

# Public engagement around using clinical free text for research: progress and opportunities

Dr Elizabeth Ford

Senior Lecturer in Primary Care Research

Brighton and Sussex Medical School

University of Sussex, Brighton, UK

Data custodians routinely remove free-text data from datasets provided for research

Little or no understanding of whether the public find sharing their healthcare text acceptable

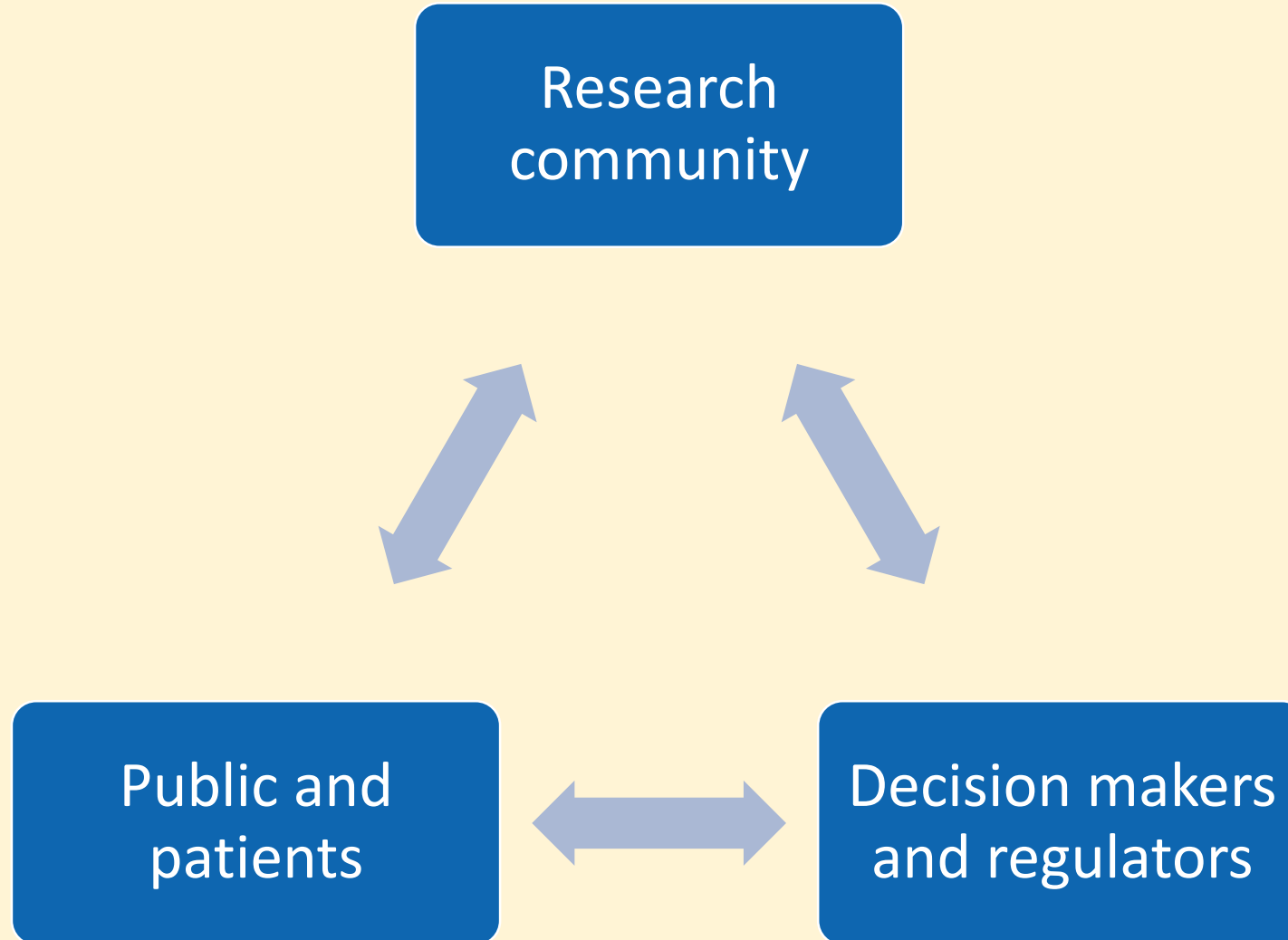
Why?

