Journal of the American Medical Informatics Association Advance Access published April 23, 2016

Public preferences for electronic health data storage, access, and sharing – evidence from a pan-European survey

RECEIVED 14 September 2015
REVISED 24 November 2015
ACCEPTED 16 January 2016

Sunil Patil,[1] Hui Lu,[1] Catherine L Saunders,[1] Dimitris Potoglou,[2] and Neil Robinson[1]

ABSTRACT

Objective To assess the public's preferences regarding potential privacy threats from devices or services storing health-related personal data.
Materials and Methods A pan-European survey based on a stated-preference experiment for assessing preferences for electronic health data storage, access, and sharing.
Results We obtained 20 882 survey responses (94 606 preferences) from 27 EU member countries. Respondents recognized the benefits of storing electronic health information, with 75.5%, 63.9%, and 58.9% agreeing that storage was important for improving treatment quality, preventing epidemics, and reducing delays, respectively. Concerns about different levels of access by third parties were expressed by 48.9% to 60.6% of respondents.
On average, compared to devices or systems that only store basic health status information, respondents preferred devices that also store identification data (coefficient/relative preference 95% CI = 0.04 [0.00-0.08], $P = 0.034$) and information on lifelong health conditions (coefficient = 0.13 [0.08 to 0.18], $P < 0.001$), but there was no evidence of this for devices with information on sensitive health conditions such as mental and sexual health and addictions (coefficient = −0.03 [−0.09 to 0.02], $P = 0.24$). Respondents were averse to their immediate family (coefficient = −0.05 [−0.05 to −0.01], $P = 0.011$) and home care nurses (coefficient = −0.06 [−0.11 to −0.02], $P = 0.004$) viewing this data, and strongly averse to health insurance companies (coefficient = −0.43 [−0.52 to 0.34], $P < 0.001$), private sector pharmaceutical companies (coefficient = −0.82 [−0.99 to −0.64], $P < 0.001$), and academic researchers (coefficient = −0.53 [−0.66 to −0.40], $P < 0.001$) viewing the data.
Conclusions Storing more detailed electronic health data was generally preferred, but respondents were averse to wider access to and sharing of this information. When developing frameworks for the use of electronic health data, policy makers should consider approaches that both highlight the benefits to the individual and minimize the perception of privacy risks.

Keywords: Health records, data privacy, stated preference, personal data, public preferences, attitudes

BACKGROUND

The electronic storage, access, and sharing of medical data through personal devices or systems of clinical electronic patient records underpin current developments in record-keeping and communication between patients and professionals,[1] calculation of payments to health care providers,[2] measurement of health care quality,[3] and patient engagement with their health and health care, and they provide a resource for medical research.[4–7] The scope and volume of information stored has grown exponentially over the last 2 decades. Data sources range from personal information uploaded voluntarily to personal devices (eg, through smartphone apps[8] or fitness gadgets) or websites (eg, PatientsLikeMe in the United States or LifeSensor in Europe),[9] or in formal records of the clinical details of interactions with health services (electronic patient records).[1] Patient health records have evolved from paper documents designed for record-keeping and communication between health care professionals, protected by doctor-patient confidentiality, to electronic documents of which patients and doctors are co-producers.[1,5,7] In parallel, exploiting the economic potential of health care data is becoming embedded into research and health service strategies.[5,6,10]

Currently, the European Data Protection Directive[11] is the overarching regulatory framework governing health care data in Europe; however, there is some divergence in EU countries as to how this directive is applied in practice,[12] and it is currently under review.[13] Specific guidance for storage, access to, and sharing of health data has been slow to develop. This is partly because the speed of technological change has outpaced law-making, but also because of the difficulties of balancing the seemingly competing priorities of individual privacy of health data against other priorities for the use of this data. No clear consensus about how these priorities should be translated into law, policy, and practice has yet emerged,[14] although reviews are in progress.[13,15,16] A recent UK information governance review acknowledges this connection directly, that the duty to share information in patients' wider interests is as important as the duty to maintain patient confidentiality.[15]

Behind the need for policy guidance within the context of these sometimes competing priorities is a need for clear, high-quality evidence about what public preferences are for electronic health data, and whether or how these public preferences illuminate trade-offs of benefits and risks, between privacy and data storage, access, and sharing.[14,17,18] Research to date has been mixed, with conflicting findings, highlighting both concern and support for sharing or use of medical records in research, for example.[17,19–26]

Stated preference (SP) experiments have been used extensively in the fields of marketing, transport economics, environmental valuation, health, and health care.[27] Briefly, compared with opinion polls or traditional survey approaches, SP experiments provide a more nuanced insight into preferences and allow a number of the attributes that may

Correspondence to Sunil Patil, RAND Europe, Westbrook Centre, Milton Road, Cambridge, CB4 1YG, UK; spatil@rand.org; Tel: +44 1223 353 329. For numbered affiliations see end of article.

© The Author 2016. Published by Oxford University Press on behalf of the American Medical Informatics Association. This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

RESEARCH AND APPLICATIONS
Downloaded from http://jamia.oxfordjournals.org/ by guest on May 17, 2016

influence decision-making to be controlled for simultaneously. In an SP experiment, respondents are asked to indicate the most preferred alternative in a series of hypothetical scenarios which are often close to real-life situations. Each alternative is described using multiple attributes. Analysis of preferences from all respondents allows estimation of relative preferences for each attribute and the values they take. In this work we use this methodology to explore levels of privacy concern across Europe and to describe European preferences regarding devices or services storing health-related personal data empirically.

METHODS

Study participants

Participants were recruited as part of PACT: *Public perception of security and privacy: Assessing knowledge, Collecting evidence, Translating research into action,* a European Commission Seventh Framework Programme funded research project. The project was designed to assess public perceptions of security and privacy in 3 contexts: travel on metro or train, choice of Internet service provider, and purchase of a device or service for storing health-related personal data. Adults 18 and over were recruited from across the 27 European Union (EU27) member states (this study was conceptualized in 2012, before Croatia acceded).

