Patient Perspective on Privacy & Data Sharing

Andrea Borondy Kitts, MS, MPH
Patient Outreach & Research Specialist
Lahey Hospital & Medical Center
Introduction

- Retired engineer with 32 year career in Aerospace
- Lost husband to lung cancer April 2013
- Lung cancer screening advocate and researcher
- Patient advocate
- Associate editor Journal American College Radiology
Disclosures

COO & Investor *Prosumer Heath*; a start-up company developing an AI driven individualized mobile health platform to help individuals manage their health and healthcare

Associate Editor Journal of American College of Radiology

Paid faculty Medtronic Global Lung Health Summit
Key Considerations, Issues and Ethics in Patient Data Sharing for AI

• Transparency & Trust
• Data Ownership
• Compensation
• Informed Consent
• Privacy
• Use of data; research, commercial
• Data Use Agreements
• Sale to outside parties
• Patient-mediated data sharing
Factors Affecting Data Sharing Preferences

Nested Classification Schema: Factors Influencing Data Sharing Preferences

1. Transparency of Exchange
2. Access & Control
3. Balance Risk & Benefits
4. Harm Threshold
5. Relationships of Trust

Moon, 2015

Who Owns the Data? Who thinks they own the data?


The Elephant in the Room: Paying Patients for Their Data

The assumption is that de-identified data doesn’t belong to patients anymore so they don’t need to get paid & don’t need informed consent.

What is Fair Value? How Will it Be Determined?

71% of Patients Think Consent Should be Required at Least Once

Table 3.
Perceptions of Patients With Cancer Regarding Balancing Needs for Secondary Use of Data for Research and Consent for That Use

<table>
<thead>
<tr>
<th>Need</th>
<th>Percentage (SE) of Patients Rating As Critically or Very Important</th>
<th>Age (years)</th>
<th>Race/Ethnicity</th>
<th>Educational Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>To be able to conduct this kind of medical research</td>
<td>71.65 (2.24)</td>
<td>≥ 65</td>
<td>Black or Hispanic</td>
<td>Higher</td>
</tr>
<tr>
<td></td>
<td></td>
<td>&lt; 65</td>
<td>Nonblack, Non-Hispanic</td>
<td>Lower</td>
</tr>
<tr>
<td>For doctors to get a patient’s permission to use their medical record each time their medical record is used for this kind of research, even if it means that a great deal of research will not be done</td>
<td>35.49 (2.36)</td>
<td>≥ 65</td>
<td>Black or Hispanic</td>
<td>Higher</td>
</tr>
<tr>
<td></td>
<td></td>
<td>&lt; 65</td>
<td>Nonblack, Non-Hispanic</td>
<td>Lower</td>
</tr>
<tr>
<td>For there to be a way to share a patient’s medical records with researchers to do this kind of research without having to ask permission each time</td>
<td>41.30 (2.43)</td>
<td>≥ 65</td>
<td>Black or Hispanic</td>
<td>Higher</td>
</tr>
<tr>
<td></td>
<td></td>
<td>&lt; 65</td>
<td>Nonblack, Non-Hispanic</td>
<td>Lower</td>
</tr>
<tr>
<td>For doctors to ask a patient at least once whether researchers can use their medical record for all future research of this kind</td>
<td>71.29 (2.25)</td>
<td>≥ 65</td>
<td>Black or Hispanic</td>
<td>Higher</td>
</tr>
<tr>
<td></td>
<td></td>
<td>&lt; 65</td>
<td>Nonblack, Non-Hispanic</td>
<td>Lower</td>
</tr>
</tbody>
</table>

NOTE. Complex survey weights applied (details provided in Appendix).

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5501364/
Is Informed Consent Required? What are Exceptions? Are IRB’s Prepared?

Fig. 3. Perception of strategies* to improve comfort with confidentially sharing health data.
*Assigning value to each likert scale response (1–5) & calculating mean (SD) of responses (lower scores = greater improvements in comfort level).

https://doi.org/10.1016/j.ijmedinf.2019.04.00
Coming to America?

The EU General Data Protection Regulation (GDPR) is the most important change in data privacy regulation in 20 years. The regulation will fundamentally reshape the way in which data is handled across every sector, from healthcare to banking and beyond.