Access to healthcare text is a major barrier to advance of healthcare text analytics

No research has asked whether the public feel differently about use of structured and unstructured data

Addressing this gap in knowledge of public opinion is a crucial step towards informing and influencing data access policies

# Why involve the public?



# 1: Systematic Review

**Wellcome Open Research**

[SUBMIT YOUR RESEARCH](#)

[BROWSE](#) [GATEWAYS & COLLECTIONS](#) [HOW TO PUBLISH](#) [ABOUT](#) [BLOG](#)


[Home](#) » [Browse](#) » ["Giving something back": A systematic review and ethical enquiry into...](#)


SYSTEMATIC REVIEW

REVISED

**"Giving something back": A systematic review and ethical enquiry into public views on the use of patient data for research in the United Kingdom and the Republic of Ireland [version 2; peer review: 2 approved]**

Previously titled: "Giving something back": A systematic review and

 Check for updates

**ALL METRICS**

2096

 VIEWS

309

 DOWNLOADS

 bright  
me

# Findings of review

- Review found 68-83% of participants willing to share their EHR data for the “common good”.
- Participants worried about hacking/re-identification leading to:
  - identity theft,
  - consequences for employment, pension & benefits eligibility, or insurance costs,
  - social discomfort and community embarrassment,
  - unnecessary stigmatising judgements in clinical settings,
  - the use of EHRs for financial gain
- TRUST was related to organisations’ competence with data handling and motivation for holding data.
- NO STUDIES separated coded data from free text.

## 2. The Healtex Brighton Citizens' Jury June 2018

### Why use citizens' juries?

- Surveys and focus groups matter
- But policy is complex
- Citizens' juries can tell us what people think when more informed and able to talk to their peers
- People often change their minds...
- Unhappy valley
  - “know a little, worry a lot”



# Findings from the Brighton Jury

18 participants representative of England demographics met together for 3 days

**Main question:** “to what degree do you support the use of free text data from patients’ records for health research?”

6 were strongly supportive, 12 were fairly supportive, 0 were neutral/unsupportive

**Case Study** (43 year old man with diabetes and mental health symptoms):

Should the data be shared with the university?

<b>Coded data/ diabetes</b>	Yes = 8	Only if patients can opt-out = 8	Only if patients can opt in = 2	No = 0
<b>Text data / diabetes</b>	Yes = 4	Only if patients can opt-out = 12	Only if patients can opt-in = 2	No = 0
<b>Text data / mental health</b>	Yes = 4	Only if patients can opt-out = 12	Only if patients can opt-in = 2	No = 0

# Jury Recommendations



Transparency of  
data flows,  
processes and uses



Culture of  
continuous  
improvement in  
technology to  
secure privacy and  
increase research  
quality



### 3. Turing Institute Public Event March 2019

- Presentations, sketches and discussions describing: ***What are the special problems with sharing the written text in letters, reports, and clinic notes?***
- Presentation of proposed safeguards
- Panel discussion and group work – request to better understand potential benefits of using free text in research.

The  
Alan Turing  
Institute

“There is enormous potential of, as yet, largely underutilized free-text data for research purposes... it is an atrocious waste”  
*Patient representative on panel.*

Home + Events

Sharing your healthcare data  
safely

Thursday 28 Mar 2019  
Time: 10:30 - 16:00

Agenda ↓

## 4. TexGov

### **Development of Data Governance Standards for Using Clinical Free-Text Data in Health Research – PI Prof Kerina Jones @ Swansea.**

- Outline data protection legislation and regulations relating to free text
- Rapid review of governance models used for free text in UK
- Engagement with public, and research community to explore barriers and solutions.
- Public respondents generally positive about data being used.
- Opt-out mechanisms preferred but are not granular, not specific to free text.
- Importance of operating within existing frameworks for all data – free text not fundamentally different.
- Need to articulate the potential benefits of research using free text.