Question development

The survey questionnaire and the SP experiment were developed through literature review, focus groups (in the UK, Lithuania, and Greece), and stakeholder consultation.[18,28] The questionnaire included 4 sections: participants were asked about demographic information, background on their own health status, attitudes toward data privacy, and the SP scenarios. Cognitive testing was carried out in the UK,

Table 1: Respondent characteristics

	Number of respondents (%)		Survey mode	Language(s)	Number of respondents (%)
All respondents	20 882	**Country**			
		Austria	Online	German	721 (3.5)
Gender		Belgium	Online	French, Dutch	698 (3.4)
Male	9960 (47.7)	Bulgaria	Face-to-face	Bulgarian	877 (4.2)
Female	10 922 (52.3)	Cyprus	Face-to-face	Greek	577 (2.8)
		Czech Republic	Face-to-face	Czech	757 (3.6)
Age		Denmark	Online	Danish	744 (3.6)
18–24	2163 (10.4)	Estonia	Online	Estonian	752 (3.6)
25–34	3512 (16.8)	Finland	Online	Finnish	712 (3.4)
35–44	3719 (17.8)	France	Online	French	730 (3.5)
45–54	3763 (18)	Germany	Mixed	German	777 (3.7)
55–64	3735 (17.9)	Greece	Face-to-face	Greek	880 (4.2)
65+	3990 (19.1)	Hungary	Face-to-face	Hungarian	831 (4.0)
		Ireland	Online	English	696 (3.3)
Household income		Italy	Mixed	Italian	784 (3.8)
<€500	3579 (17.1)	Latvia	Face-to-face	Latvian, Russian	868 (4.2)
€500–€1500	6425 (30.8)	Lithuania	Face-to-face	Lithuanian	1014 (4.9)
€1500–€3000	4444 (21.3)	Luxembourg	Online	French, German, Luxembourgish	551 (2.6)
€3000–€9000	2846 (13.6)	Malta	Face-to-face	Maltese	650 (3.1)
>€9000	160 (0.8)	Netherlands	Online	Dutch	771 (3.7)
(Missing)	3428 (16.4)	Poland	Face-to-face	Polish	863 (4.1)
		Portugal	Face-to-face	Portuguese	901 (4.3)
Survey Approach		Romania	Face-to-face	Romanian	
Online	9198 (44.0)	Slovakia	Face-to-face	Slovak	846 (4.1)
Offline	11684 (56.0)	Slovenia	Face-to-face	Slovenian	885 (4.3)
		Spain	Online	Spanish	685 (3.3)
Self-rated health		Sweden	Online	Swedish	717 (3.4)
Good or very good	12823 (61.4)	United Kingdom	Online	English	714 (3.4)

Downloaded from http://jamia.oxfordjournals.org/ by guest on May 17, 2016

Hungary, and Portugal once the first draft was finalized, focusing on whether scenarios were understood by the respondents across the European countries.[29] Surveys were translated into the official language(s) of each country (Table 1).

Demographic questions

Participants were asked their age in years, gender, and household income (in 17 categories from <200 Euros to >12 000 Euros per month).

Privacy concerns

Respondents' concern about privacy of information stored on a health device was collected based on their responses about levels of concern to the following three items: *Personal information (name, address, health conditions) accessed by non-medical personnel (eg, police); personal information (name, address, health conditions) accessed by private companies (such as pharmaceutical and insurance companies); misuse of personal information for harassment based on race, health status, sexual orientation, etc.,* with 5-point Likert scale options ranging from "not concerned at all" to "very concerned." Additional questions explored concerns about access to data and opinions of current levels of health data security, with 5-point Likert-type responses ranging from "agree strongly" to "disagree strongly." Questions were developed from existing tools, and[18,30,31] full wording is given in the footnote to Table 2.

Stated-preference scenarios

The SP experiment was framed around the choice of a health data storage device/system involving 5 attributes: the level of personal health information stored, the authorities, and their geographical regions, that can access the information, preferences for whom this data can be shared with apart from medical personnel, and an associated cost for using the device. Briefly, in a SP experiment, respondents are asked to indicate the most preferred alternative in a series of hypothetical scenarios which are close to real-life situations. Each alternative is described using a combination of options from multiple attributes, and an analysis of preferences from all respondents allows estimation of relative preferences for each attribute and the values it takes. Each attribute is described using a level in each choice task; a full list of attributes and levels is given in Box 1 and an example scenario is presented in Figure 1. In the design of the experiment, out of all possible scenarios involving different combinations of all attribute levels, 120 were selected, using the Ngene survey design software.[32] These 120 scenarios were chosen to reduce the standard errors of estimated parameters. Each respondent considered 5 different choices, between 2 of these scenarios.

Data collection

Data collection across the EU 27 countries was carried out by Ipsos, a market research company. Online interviews were conducted in 12 countries (Table 1), using quota sampling for age, gender, and region of residence within each country. Face-to-face interviews were conducted in 13 countries with the lowest Internet access. In these 13 countries, areas were stratified by region and level of urbanization; households were selected at random within each selected area; and within each household 1 individual was sampled, with national quotas for age and gender. In Germany and Italy, both survey modes were used in order to allow for potential differences between survey approaches to be accounted for at the analysis stage. Full sampling details are available.[33] Pilot surveys were conducted in Denmark, Italy, and Romania between May and July 2013. The main survey data collection was carried out from August to November 2013. Respondents were made aware of the data handling and anonymity protocols and

Table 2: Health and health data storage, privacy, and access attitudes and opinions

	N	Responses (%)	
Attitudes to data storage			
Storing health information is useful for improving treatment quality[a]	20464	75.5	Agree or agree strongly
Storing health information is useful for preventing health epidemics[b]	20361	63.9	Agree or agree strongly
Lack of personal and health information leads to delays in treatment in health emergencies[c]	20368	58.9	Agree or agree strongly
Health providers collect too much personal information[d]	20391	37.0	Agree or agree strongly
Concerns about access to data			
Access to personal information by non-medical personnel[e]	20696	48.9	Concerned or very concerned
Access to personal information by private companies[e]	20676	60.6	Concerned or very concerned
Misuse of personal information for harassment[e]	20572	54.5	Concerned or very concerned
Opinions about data security			
Healthc are providers are successful in preventing unauthorized access[f]	19372	38.4	Agree or agree strongly
Computer databases should be protected from unauthorized access, regardless of cost[g]	20134	73.4	Agree or agree strongly

[a]"A system which stores health information (such as your blood group, allergies, and health conditions) can be useful in providing higher-quality treatments"[b]"A system which stores health-related information (such as your blood group, allergies, and health conditions) can be useful in preventing health epidemics (eg, H1N1/swine flu)"[c]"I am concerned that in a health emergency there could be an unacceptable delay due to the time spent in identifying the person needing help and their health conditions before the treatment begins"[d]"I'm concerned that health care providers (such as hospitals and health insurance companies) are collecting too much personal information about me" (Note: this is a negatively worded question)[e]These 3 items combine to form the "Health care privacy index"[f]"Health care providers (such as hospitals and health insurance companies) are successful in preventing unauthorized access to personal information"[g]"Computer databases that contain health information (including health conditions, allergies, and identification) should be protected from unauthorized access no matter how much it costs"