Right to Access

Part of the expanded rights of data subjects outlined by the GDPR is the right for data subjects to obtain confirmation from the data controller as to whether or not personal data concerning them is being processed, where and for what purpose. Further, the controller shall provide a copy of the personal data, free of charge, in an electronic format. This change is a dramatic shift to data transparency and empowerment of data subjects.

https://eugdpr.org/
# Privacy

## Social security numbers, health info and phone conversations among the most sensitive data

<table>
<thead>
<tr>
<th>Data Category</th>
<th>Very Sensitive</th>
<th>Somewhat Sensitive</th>
<th>Not Too Sensitive</th>
<th>Not At All Sensitive</th>
</tr>
</thead>
<tbody>
<tr>
<td>Your social security number</td>
<td>90</td>
<td>5</td>
<td>21</td>
<td>0</td>
</tr>
<tr>
<td>State of your health and medications</td>
<td>55</td>
<td>26</td>
<td>12</td>
<td>5</td>
</tr>
<tr>
<td>Content of your phone conversations</td>
<td>54</td>
<td>27</td>
<td>13</td>
<td>4</td>
</tr>
<tr>
<td>Content of your email messages</td>
<td>52</td>
<td>25</td>
<td>13</td>
<td>7</td>
</tr>
<tr>
<td>Details of your physical location over time</td>
<td>50</td>
<td>32</td>
<td>11</td>
<td>5</td>
</tr>
<tr>
<td>Content of your text messages</td>
<td>49</td>
<td>26</td>
<td>13</td>
<td>8</td>
</tr>
<tr>
<td>Numbers you have called or texted</td>
<td>45</td>
<td>30</td>
<td>16</td>
<td>6</td>
</tr>
<tr>
<td>Your birth date</td>
<td>41</td>
<td>25</td>
<td>19</td>
<td>14</td>
</tr>
<tr>
<td>Your relationship history</td>
<td>40</td>
<td>31</td>
<td>14</td>
<td>12</td>
</tr>
<tr>
<td>Websites you have visited</td>
<td>27</td>
<td>43</td>
<td>20</td>
<td>8</td>
</tr>
<tr>
<td>Searches you have made using search engines</td>
<td>24</td>
<td>41</td>
<td>22</td>
<td>10</td>
</tr>
<tr>
<td>Your religious and spiritual views</td>
<td>22</td>
<td>23</td>
<td>29</td>
<td>25</td>
</tr>
<tr>
<td>Your friends and what they are like</td>
<td>22</td>
<td>46</td>
<td>23</td>
<td>7</td>
</tr>
<tr>
<td>Your political views and the candidates you support</td>
<td>20</td>
<td>31</td>
<td>30</td>
<td>17</td>
</tr>
<tr>
<td>The media you like</td>
<td>9</td>
<td>22</td>
<td>45</td>
<td>21</td>
</tr>
<tr>
<td>Your basic purchasing habits</td>
<td>8</td>
<td>33</td>
<td>44</td>
<td>14</td>
</tr>
</tbody>
</table>


[PEW RESEARCH CENTER](https://www.pewresearch.org/fact-tank/2016/09/21/the-state-of-privacy-in-america/)
Challenges to De-identification

• Aggregated Information required for algorithm development - images, pathology, genomics, social determinants of health
• Secondary use and aggregation with other data sets
Privacy - Is it Possible to De-identify Data?

- Distributed learning – data stays at originating institution
- Transparency on risk of breach
- Transparency and verification on use (Deep Mind Smart Auditing)

https://www.wired.co.uk/article/ai-healthcare-gp-deepmind-privacy-problems
Sharing health information

A new health information website is being used by your doctor’s office to help manage patient records. Your participation would allow you to have access to your own health records and make scheduling appointments easier. If you choose to participate, you will be allowing your doctor’s office to upload your health records to the website and the doctor promises it is a secure site.

Would this scenario be acceptable to you, or not?

Note: Refused responses not shown.
PEW RESEARCH CENTER

https://www.pewinternet.org/2016/01/14/scenario-health-information-convenience-and-security/
### Level of Concern about Potential Consequences of Data Sharing.

