Digital health data governance

Nicki Tiffin
SANBI
University of the Western Cape
South Africa

ntiffin@uwc.ac.za

Rapidly expanding volumes of health data

- Increased data volumes
- Increased opportunity for evidence-based decision making
- Increased risk of data breach, privacy violation
- Vulnerable populations are disadvantaged in negotiating data protection

Tiffin N, George A, LeFevre AE. **How to use relevant data for maximal benefit with minimal risk: digital health data governance to protect vulnerable populations in low-income and middle-income countries.** BMJ Glob Health. 2019 Apr 11;4(2):e001395

Analysis

BMJ Global Health

How to use relevant data for maximal benefit with minimal risk: digital health data governance to protect vulnerable populations in low-income and middle-income countries

Nicki Tiffin,[1,2,3] Asha George,[4] Amnesty Elizabeth LeFevre[5,6]

To cite: Tiffin N, George A, LeFevre AE. How to use relevant data for maximal benefit with minimal risk: digital health data governance to protect vulnerable populations in low-income and middle-income

ABSTRACT

Globally, the volume of private and personal digital data has massively increased, accompanied by rapid expansion in the generation and use of digital health data. These technological advances promise increased opportunity for data-driven and evidence-based health programme

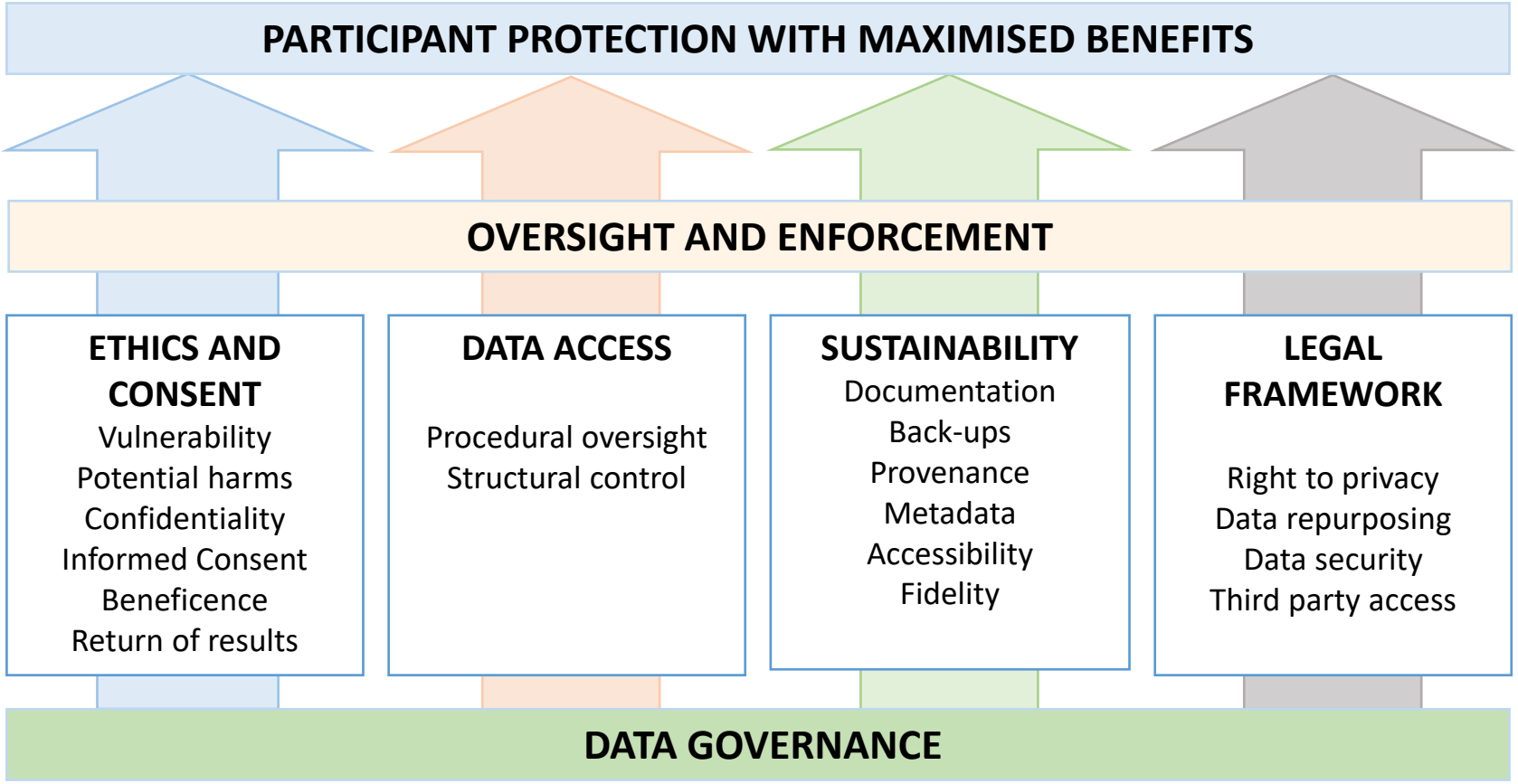
Summary box

- Digital health data provide both opportunities for benefit and risks for vulnerable populations. We propose a data governance framework that can both reduce the risk of digital health data misuse while

BMJ Glob Health: first published as 10.1136/bmjgh-2019-001395

Nicki Tiffin

A framework to guide governance



Pillars of data governance

1. Ethics and informed consent
2. Legislation
3. Data Access Controls
    - Procedural
    - Structural
4. Sustainability

Pillars of data governance

1. Ethics and informed consent
2. Legislation
3. Data Access Controls
    - Procedural
    - Structural
4. Sustainability

“Ethics” is much more than a review board

PARTICIPANT
COMMUNITY

It is about protecting the participant and their community,

and ensuring that the benefits of research outweigh their exposure to risk.