RESEARCH AND APPLICATIONS Downloaded from http://jamia.oxfordjournals.org/ by guest on May 17, 2016

Box 1: Attributes and levels used in the SP experiment

What information is stored on the device/system?
Only basic health status information
Basic health status information
Identification
Lifelong health conditions
All other health conditions and medical history
Who can access the information?
Only doctors and nurses
Doctors, nurses, and emergency medical personnel (paramedics)
Doctors, nurses, emergency medical personnel (paramedics), and other non-medical emergency personnel (fire and rescue)
In which countries can your information can be accessed?
Only in [country of residence]
Across Europe (EU)
Worldwide
Who else can view this information apart from the medical specialists?
No one
Immediate family
Nurses providing home care
Health insurance companies
Private sector pharmaceutical companies
Academic researchers (If your name is not connected to the data)
Cost
Free (given by your hospital/national government)
0.5 €/month
1 €/month
2 €/month
3 €/month

study purpose prior to the survey. Participation was voluntary; data collection commenced only after obtaining individual informed consent.

Analysis

In order to describe variation in the levels of sensitivity to the privacy of health care information across Europe, a composite measure of concern for privacy of the information stored on the health device was computed by country using Westin's methodology.[34] Using this method, respondents were categorized into 3 groups, high, medium, and low concerns, according to the number of "concerned/very concerned" responses they gave to the 3 questions (Table 2).

The SPs were analyzed using a multinomial logit model (Box 2). The model is based on the principle of Random Utility Maximization; that is, a respondent chooses an alternative (among a set of available alternatives) in order to maximize its "utility," which represents the satisfaction or benefit received by the person from the alternative. Differences in preferences across countries were included by testing whether each country required a separate coefficient in the model, and using a similar approach, the effects of age, gender, and household income were included. Education level and employment status were also considered, but there was no evidence that they improved the fit of the model and hence they were not included in the final analysis; only coefficients for country and socio-demographic characteristics for each preference where $P < 0.05$ were retained. Country scales were included to control for country-specific unobserved factors, and scales by survey methodology were included to control for variation in response quality between the online and face-to-face surveys. Bootstrap resampling was used to account for multiple preference responses from each participant.[35] Descriptive analysis was carried out using SPSS 21,[36] and the choice modeling analysis was performed with ALOGIT 4.2.[37]

RESULTS

A total of 20 882 survey responses were obtained from 27 EU member countries, with 551 to 1014 responses per country (Table 1).

Attitudes and privacy concerns

Overall, respondents recognized the benefits of storing electronic health information, with 75.5%, 63.9%, and 58.9% of respondents agreeing that storing electronic health data was important for improving treatment quality, preventing health epidemics, and reducing treatment delays, respectively. However, between 48.9% and 60.6% of respondents also expressed concerns about different levels of access to this data, and only 38.4% of respondents agreed that health care providers are currently successful in providing effective data security (Table 2). Levels of health privacy concern varied strongly across Europe (Figure 2). The highest proportions of respondents answering "Not concerned at all" about privacy were observed in Sweden, Slovenia, and Denmark, while the highest proportion of respondents answering "Very concerned" was observed in Lithuania.

Preferences

Multinomial logit model estimates based on the SP data are presented in Table 3 (baseline preferences across Europe) and Table 4 (country, age, and gender-specific effects). The coefficients from SP models capture the weight respondents placed on attribute levels. The coefficients thus represent the marginal utility of respondents' gain or loss (Box 2). In general, when coefficients are greater than zero, this means that the attribute level is preferred relative to the reference level (where the coefficient is set to zero) and it is seen as contributing to service improvement or utility gain. Similarly, when coefficients are less than zero, this means that respondents were "averse" to a particular option. For a given attribute, the relative size of coefficients can also be compared. For example, a coefficient of 0.4 means that the preference for a particular attribute level is twice as much as the preference for a level where the coefficient is 0.2 (ie, half the size) relative to the reference level of that attribute; these coefficients are the "betas" from Box 1. The coefficient for the cost is included in the model as a continuous variable which describes the additional disutility of paying for the health device.

For each preference option, the country, age, and gender groups presented in Table 4 were those where preferences were different ($P < 0.05$) from the baseline across Europe. That is, in this table, positive coefficients for males indicate that preferences for the group were more positive, or stronger, than for females across Europe.

Findings from 94 606 SPs from 20 882 survey respondents are presented; 9804 out of 104 410 obtained responses were excluded, because responses were missing or because the length of time taken to respond to the scenario was below the minimum threshold (of less than a minute in the case of the online survey).

Preferences for each attribute are summarized below.

Downloaded from http://jamia.oxfordjournals.org/ by guest on May 17, 2016

Figure 1: Stated preference experiment: Introductory text and example scenario presented to participants

Currently, there are devices or services in the form of health ID, cards, tags, etc. that can store and allow access to your health information. This information may include personal details such as: your name, address, blood group/type, allergies, current health conditions and medical history. These devices or services are available for purchase or may be provided by your national government at some cost to you. Payment may be direct or through your income tax or health insurance contributions.

With these devices, for example, doctors can become aware of any existing medical conditions (e.g., allergies to medicine and food), authorities can identify you and respond faster in an medical-emergency situation (e.g. pandemic, flu) or in emergencies when you are traveling abroad.

On the other hand, information stored on such a device or system may be misused or may affect your privacy, for example when unauthorized users get access to such information.

Which of the following options would you prefer for a health record storage device/system?