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Very concerned</th>
<th>Somewhat concerned</th>
<th>Not very concerned</th>
<th>Not at all concerned</th>
</tr>
</thead>
<tbody>
<tr>
<td>It could be harder to get people to agree to be in clinical trials if they know their data will be shared</td>
<td>9.5</td>
<td>27.1</td>
<td>31.4</td>
<td>32.0</td>
</tr>
<tr>
<td>Companies might use the information for marketing purposes instead of scientific purposes</td>
<td>11.2</td>
<td>22.6</td>
<td>32.0</td>
<td>34.2</td>
</tr>
<tr>
<td>My information might be stolen</td>
<td>12.9</td>
<td>17.6</td>
<td>33.8</td>
<td>35.7</td>
</tr>
<tr>
<td>Scientists and companies might have less incentive to invest time and money in doing clinical trials</td>
<td>8.7</td>
<td>19.3</td>
<td>33.2</td>
<td>38.8</td>
</tr>
<tr>
<td>Someone who is good with computers could identify me</td>
<td>7.1</td>
<td>19.2</td>
<td>35.0</td>
<td>38.7</td>
</tr>
<tr>
<td>Scientists or companies could unfairly “free ride” on the work of others</td>
<td>8.2</td>
<td>16.6</td>
<td>35.5</td>
<td>39.7</td>
</tr>
<tr>
<td>People might use the data to do poor-quality science</td>
<td>8.4</td>
<td>15.4</td>
<td>32.1</td>
<td>44.1</td>
</tr>
<tr>
<td>I could be discriminated against if the information was linked back to me</td>
<td>6.6</td>
<td>14.9</td>
<td>27.9</td>
<td>50.7</td>
</tr>
<tr>
<td>Some person or company could make a lot of money developing products using my information</td>
<td>6.1</td>
<td>14.0</td>
<td>32.3</td>
<td>47.7</td>
</tr>
<tr>
<td>My information might be used in scientific projects that I wouldn’t approve of</td>
<td>6.3</td>
<td>12.3</td>
<td>34.6</td>
<td>46.8</td>
</tr>
<tr>
<td>It could be embarrassing if the information was linked back to me</td>
<td>4.7</td>
<td>9.5</td>
<td>28.6</td>
<td>57.2</td>
</tr>
</tbody>
</table>
**Table 2. Willingness of Clinical Trial Participants to Share Their Data, According to Type of Use and Recipient.**

<table>
<thead>
<tr>
<th>Type of Use or Recipient</th>
<th>Very Likely</th>
<th>Somewhat Likely</th>
<th>Neither Likely nor Unlikely</th>
<th>Somewhat Unlikely</th>
<th>Very Unlikely</th>
</tr>
</thead>
<tbody>
<tr>
<td>To help patients and groups of patients learn more about health problems that affect them</td>
<td>77.8</td>
<td>18.8</td>
<td>2.6</td>
<td>0.4</td>
<td>0.4</td>
</tr>
<tr>
<td>To do research on health problems that affect my family or me</td>
<td>78.3</td>
<td>17.1</td>
<td>3.2</td>
<td>1.1</td>
<td>0.5</td>
</tr>
<tr>
<td>To do research that will help others</td>
<td>79.9</td>
<td>17.1</td>
<td>2.0</td>
<td>0.4</td>
<td>0.7</td>
</tr>
<tr>
<td>To help get answers to scientific questions faster using information that others have already gathered</td>
<td>72.2</td>
<td>22.6</td>
<td>3.4</td>
<td>1.2</td>
<td>0.5</td>
</tr>
<tr>
<td>To help scientists check the accuracy of research results announced by other scientists or companies (by redoing the analyses)</td>
<td>70.9</td>
<td>22.6</td>
<td>3.8</td>
<td>1.5</td>
<td>1.2</td>
</tr>
<tr>
<td>To learn more about diseases that only a small number of people have (by combining data from many clinical trials)</td>
<td>69.1</td>
<td>22.1</td>
<td>5.9</td>
<td>1.7</td>
<td>1.2</td>
</tr>
<tr>
<td>To help lawyers prove their case in lawsuits claiming that medical products are unsafe</td>
<td>27.9</td>
<td>24.5</td>
<td>26.9</td>
<td>12.7</td>
<td>8.0</td>
</tr>
<tr>
<td>Recipient</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scientists in universities and other not-for-profit organizations</td>
<td>69.2</td>
<td>24.0</td>
<td>3.3</td>
<td>1.7</td>
<td>1.8</td>
</tr>
<tr>
<td>Scientists in companies developing medical products, such as prescription drugs</td>
<td>53.4</td>
<td>28.5</td>
<td>10.6</td>
<td>5.4</td>
<td>2.1</td>
</tr>
</tbody>
</table>

* Shown are the responses to items worded as “How likely would you be to allow your anonymous, individual clinical trial data to be used in the following ways?” (for type of use) or “How likely would you be to allow your anonymous, individual clinical trial data to be shared with . . . .” (for recipient). Numbers were rounded to the nearest tenth. The accuracy (95% confidence interval) of percentages close to 50% is ±3.6 percentage points, diminishing to ±2.2 for percentages close to 10%.
Will Patients Want to Share Data for Commercial Uses?

https://www.pewinternet.org/2015/05/20/americans-attitudes-about-privacy-security-and-surveillance/
Data Use Agreements with Patients?

