Ethical data use, Good data governance

Nicki Tiffin
nicki.tiffin@uct.ac.za
Governance structures:

- SOPs for data access
- Data protection
- Documentation
- Ethical oversight
The index card of good data governance

With a nod to Harold Pollack....
KEY POINTS FOR GOOD DATA GOVERNANCE

1. The study participant knows what I am using their health data for, and is ok with it
Informed consent process:

Explicit
- Which data
- Where are the data collected
- What time frame

Simple language
- Research question
- Opt-in, not opt-out
- Withdrawal from study at any time
- No negative consequences of declining
KEY POINTS FOR GOOD DATA GOVERNANCE

1. The study participant knows what I am using their health data for, and is ok with it

2. The ethics board knows what health data I am using, and how; and is ok with it
GOVERNANCE – Ethics

- Confidentiality
- Beneficence
- Potential harms
- Vulnerable populations
  
  *Poor health, low SES*

DoH Research approval
HRECS and Ethics approval
Informed consent/anonymisation/aggregation
KEY POINTS FOR GOOD DATA GOVERNANCE

1. The study participant knows what I am using their health data for, and is ok with it
2. The ethics board knows what health data I am using, and how; and is ok with it
3. This data use complies with legislation (POPI, National Health Act)
GOVERNANCE – legal compliance

**POPI** Protection of Personal Information act

Responsible Party: DoH

Primary purpose: data collected for provision of health care

**PAIA** Promotion of access to information act

Record keeping of how individuals’ data are used

- records of all data access
<table>
<thead>
<tr>
<th>KEY POINTS FOR GOOD DATA GOVERNANCE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. The study participant knows what I am using their health data for, and is ok with it</td>
</tr>
<tr>
<td>2. The ethics board knows what health data I am using, and how; and is ok with it</td>
</tr>
<tr>
<td>3. This data use complies with legislation (POPI, Healthcare Act, in SA)</td>
</tr>
<tr>
<td>4. All datasets and drives are password protected and encrypted</td>
</tr>
</tbody>
</table>
Easy software solutions, e.g.

7zip
Password protection and AES encryption for all data files

Excel (RedCAP or MySQL is better!)
Spreadsheets with password protection

Bitlocker
Encrypt all drives where data are stored, including flash drives for data transfer
Ask your departmental admin to file codes securely
**KEY POINTS FOR GOOD DATA GOVERNANCE**

1. The study participant knows what I am using their health data for, and is ok with it.
2. The ethics board knows what health data I am using, and how; and is ok with it.
3. This data use complies with legislation (POPI, Healthcare Act, in SA).
4. All datasets and drives are password protected and encrypted.
5. All participant identifiers are stored and transferred in a separate file to any clinical data.
Separation of identifying and clinical data

**File 1**
Medical and clinical data linked to study ID

**File 2**
Patient identifying data linked to study ID
KEY POINTS FOR GOOD DATA GOVERNANCE

1. The study participant knows what I am using their health data for, and is ok with it
2. The ethics board knows what health data I am using, and how; and is ok with it
3. This data use complies with legislation (POPI, Healthcare Act, in SA)
4. All datasets and drives are password protected and encrypted
5. All participant identifiers are stored and transferred in a separate file to any clinical data
6. All anonymised data cannot be re-identified
Can data be easily re-identified?

**e.g. Genomic data**
Anonymising genomic data is like trying to anonymise a fingerprint

**e.g. Clinical data**
A 37-year old lady with epilepsy - with a 9yr old daughter and a 5yr old son and frequently attending a particular clinic - can be easily re-identified.

Do not manage de-identified data in the same way as truly anonymised data, because it can be re-identified.
<table>
<thead>
<tr>
<th>KEY POINTS FOR GOOD DATA GOVERNANCE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. The study participant knows what I am using their health data for, and is ok with it</td>
</tr>
<tr>
<td>2. The ethics board knows what health data I am using, and how; and is ok with it</td>
</tr>
<tr>
<td>3. This data use complies with legislation (POPI, Healthcare Act, in SA)</td>
</tr>
<tr>
<td>4. All datasets and drives are password protected and encrypted</td>
</tr>
<tr>
<td>5. All participant identifiers are stored and transferred in a separate file to any clinical data</td>
</tr>
<tr>
<td>6. All anonymised data cannot be re-identified</td>
</tr>
<tr>
<td>7. I never use email to transfer identifiable health information, and I always send passwords separately (not by email).</td>
</tr>
</tbody>
</table>
**Passwords**

**Never email passwords**
Use sms or telephone call

**Never send passwords together with data**
Just don’t.

**UCT: Use filesender (filesend.uct.ac.za) to send password-protected, encrypted files with sensitive data**
**KEY POINTS FOR GOOD DATA GOVERNANCE**

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>The study participant knows what I am using their health data for, and is ok with it</td>
</tr>
<tr>
<td>2.</td>
<td>The ethics board knows what health data I am using, and how; and is ok with it</td>
</tr>
<tr>
<td>3.</td>
<td>This data use complies with legislation (POPI, Healthcare Act, in SA)</td>
</tr>
<tr>
<td>4.</td>
<td>All datasets and drives are password protected and encrypted</td>
</tr>
<tr>
<td>5.</td>
<td>All participant identifiers are stored and transferred in a separate file to any clinical data</td>
</tr>
<tr>
<td>6.</td>
<td>All anonymised data cannot be re-identified</td>
</tr>
<tr>
<td>7.</td>
<td>I never use email to transfer identifiable health information, and I always send passwords separately (not by email).</td>
</tr>
<tr>
<td>8.</td>
<td>If I were a study participant I would be happy with the way my personal health data are being used.</td>
</tr>
</tbody>
</table>
### KEY POINTS FOR GOOD DATA GOVERNANCE

1. The study participant knows what I am using their health data for, and is ok with it
2. The ethics board knows what health data I am using, and how; and is ok with it
3. This data use complies with legislation (POPI, Healthcare Act, in SA)
4. All datasets and drives are password protected and encrypted
5. All participant identifiers are stored and transferred in a separate file to any clinical data
6. All anonymised data cannot be re-identified
7. I never use email to transfer identifiable health information, and I always send passwords separately (not by email).
8. If I were a study participant I would be happy with the way my personal health data are being used.
Thank you