Description	Option A	Option B	Option C
What information is stored on the device/system Basic health status: blood group, allergies, diabetic status Identification: name, address, age, photograph, nationality Lifelong health conditions: asthma, disabilities, cancer, etc. All health conditions: mental health, sexual health, addictions, and medical history	**Information on device/service** Basic health status information Identification Lifelong health conditions	**Information on device/service** Basic health status information Identification Lifelong health conditions All other health conditions and medical history	I would not purchase any such device/service
Who can access the information	Doctors, nurses, and emergency medical personnel (paramedics)	Only doctors and nurses	
In which countries your information can be accessed?	Worldwide	Across Europe (EU)	
Who else can view this information apart from the medical specialists?	Health insurance companies	Private sector pharmaceutical companies	
Cost	£ 0.58 per month	£ 2.32 per month	

Previous Next

Storage: On average, compared to a device/system that "only stores basic health status information," respondents preferred health devices that "also store identification data," and "identification data with information on lifelong health conditions," coefficients: 0.04 ($P = 0.034$) and 0.13 ($P < 0.001$), respectively (Table 3). Respondents disliked that devices store "additional information on all health conditions (such as mental, sexual health, and addiction, etc.) and medical history"; however, the aversion is not statistically significant, coefficient: −0.03 ($P = 0.24$). With very few exceptions (Czech Republic, Lithuania, Cyprus, and UK), preferences are consistent across countries. Younger people aged 18 to 24 had consistently stronger preferences for the storage of all 4 levels of information (coefficients: from 0.10 to 0.42) compared with other age groups (Table 4).

Access: Across Europe, on average, respondents preferred that "doctors, nurses, and paramedics" be able to access information, when compared with "only doctors and nurses," the reference level, but were averse to emergency services additionally being able to access the information, coefficients: 0.07 ($P < 0.001$) and −0.06 ($P = 0.017$), respectively. Older respondents (65 and over) had a stronger aversion to widening the access to emergency services, coefficient: −0.11 ($P = 0.001$). Further, with respect to geographic access, respondents in general preferred that information be accessible

RESEARCH AND APPLICATIONS Downloaded from http://jamia.oxfordjournals.org/ by guest on May 17, 2016

Box 2: Multinomial logit model specification

The responses to the SP exercises were modelled in a random utility theory framework using discrete choice models. The overall utility of choice alternative i for respondent j is represented below:

$$U_{ij} = V_{ij} + \varepsilon_{ij}$$

Where V_{ij} represents the deterministic or measurable part of the utility and ε_{ij} is the error term, which represents the unobserved component of the utility and is assumed to independent and identically distributed (iid) over all the alternatives and respondents. In this study there are 3 alternatives: Alt A, Alt B, and "none of these," and it is assumed that respondents choose the alternative which maximizes their utility. The distributional assumptions on ε_{ij} give rise to different types of models. In case of a multinomial logit model, the error term is assumed to follow a Type 1 extreme value distribution (Gumbel) and the resulting probabilities are shown below.

$$V[\text{Alt A}] = \beta * X_{iA};\ P[\text{Alt A}] = e^{V[\text{Alt A}]} / \left(e^{V[\text{Alt A}]} + e^{V[\text{Alt B}]} + e^{V[\text{none}]}\right)$$

$$V[\text{Alt B}] = \beta * X_{iB};\ P[\text{Alt A}] = e^{V[\text{Alt B}]} / \left(e^{V[\text{Alt A}]} + e^{V[\text{Alt B}]} + e^{V[\text{none}]}\right)$$

$$V[\text{none}] = \beta_{\text{none}};\ P[\text{none}] = e^{V[\text{none}]} / \left(e^{V[\text{Alt A}]} + e^{V[\text{Alt B}]} + e^{V[\text{none}]}\right)$$

Where β is the vector of parameters of the attributes and their levels, and X_{iA} and X_{iB} represent matrices of attributes and their levels presented in each of the choice scenarios. The product of probabilities across all respondents is maximized (maximum likelihood method) to estimate the vector of parameters.

across the EU rather than their home country alone. However, they were averse to widening the access beyond the EU (ie, worldwide access). Again, older respondents (65 and over) preferred to limit access to their home country when compared with other age groups.

Sharing: Respondents disliked "immediate family" and "nurses providing home care" being able to access their health information, coefficients: -0.05 ($P = 0.011$) and -0.06 ($P = 0.004$), respectively (Table 3). They were very strongly averse to "health insurance companies" and "private sector pharmaceutical companies" being able to view their health information, coefficients: -0.43 ($P < .001$) and -0.82 ($P < .001$). Respondents were also less likely to select "academic researchers" being able to view the information, even though in this scenario the respondent's name was not connected to the data (ie, anonymous health information), coefficient: -0.53 ($P < .001$).

Cost: Respondents were sensitive to the additional cost per month associated with the devices; people prefer not to pay, and if they have to pay, they prefer a cheaper option. People from households with lower incomes were more sensitive to price than those from households with high incomes. For example, households with an income of $<$500 €/month were more sensitive to cost than households with income $>$9000 €/month, coefficients: -0.0042 and -0.0016, respectively. Respondents who declined to respond to the income question were most cost sensitive, coefficient: -0.0044. Full multinomial logit model outputs are available.[38]

DISCUSSION

We found that respondents across Europe preferred to store more, rather than less, electronic health information and preferred access by all health professionals, rather than just doctors and nurses. However, wider access to include fire and rescue personnel (which may come with higher risk to privacy) and anonymized sharing of this information with academic researchers was not preferred, thus pointing toward a preference for individual-level benefits over broader population-level benefits. Respondents were also averse to sharing their health information with other third parties, such as insurance providers and pharmaceutical companies. Compared to limiting access in the country of residence, access across Europe was preferred, although access worldwide was not. Findings were consistent across most countries in Europe, which is surprising, particularly given country-specific differences including history, legal and political environment, and provision of health care.[12]