• How is the Data Being Used? For How Long?
• Will Data Be Reused? Will it Be Aggregated with Other Data?
• Exclusivity
• Compensation
• What Happens to Data if Patients Decide to Opt Out of Study/Venture?
Will Patients Sell Their Data to Commercial Vendors?

Dollars for Donating

$1000 DONATE SPERM

WHO PAYS THE MOST?
* Seattle Sperm Bank $70 per donation.
* Massachusetts New England Cryogenic $100 per donation.
* The Sperm Bank of California $60 per donation.
* Washington Northwest Andrology & Cryobank Inc. $1000 per month.
* European Sperm Bank $35 per donation plus a bonus.
* The London Women’s Clinic £35 per donation.

Visit Co-ParentMatch.com
Find women who are searching for sperm donors right now!

Images & Medical Data
Patient Mediated Data Sharing

- No standard process or protocol
- No central secure data repository
- Self-selection may lead to algorithm bias
  - Differences in trust in research by race/ethnicity
- Need education for patients on importance of research & sharing their data
- Crowd sourcing approaches

HealthIT.gov Artificial intelligence for health and health care. Available at: https://www.healthit.gov/sites/default/files/jsr-17-task-002_aiforhealthandhealthcare12122017.pdf
Accessed June 6, 2019

https://pubs.rsna.org/doi/full/10.1148/radiol.2019190613
Patient Partners in Lung Cancer Research – Targeted crowdsourcing? Scalable Model for AI Algorithm Development?

LUNGevity Partners with ALK Positive Patient Group to Fund Lifesaving Research

Collaboration for patient-driven research that has the potential to save their lives

FOR IMMEDIATE RELEASE

Media Contact:
Linda Wenger
ljwenger@LUNGevity.org
(973) 649-3214

WASHINGTON, DC [November 15, 2017] – LUNGevity Foundation is honored to announce that it is partnering with the ALK Positive patient group to support the new ALK-Positive Lung Cancer Transformational Research Award Program. ALK Positive members are the first group of ALK-positive patients to influence the direction of research into their specific mutation that will, one day, save their lives. The ultimate goal of the program is to transform ALK-positive lung cancer into a chronic or curable condition.

ALK Positive is a group of 700+ lung cancer patients, and caregivers of these patients, whose cancer has tested positive for a mutation in the anaplastic lymphoma kinase (ALK) gene. ALK-positive patients account for approximately 4% of patients with non-small cell lung cancer. The group was initially formed on Facebook, by Tom and Merita Carroll, and has quickly grown into its current proactive form with their direct funding of research.

“We have members who are on their last treatment option, and they are wondering what will happen when this treatment stops working,” explains Laura Green, the leader of ALK Positive’s Executive Outreach Committee.

Purpose

The EGFR Resisters is a group of people living with and/or personally affected by EGFR positive lung cancer. The purpose of founding the group is to build a community of survivors and caregivers to share knowledge, provide support, and confidentially collect patient data.

https://ros1cancer.com/ros1-patient-driven-research/

https://www.cancer.gov/about-nci/organization/ccg/blog/2017/ros1-patient-driven-research


https://egfrcancer.org/
The New Deal on Data to Enable Open Information Markets

Individual Ownership of Data
- Right to Possess Your Data
- Full Control Over the Use of Your Data
- Right to Dispose of or Distribute Your Data

Policies to encourage combination of massive amounts of anonymous data to promote the Common Good

https://hd.media.mit.edu/wef_globalit.pdf
Future => One Patient, One Record Owned by Patient? Helping Facilitate Data Sharing

“As mobile devices become more powerful, healthcare professionals envision a world where the patients own and store the data on their devices, leaving health institutions responsible to create a system where the data can be anonymized, shared, and exchanged.”

Additional Benefit: Improved Data Quality

Summary

• Transparency critical for gaining trust for patients to feel comfortable sharing their data
  • Transparency needs to extend to all uses and reuses of patient data & risk of privacy breach
• Need to develop an approach for valuation and fair compensation to patients for data that is sold for profit by the controlling institution
• Frameworks for broad, population level consent
• Patient mediated data sharing in the future