses: systematic review. *J. Med. Internet Res.* **26**, e50421 (2024).
84. Benevento, M. et al. Measuring the willingness to share personal health information: a systematic review. *Front. Public Health* **11**, 1213615 (2023).
85. Cumyn, A. et al. Patients' and members of the public's wishes regarding transparency in the context of secondary use of health data: scoping review. *J. Med. Internet Res.* **25**, e45002 (2023).
86. Fischer, B. A. t. A summary of important documents in the field of research ethics. *Schizophr. Bull.* **32**, 69–80 (2006).
87. Bell, E. A., Ohno-Machado, L. & Grando, M. A. Sharing my health data: a survey of data sharing preferences of healthy individuals. *AMIA Annu Symp. Proc.* **2014**, 1699–1708 (2014).
88. Botkin, J. R., Rothwell, E., Anderson, R., Stark, L. A. & Mitchell, J. Public attitudes regarding the use of electronic health information and residual clinical tissues for research. *J. Community Genet.* **5**, 205–213 (2014).
89. Sterckx, S., Rakic, V., Cockbain, J. & Borry, P. "You hoped we would sleep walk into accepting the collection of our data": controversies surrounding the UK care.data scheme and their wider relevance for biomedical research. *Med. Health Care Philos.* **19**, 177–190 (2016).

90. Stall, R., Dodge, B., Bauermeister, J. A., Poteat, T. & Beyrer, C. *LGBTQ Health Research: Theory, Methods, Practice* (John Hopkins University Press, 2020).
91. Khatri, R. B. & Assefa, Y. Access to health services among culturally and linguistically diverse populations in the Australian universal health care system: issues and challenges. *BMC Public Health* **22**, 880 (2022).
92. Lounsbury, O. et al. Opening a "can of worms" to explore the public's hopes and fears about health care data sharing: qualitative study. *J. Med. Internet Res.* **23**, e22744 (2021).
93. Rechel, B., Mladovsky, P. & Deville, W. Monitoring migrant health in Europe: a narrative review of data collection practices. *Health Policy* **105**, 10–16 (2012).
94. de Freitas, C. & Martin, G. Inclusive public participation in health: policy, practice and theoretical contributions to promote the involvement of marginalised groups in healthcare. *Soc. Sci. Med.* **135**, 31–39 (2015).
95. Goodman, D. et al. De-identified genomic data sharing: the research participant perspective. *J. Community Genet.* **8**, 173–181 (2017).
96. Garrison, N. A. et al. A systematic literature review of individuals' perspectives on broad consent and data sharing in the United States. *Genet. Med.* **18**, 663–671 (2016).
97. Wang, X. & Cheng, Z. Cross-sectional studies: strengths, weaknesses, and recommendations. *Chest* **158**, S65–S71 (2020).
98. Davenport, T. & Kalakota, R. The potential for artificial intelligence in healthcare. *Future Health J.* **6**, 94–98 (2019).
99. Dron, L. et al. Data capture and sharing in the COVID-19 pandemic: a cause for concern. *Lancet Digit. Health* **4**, e748–e756 (2022).
100. Mutharasan, R. K. & Walradt, J. Population Health and Artificial Intelligence. *JACC Adv.* **3**, 101092 (2024).
101. Shipman, F. M. & Marshall, C. C. Ownership, Privacy, and Control in the Wake of Cambridge Analytica: The Relationship between Attitudes and Awareness. In *Proc. 2020 CHI Conference on Human Factors in Computing Systems* 1–12 (Association for Computing Machinery, 2020).
102. Nagappan, A. & Zhu, X. Patterns of willingness to share health data with key stakeholders in US consumers: a latent class analysis. *J Am. Med. Inform. Assoc.* **32**, 702–711 (2025).
103. Khan, N. et al. Diversity, community engagement and co-design in research: a rapid review. *BMJ Lead*. https://doi.org/10.1136/leader-2024-001046 (2024).
104. Page, M. J. et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* **372**, n71 (2021).
105. Veritas Health Innovation. Covidence systematic review software. Melbourne, Australia. Available at www.covidence.org (2024).
106. Bank, T. W. The world by income. World Development Indicators. World Bank Group Archives, 2023.
107. Hong, Q. N. et al. Mixed methods appraisal tool (MMAT), version 2018. *Regist. Copyr.* **1148552**, 1–7 (2018).
108. Furuya-Kanamori, L., Barendregt, J. J. & Doi, S. A. R. A new improved graphical and quantitative method for detecting bias in meta-analysis. *Int. J. Evid. Based Health* **16**, 195–203 (2018).
109. Pinto-Coelho, L. How artificial intelligence is shaping medical imaging technology: a survey of innovations and applications. *Bioengineering* **10**, 1435 (2023).
110. Eloranta, S. & Boman, M. Predictive models for clinical decision making: deep dives in practical machine learning. *J. Intern. Med.* **292**, 278–295 (2022).
111. Maro, J. C. et al. Six years of the US Food and Drug Administration's postmarket active risk identification and analysis system in the sentinel initiative: implications for real world evidence generation. *Clin. Pharmacol. Ther.* **114**, 815–824 (2023).
112. Verweij, S., Weemers, J., Jonker, C. & Pasmooij, A. M. G. Learning from European regulator-initiated studies for regulatory decision making. *Drug Discov. Today* **30**, 104256 (2025).
113. Meystre, S. M. et al. Clinical data reuse or secondary use: current status and potential future progress. *Yearb. Med. Inf.* **26**, 38–52 (2017).
114. Birkhead, G. S., Klompas, M. & Shah, N. R. Uses of electronic health records for public health surveillance to advance public health. *Annu. Rev. Public Health* **36**, 345–359 (2015).
115. Glick, T. H. Viewpoint: evidence-guided education: patients' outcome data should influence our teaching priorities. *Acad. Med.* **80**, 147–151 (2005).
116. Chen, C. et al. Merging data curation and machine learning to improve nanomedicines. *Adv. Drug Deliv. Rev.* **183**, 114172 (2022).

Acknowledgements

The authors would like to thank Danyang Dai for her assistance with the analysis, Dr Zahed Lambat for assistance with data extraction and Dr Tuan Duong for support and editorial assistance. Funded in kind by Queensland Digital Health Centre (QDHeC), University of Queensland.

Author contributions

L.W., J.P., A.D., Q.O., R.E., B.R. and N.P. designed the study and C.S. and J.P. supervised. Q.O., E.L., L.W., A.D., R.E., J.P. and B.R. conducted the screening. Q.O., A.D., J.P. and R.E. undertook quality analysis. Q.O. and A.D. extracted the data. Q.O. wrote the manuscript and analysed the data. All authors (including M.K. and L.M.) contributed to editing and manuscript preparation.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41746-025-01868-9.

Correspondence and requests for materials should be addressed to Jason D. Pole.