Original Paper

# Toward the Development of Data Governance Standards for Using Clinical Free-Text Data in Health Research: Position Paper

---

Kerina H Jones<sup>1</sup>, BSc, PhD; Elizabeth M Ford<sup>2</sup>, MA, DPhil; Nathan Lea<sup>3</sup>, BA, MSc, PhD; Lucy J Griffiths<sup>1</sup>, BSc, MSc, PhD; Lamiece Hassan<sup>4</sup>, BSc, PhD; Sharon Heys<sup>1</sup>, BSc; Emma Squires<sup>1</sup>, BSc; Goran Nenadic<sup>5</sup>, BSc, MSc, PhD

---

<sup>1</sup>Population Data Science, Medical School, Swansea University, Swansea, United Kingdom

<sup>2</sup>Brighton and Sussex Medical School, Brighton, United Kingdom

<sup>3</sup>Institute of Health Informatics, University College London, London, United Kingdom

<sup>4</sup>Division of Informatics, Imaging & Data Sciences, University of Manchester, Manchester, United Kingdom

<sup>5</sup>Department of Computer Science, University of Manchester & The Alan Turing Institute, Manchester, United Kingdom

**Corresponding Author:**

Kerina H Jones, BSc, PhD  
Population Data Science  
Medical School  
Swansea University

# Recommendations

- Clearer regulatory guidance based on type of data extraction and location of storage
- Better information for the public
- Use and results of free text data should be publicised to all stakeholders
- Further improvements in deidentification and information extraction NLP
- Identifier-redacted full text data should be treated as still potentially identifiable.
- Governance statements for use of free text to be published
- Consider creating a databank of donated/consented clinical free text.



# 5. Patient Benefits?

Systematic review  
providing a summary and  
taxonomy of possible  
benefits to patients of  
research using free text:

- **Quality or service improvement**
- **Exposure/outcome relationships**
- **Drug prescribing safety**
- **Clinical decision support**
- **Identification for clinical trials**



## The Potential of Research Drawing on Clinical Free Text to Bring Benefits to Patients in the United Kingdom: A Systematic Review of the Literature

Elizabeth Ford<sup>1</sup>, Keegan Curlewis<sup>1</sup>, Emma Squires<sup>2</sup>, Lucy J. Griffiths<sup>2</sup>, Robert Stewart<sup>3,4</sup> and Kerina H. Jones<sup>2\*</sup>

<sup>1</sup> Department of Primary Care and Public Health, Brighton and Sussex Medical School, Brighton, United Kingdom, <sup>2</sup> Swansea Medical School, University of Swansea, Swansea, United Kingdom, <sup>3</sup> King's College London, London, United Kingdom, <sup>4</sup> South London and Maudsley NHS Foundation Trust, London, United Kingdom

### OPEN ACCESS

**Edited by:**  
Aurélien Névelo,  
Université Paris-Saclay, France

**Reviewed by:**

**Background:** The analysis of clinical free text from patient records for research has potential to contribute to the medical evidence base but access to clinical free text is frequently denied by data custodians who perceive that the privacy risks of data-sharing are too high. Engagement activities with patients and regulators, where views on the sharing of clinical free text data for research have been discussed, have identified that stakeholders would like to understand the potential clinical benefits that could be achieved if access to free text for clinical research were improved. We aimed to

# What next?

- Panel discussion at HealTAC 2021 (London/Zoom June 16-18)
  - **“Breaking the deadlock: working towards better access to clinical free text data for research**
- Grant application to Information Commissioners Office for an online tool to produce risk assessments and educational information about text analytics projects
  - Led by Prof Kerina Jones
- Need a plan for how the public can be involved in free text research – as collaborators and research team members.
- Public members on decision making panels – what information do they need about our research?