Findings in the context of other research

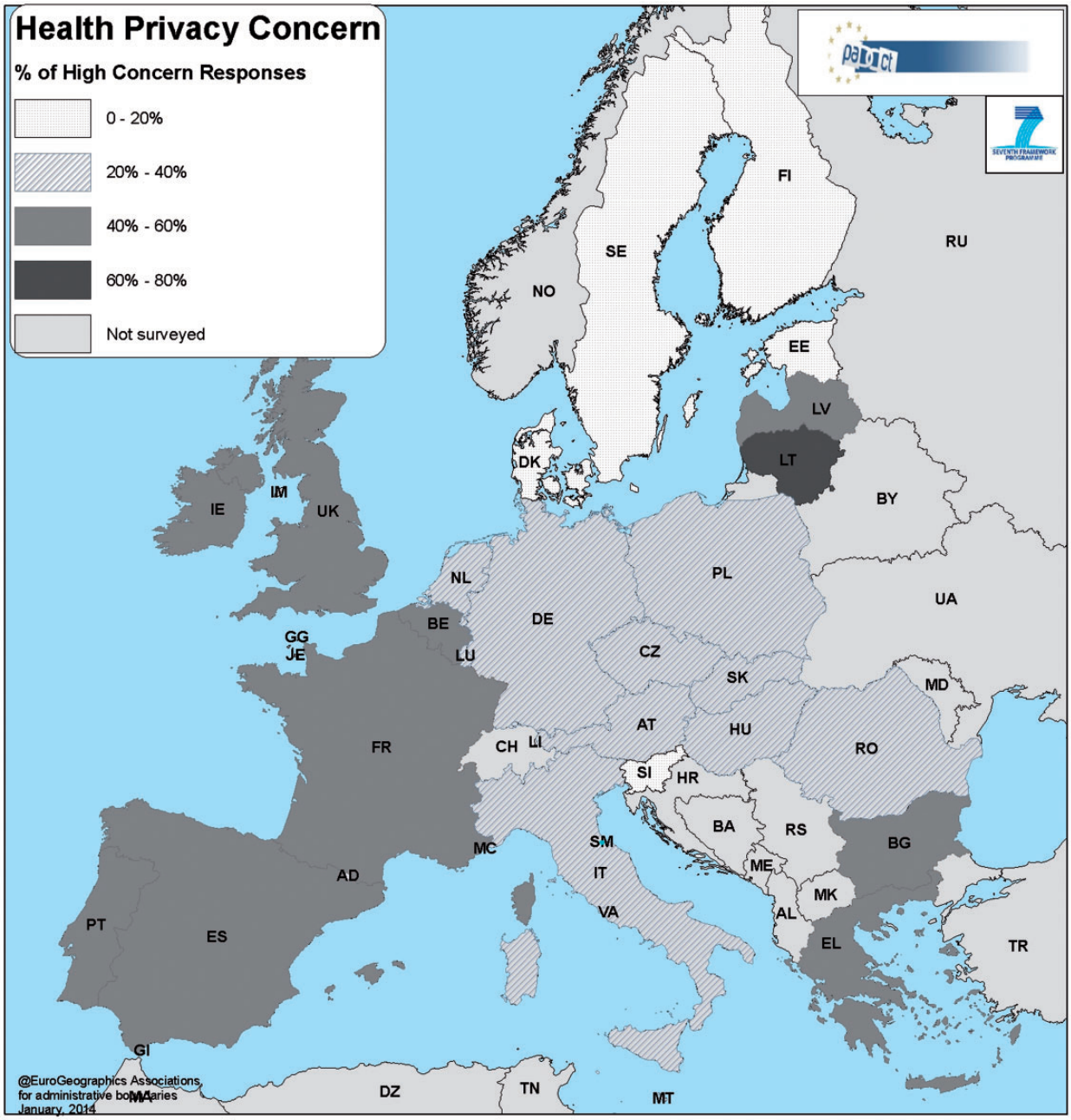
The lack of consensus about public preferences for electronic data storage, access, and sharing in previous research[17,19–26] probably reflects in part the privacy valuation paradox[18]; that is, people value privacy very highly, but express differing preferences when it concerns priorities that they deem important in the immediate context, usually involving money or time, attitudes, and acceptability. The fact that, in general, people have poor understanding of risk in decision-making and a lack of knowledge about how electronic health data is used[17,39] is also important. Previous work illustrates that in the context of electronic health data, people express "concerns" if asked about "concerns" but are able to articulate benefits if asked about these.[19] This pattern is also reflected in our findings, which add additional robust, empirical insight to these perceptions: respondents identify both the benefits and risks to privacy of the storage of electronic health data; however, the range of preferences identified in this work indicate that respondents possess a nuanced understanding of privacy and are able to understand both the risks to privacy and the benefits of electronic health data storage.

Strengths and limitations

To our knowledge, this is the first Europe-wide SP study of this scale, in this context; the methodological approach and findings presented above are innovative and original. A second strength is the large sample size, allowing reliable estimates of preferences from across Europe. Our finding that, with a few exceptions, preferences were consistent across Europe strengthens the generalizability of this work beyond specific national contexts.

Downloaded from http://jamia.oxfordjournals.org/ by guest on May 17, 2016

Figure 2: Levels of high health privacy concern across Europe



Missing data and refusal to participate are the main challenges to any survey research of attitudes and preferences. In this research, not responding to questions requiring personal information (eg, income) was associated with a more risk-averse approach to electronic data storage and access. We would also expect people who were more concerned about data security to be underrepresented overall among respondents. However, using an approach taken in other health survey contexts,[40] by having included "missing income" (representing a proxy for concerns about data privacy and non-response) as a covariate in our choice model, we partly accounted for this in the analysis.

Including attributes in the SP experiment relating to sharing of non-anonymized data with health insurance companies and private pharmaceutical companies, and sharing with academic researchers with anonymization specifically articulated, means that preferences across these 3 groups are less comparable. Nonetheless, they represent real-world situations in which electronic health data may be shared, and the findings from this research provide insights into public preferences across these scenarios.

A third limitation is that the time period for the main stage fieldwork overlapped with publications in the news about US National Security Agency and UK Government Communications Headquarters surveillance practices revealed by Edward Snowden, which conceivably might have influenced some responses due to heightened awareness of issues relating to privacy.

RESEARCH AND APPLICATIONS
Downloaded from http://jamia.oxfordjournals.org/ by guest on May 17, 2016

Table 3: Stated preference choice modelling results of preferences across Europe (baseline group)[a]

Model parameter	Coefficient (preference)	*P*-value
Information stored		
Basic health status	Reference	
[information above +] identification	0.04 (0.00–0.08)	0.34
[information above +] lifelong health conditions	0.13 (0.08–0.18)	<.001
[information above +] all other health conditions and medical history	−0.03 (−0.09 to 0.02)	0.24
Access (personnel)		
Only doctors and nurses	Reference	
Doctors, nurses, and paramedics	0.07 (0.04 to 0.10)	<.001
Doctors, nurses, paramedics, and fire and rescue	−0.06 (−0.11 to −0.01)	0.17
Access (location)		
Only in the home country	Reference	
Across Europe (EU)	0.06 (0.02 to 0.10)	0.02
Worldwide	−0.14 (−0.19 to −0.09)	<.001
Sharing		
No one	Reference	
Immediate family	−0.05 (−0.09 to −0.01)	0.11
Nurses providing home care	−0.06 (−0.11 to −0.02)	0.04
Health insurance companies	−0.43 (−0.52 to −0.34)	<.001
Private sector pharmaceutical companies	−0.82 (−0.99 to −0.64)	<.001
Academic researchers (if your name is not connected to the data)	−0.53 (−0.66 to −0.40)	<.001
Cost		
HH income less than €500	−0.0042 (−0.0050 to −0.0034)	0.00
HH income from €500 to €1500	−0.0036 (−0.0043 to −0.0029)	0.00
HH income from €1500 to €3000	−0.0031 (−0.0037 to −0.0025)	0.00
HH income from €3000 to €9000	−0.0028 (−0.0033 to −0.0023)	0.00
HH income greater than €9000	−0.0016 (−0.0029 to −0.0003)	0.20
Missing income	−0.0044 (−0.0053 to −0.0035)	0.00

[a]In all analyses, differences in preferences for each country, age group, and gender were tested. Where differences did occur, the group was modelled separately for that particular preference. These are given in Table 4. The coefficients in this table can therefore be interpreted as the preferences in all other countries, or groups not covered by the effects given in Table 4.

Finally, some specific issues, such as attitudes toward linkage of health data across multiple sources, were not explored.

SIGNIFICANCE

We now turn to a brief overview of the policy implications of each of the findings.

Implications for policy: storage

In general, this research finds that the perceived benefits of storing electronic health information outweigh the perceived risks to privacy of storing any information at all. The massive shift to increased electronic storage of health information in the past 20 years is broadly in line with public preferences.

However, this evidence also has specific implications for policy regarding the storage of sexual health data. We find that, although respondents preferred storage of information on lifelong health conditions in addition to basic health data (except among 18- to 34-year-olds), the level "All health conditions: mental health, sexual health, addictions, and medical history" was not preferred. In the UK, sexual health services are not stored as part of a patient's health record, and this approach has been criticized for not enabling adequate information for performance management of sexual health services.[41] The findings presented here provide some support for the current system of separation and anonymity in the storage of sensitive medical records.

Implications for policy: access

One area where the findings from this work regarding access to data are particularly relevant is in the development of frameworks for access to European electronic health data worldwide. The recent memorandum of understanding which was signed by the US Department of Health and Human Services and the UK Secretary of State for Health paves the way for the sharing of electronic health data from the UK outside Europe,[42,43] and so it is particularly relevant to note that in this work, access to information within Europe was preferred, but worldwide access was not. Additionally, health information is increasingly being accessed by the police and emergency services (to enable a more effective response to emergencies) and we found in general that this is not preferred.[44]

Implications for policy: sharing

Sharing of electronic health data beyond health care professionals was not preferred. Our work, however, provides insight across several areas in this ongoing debate, particularly around the sharing of electronic health data for research.

Not sharing data at all, or not sharing data without explicit individual consent, is one approach. However, limitations to opt-in approaches to data sharing have been highlighted, particularly gaming, by including a simple tick box at the end of extensive terms and conditions, which are typically not read or fully understood.[45] Our finding that there are variations in preference between men and women and by age also highlights that if this approach were taken, it could not be expected that a representative sample of records would be obtained.

Policy approaches to the sharing of electronic health data could also learn from other sensitive methods used in medical research where personal dislike of an approach (namely our finding that sharing of electronic records for research is not preferred) is at odds with broader research goals designed for the public good. For example, there remains in the United States strong opposition to the use of embryonic stem cells[46] and, particularly in the 1980s and 1990s, there was strong public opinion in the UK against the use of animals in research.[47] The poor understanding of animal research parallels the low public knowledge of how electronic health information is used today.[19,47] It is perhaps relevant, therefore, that in this work the countries found to have the weakest aversion to the sharing of electronic health data for research were Estonia and Denmark. Denmark in particular has a long history of using population databases in medical research, and Estonia has developed a high-profile national genetic database in the past 15 years and is a renowned leader in digitization of many societal functions. It is possible that, over time, as electronic data health research becomes more established, aversion to this use of health data may tail off. Although options for data sharing may not be preferred, in some situations they may become acceptable.[45]

Strong regulatory frameworks, however, have also been part of the national policy landscape governing both animal and embryonic stem

Downloaded from http://jamia.oxfordjournals.org/ by guest on May 17, 2016

Table 4: Preference choice modelling results (summary of differences by country, age and gender)

Model parameter	Differences by country		Differences by age and gender	
	Stronger preference or weaker aversion compared with the baseline group[a]	Weaker preference or stronger aversion compared with the baseline group[a]	Stronger preference or weaker aversion compared with the baseline group[a] (coefficient)	Weaker preference or stronger aversion compared with the baseline group[a] (coefficient)
Information stored				
Basic health status	Czech Republic (0.19)	United Kingdom (−0.11)	18–24 (0.28)	
	Lithuania (0.25)			
[information above +] identification			18–24 (0.23)	
[information above +] lifelong health conditions			18–24 (0.29)	55–64 (−0.09)
			25–34 (0.10)	
[information above +] all other health conditions and medical history	Cyprus (0.72)		Men (0.06)	
			18–24 (0.42)	
			25–34 (0.18)	
Access (personnel)				
Only doctors and nurses	Slovenia (0.30)			
Doctors, nurses, and paramedics	Estonia (0.24)			
Doctors, nurses, paramedics, and fire and rescue		Denmark (−0.11)		65+ (−0.11)
Access (location)				
Only in the home country	Czech Republic (0.18)	Belgium (−0.25)	65+(0.13)	
	Slovakia (0.11)	Ireland (−0.19)		
		Romania (−0.34)		
		Spain (−0.25)		
Across Europe (EU)		Austria (−0.21)		
Worldwide		Slovakia (−0.10)	18–24 (0.11)	65+ (−0.098)
			25–34 (0.14)	
Sharing				
No one	Austria (0.21)	Lithuania (−0.43)		
	Luxembourg (0.21)	Romania (−0.39)		
	Netherlands (0.27)	Slovakia (−0.19)		
Immediate family	Slovenia (0.52)			
Nurses providing home care	Belgium (0.17)	Bulgaria (−0.24)		
		Slovakia (−0.27)		
	France (0.21)	Lithuania (−0.31)		
Health insurance companies	Czech Republic (0.20)	France (−0.40)		
	Latvia (0.36)	Greece (–0.53)		
	Slovakia (0.19)	Italy (−0.29)		
	Hungary (0.22)			
		Ireland (−0.24)		
		UK (−0.25)		
Private sector pharmaceutical companies	Bulgaria (0.35)	Austria (−0.23)		
	Hungary (0.41)	Belgium (−0.20)		
	Latvia (0.47)	Denmark (−0.14)		
	Lithuania (0.22)	Estonia (−0.25)		
	Portugal (0.36)	Germany (−0.47)		
	Romania (0.46)	Luxembourg (−1.56)		
	Slovakia (0.21)	Slovenia (−0.32)		
Academic researchers (if your name is not connected to the data)	Estonia (0.36)	Romania (−0.23)		
	Denmark (0.33)			

[a]In all analyses, differences in preferences for each country, age group, and gender were tested. Where differences did occur, the group was modelled separately for that particular preference and these differences are presented in this table.

RESEARCH AND APPLICATIONS Downloaded from http://jamia.oxfordjournals.org/ by guest on May 17, 2016

cell research, and this may have led to some degree of public acceptance of these approaches. There are strong concerns that anonymized patient records can be uniquely identified,[48] and there have been reported high-profile lapses in security for electronic health information[16,49] (along with findings from this work that health care providers are not perceived as successful in preventing unauthorized access) and technology failures.[50,51] Strong regulatory frameworks and effective sanctions where best practices are not in place are relatively weak in the context of electronic health data sharing.[49,52] Development, however, is not simple; for example, in the context of rapidly developing technology, it can often be unclear what regulations should cover.[53]

CONCLUSIONS

In this large objective pan-European stated preference study, respondents gave a range of preferences regarding health information, storage access, and sharing, indicating a nuanced understanding of privacy. This work clarifies and explains previously inconsistent evidence regarding public perceptions toward privacy, particularly how the risks to privacy in electronic data storage are understood in the context of perceived individual and collective benefits. Access to and sharing of health information beyond those who are involved in immediate care is perceived negatively. When developing frameworks for the use of electronic health data in this rapidly developing field, assurances of accountability and conditionality, and clear descriptions of the benefits to individuals, will be required by policy-makers in further digitization of health information.

FUNDING

This study has received funding from the European Commission's Seventh Framework Programme for research, technological development, and demonstration under grant agreement No. 285635 (Public perception of security and privacy: Assessing knowledge, Collecting evidence, Translating research into action).

ETHICAL APPROVAL

The survey was reviewed by an independent ethics expert, via the funder for this work, the European Commission.

COMPETING INTERESTS

The authors have no competing interests to declare.

AUTHORSHIP

All listed authors meet the criteria for authorship. S.P. led the questionnaire development, analysis, interpretation, and reporting. H.L. estimated the models and undertook analysis for the quantitative part of this paper. C.S. conducted the literature review for this paper and developed the discussion. D.P. and N.R. led the study conception and design, and contributed to the survey design and analysis.

ACKNOWLEDGEMENTS

The authors would like to express their gratitude to RAND Europe's Quality Assurance team, especially Ms. Charlene Rohr and Dr Matthew Bassford, for their useful comments during the project work. We also wish to thank Prof Andrew Daly (RAND Europe) for his guidance on the analysis and Peter Burge for his suggestions on survey design. Finally, we thank the PACT project consortium partners for their comments on the survey questionnaire. Any errors or omissions herein remain the responsibility of the authors.

REFERENCES

1. Schwartz PH, Caine K, Alpert SA, *et al.* Patient preferences in controlling access to their electronic health records: a prospective cohort study in primary care. *J General Int Med.* 2014. Published Online First: doi: 10.1007/s11606-014-3054-z.
2. Adler-Milstein J, Jha AK. No evidence found that hospitals are using new electronic health records to increase Medicare reimbursements. *Health Affairs.* 2014;33(7):1271–1277.
3. Ioannidis JP. Are mortality differences detected by administrative data reliable and actionable? *JAMA.* 2013;309(13):1410–1411.
4. Robinson N, Potoglou D, Kim CW, *et al.* Security at what cost? Quantifying people's trade-offs across liberty, privacy, and security. 2010. http://www.rand.org/pubs/technical_reports/TR664.html. Accessed July 25, 2014.
5. National Information Board. *Personalised Health and Care* 2020. 2014.
6. NHS Five Year Forward View. 2014.
7. Greenhalgh T, Keen J. "Personalising" NHS information technology in England. *BMJ.* 2014;349:g7341.
8. Gerada C. 10 health apps to help you stay fitter and feel better, 2014 http://www.theguardian.com/technology/2014/feb/08/health-apps-smartphones-top-10-gp-recommends. Accessed September 14, 2015.
9. Pagliari C, Detmer D, Singleton P. Potential of electronic personal health records. *BMJ.* 2007;335(7615):330–333.
10. Department of Health. *Innovation, Health, and Wealth: Adoption and Diffusion of Innovation in the NHS.* 2012.
11. European Directive 95/46/EC on the protection of individuals with regard to the processing of personal data and on the free movement of such data. http://europa.eu/legislation_summaries/information_society/data_protection/l14012_en.htm. Accessed September 14, 2015.
12. Oderkirk J, Ronchi E, Klazinga N. International comparisons of health system performance among OECD countries: opportunities and data privacy protection challenges. *Health Policy.* 2013;112(1-2):9–18.
13. *Reform of Data Proection Legislation.* http://ec.europa.eu/justice/data-protection/. Accessed September 14, 2015.
14. Keen J, Calinescu R, Paige R, *et al.* Big Data + Politics = Open Data: The Case of Health Care Data in England. *Policy Internet.* 2013;5(2):228–243.
15. Nissembaum H. *Privacy in Context. Technology, Policy and the Integrity of Social Life.* Stanford: Stanford Unversity Press, 2010.
16. Sir Nick Partridge. *Review of data releases by the NHS Information Centre.* Location: Health and Social Care Information Centre; 2014. https://www.gov.uk/government/publications/review-of-data-releases-made-by-the-nhs-information-centre.
17. Stone MA, Redsell SA, Ling JT, *et al.* Sharing patient data: competing demands of privacy, trust and research in primary care. *Brit J General Pract.* 2005;55(519):783–789.
18. PACT Dimitris Potoglou; Robinson N, Hellgren T, Kobzar S, Patil S. Literature Review of Approaches for Measuring Preferences with Respect to Privacy, Security and Surveillance. The Privacy & Security Research Paper Series issue 10. 2010. http://www.projectpact.eu/privacy-security-research-paper-series/privacy-security-research-paper-series/PACT_ResearchPapers_10_FINAL.pdf. Accessed September 14, 2015.
19. Grossklags J, Acquisti A. When 25 Cents is too much: An Experiment on Willingness-To-Sell and Willingness-To-Protect Personal Information. 2007. http://weis2007.econinfosec.org/papers/66.pdf. Accessed September 14, 2015.
20. Allenby GM, Shively TS, Yang S, *et al.* A choice model for packaged goods: dealing with discrete quantities and quantity discounts. *Marketing Sci.* 2004;23(1):95–108.
21. 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 Research Methodol.* 2013;13:72.
22. Perera G, Holbrook A, Thabane L, *et al.* Views on health information sharing and privacy from primary care practices using electronic medical records. *Int J Med Informatics.* 2011;80(2):94–101.
23. van Panhuis WG, Paul P, Emerson C, *et al.* A systematic review of barriers to data sharing in public health. *BMC Public Health.* 2014;14(1):1144.
24. Whiddett R, Hunter I, Engelbrecht J, *et al.* Patients' attitudes towards sharing their health information. *Int J Med Informatics.* 2006;75(7):530–541.

Downloaded from http://jamia.oxfordjournals.org/ by guest on May 17, 2016

25. Luchenski SA, Reed JE, Marston C, *et al.* Patient and public views on electronic health records and their uses in the United Kingdom: cross-sectional survey. *J Med Internet Res.* 2013;15(8):e160.
26. Boyle KJ, Bishop RC, Welsh MP. Starting point bias in contingent valuation bidding games. *Land Econ.* 1985;61(2):188–194.
27. Louviere JJ, Hensher DA, Swait JD. *Stated Choice Methods: Analysis and Application.* Cambridge: Cambridge University Press; 2000.
28. PACT D2.2 Methodological report on survey design and survey questionnaire: http://www.projectpact.eu/deliverables/wp3-fieldwork. Accessed September 14, 2015
29. PACT D2.3 Report on the analysis of pilot data. http://www.projectpact.eu/deliverables/wp3-fieldwork. Accessed September 14, 2015
30. Angst CM, Agarwal R. Adoption of electronic health records in the presence of privacy concerns: The elaboration likelihood model and individual persuasion. *MISQuart.* 2009;33:339–370.
31. Hsin-Ginn H, Hwai-En H, Kuang-Ming K, Chung-Feng L. The Differing Privacy Concerns Regarding Exchanging Electronic Medical Records of Internet Users in Taiwan. *J Med Syst.* 2012;36 (6):3783–3793.
32. ChoiceMetrics. *Ngene 1.1.1 User Manual & Reference Guide.* Australia; 2012.
33. PACT D3.1. *Summary of findings from the Survey Testing Phase.* http://www.projectpact.eu/deliverables/wp3-fieldwork. Accessed September 14, 2015
34. Kumaraguru P, Cranor LF. *Privacy Indexes: A Survey of Westin's Studies.* 2005. http://www.cs.cmu.edu/~ponguru/CMU-ISRI-05-138.pdf. Accessed September 14, 2015.
35. Efron B. Bootstrap methods: another look at the Jackknife. *Ann Stat.* 1979;7:1–26
36. IBM Corp. IBM SPSS Statistics for Windows, Version 21.0. Armonk, NY: IBM Corp.; 2012.
37. ALOGIT 4.2. ALOGIT Software & Analysis Ltd. 2007.
38. PACT D 4.2 Public Perception of Security and Privacy: Results of the comprehensive analysis of PACT's pan-European Survey: http://www.projectpact.eu/deliverables/wp3-fieldwork. Accessed September 14, 2015
39. Shaw C, Abrams K, Marteau TM. Psychological impact of predicting individuals' risks of illness: a systematic review. *Soc Sci Med.* 1999; 49(12):1571–1598.
40. Elliott MN, Zaslavsky AM, Goldstein E, *et al.* Effects of survey mode, patient mix, and nonresponse on CAHPS hospital survey scores. *Health Services Res.* 2009;44(2 Pt 1):501–518.
41. Kelsey T. Long live the database state. *Prospect Magazine.* 2009;161. http://www.prospectmagazine.co.uk/features/longlivethedatabasestate. Accessed September 22, 2015.
42. Kahneman D, Tversky A. Prospect theory analysis of decision under risk. *Econometrica.* 1979;47(2):263–291.
43. Pollock AM, Roderick P. Trust in the time of markets: protecting patient information. *Lancet.* 2014;383(9928):1523–1524.
44. Florence C, Shepherd J, Brennan I, *et al.* An economic evaluation of anonymised information sharing in a partnership between health services, police and local government for preventing violence-related injury. *Injury Prevention.* 2014;20(2):108–114.
45. 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;10:1–24.
46. Embryonic stem cells, Francis Collins, and the NIH. *Lancet.* 2009;374(9685):175.
47. Kingman S, Zimmerman D, Dolley M, *et al.* Using animals in medical research. *BMJ.* 1993;306(6884):1019–1023.
48. Sweeney L. *Simple Demographics Often Identify People Uniquely.* Carnegie Mellon University, Data Privacy Working Paper 3. Pittsburgh, 2000.
49. Goldacre B. Care.data is in chaos. It breaks my heart. http://www.theguardian.com/commentisfree/2014/feb/28/care-data-is-in-chaos. 2014 Accessed September 14, 2015.
50. Fernandez-Aleman JL, Senor IC, Lozoya PA, *et al.* Security and privacy in electronic health records: a systematic literature review. *J Biomed Informatics.* 2013;46(3):541–562.
51. O'Dowd A. New e-records system leads to 20% drop in emergency department performance at Addenbrooke's. *BMJ.* 2014;349:g7537.
52. Nuffield Council on Bioethics. The collection, linking and use of data in biomedical research and health care: ethical issues, 2015. http://nuffieldbioethics.org/project/biological-health-data/#sthash.pdXEZZ7t.dpuf http://nuffieldbioethics.org/project/biological-health-data/. Accessed September 14, 2015.
53. Schindler HR, Cave J, Robinson N, *et al.* Europe's policy options for a dynamic and trustworthy development of the Internet of Things SMART 2012/0053, 2012. www.randeurope.org/iot. Accessed September 14, 2015.

AUTHOR AFFILIATIONS

[1]RAND Europe

[2]School of Geography and Planning Cardiff University, Cardiff, UK

RESEARCH AND APPLICATIONS Downloaded from http://jamia.oxfordjournals.org/ by guest on May 17, 2016