Not just: “Can I get to do the cool research I want to do?”

Protecting the participant

- Confidentiality
- Beneficence
- Potential harms
- Vulnerable populations

  Poor health, low SES, limited access to care

DoH Research approval

HRECS and Ethics approval

Informed consent/anonymisation/aggregation

De-identification is not equivalent to anonymisation

Protecting the community

Balancing benefits and risks

- Understanding what creates risks
- Understanding what is considered a benefit

Local context and local understanding is essential

- Local consultation with stakeholders from planning stages
- Review of proposal by a local Ethics Review Board
- Understand what is culturally acceptable/not acceptable
- Sensitive reporting practices to mitigate risks
- Agree appropriate informed consent protocols
- Agree how data and biospecimens will be shared

Participant information and informed consent

Consent models:

Broad consent

- One consent for all possible future use cases
- Oversight should be provided by an access committee

Tiered consent

- Multiple questions and consents about different levels of data use
  This study, other studies on this disease, other health studies, all other use

Dynamic consent

- Online platform describing ongoing research
- Tracking by participant of data/sample use
- Approve or withhold consent for each new study
  Limited by smartphone, internet access, digital literacy, not feasible in Africa yet

Respect for autonomy

- Ensure that research participants know exactly how their data and biospecimens will be used, including onward sharing
- Provide opportunities for participants to select their own level of comfort with onward use

Tiered consent

"I agree for my samples and data to be used in this diabetes study" YES/NO

"I agree for my samples and data to be used in other diabetes studies" YES/NO

"I agree for my samples and data to be used in other health-related studies" YES/NO

[…]

Electronic recording keeping of consent tiers: identify what onward use each participant agreed to

Respect for autonomy

- Return of results (actionable/unactionable)
- Re-contact for recruitment into future studies

Population diversity/ ancestry studies

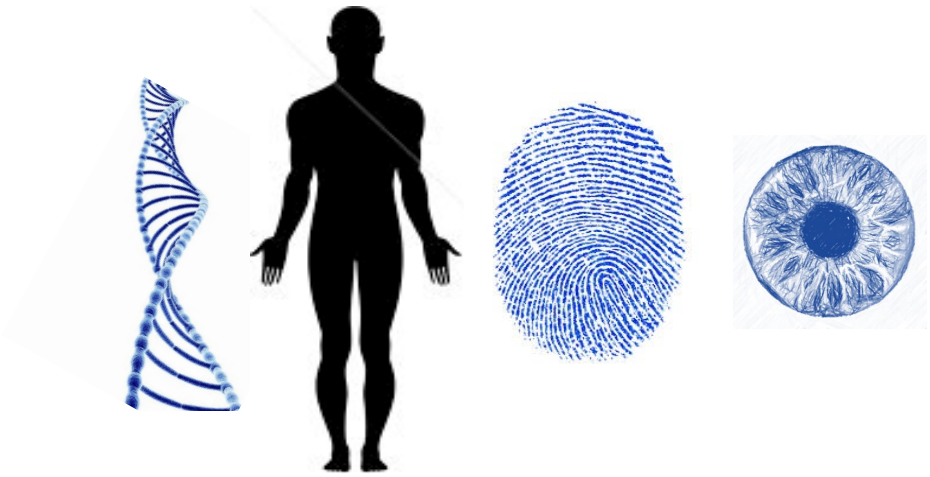
- Separate consent
- Different potential harms – family, community origins

Right to privacy

- Confidentiality

- Anonymisation/de-identification
  - DNA cannot be anonymised
  - Granular, longitudinal data – re-identification risk

- Data perturbation

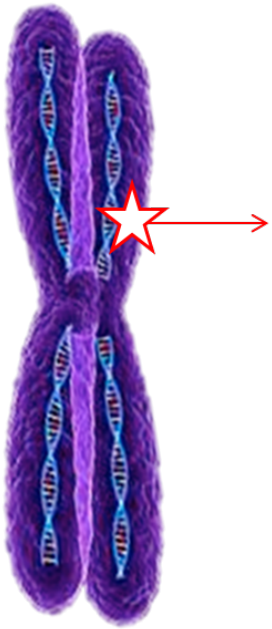


Nicki Tiffin

Right to privacy

We do not yet know what potential harms might come from sequencing human genomes

The genomic sequence of an individual does not change, but what we can read from it will change in ways we do not yet understand.



1972: Junk DNA

1980: Non-coding region

1981: Regulatory Sequence

2010: Enhancer

2013: Super Enhancer

2015: Implicated in autoimmune disease

2025: ??

Nicki Tiffin

Protection from harms

Personal harms – stigma, discrimination

- Re-identification
- Data breach, sensitive clinical data

Community harms

- African populations are genetically very distinct
- Genotypes can be associated with specific groups

(In)famous examples: Havasupai Native Americans

San in South Africa

Community engagement

- Respecting autonomy

- Ensuring acceptability of research and intended activities

Ask, and discuss, whether research outputs might expose the local community to risks of stigma or negative reporting, and whether survey questions and biological samples are culturally acceptable

- Promoting global equity

- Equitable participation by all local stakeholders, not only potential participants

Involve local researchers, local leaders, local institutions and biotech, local service providers e.g. health care service

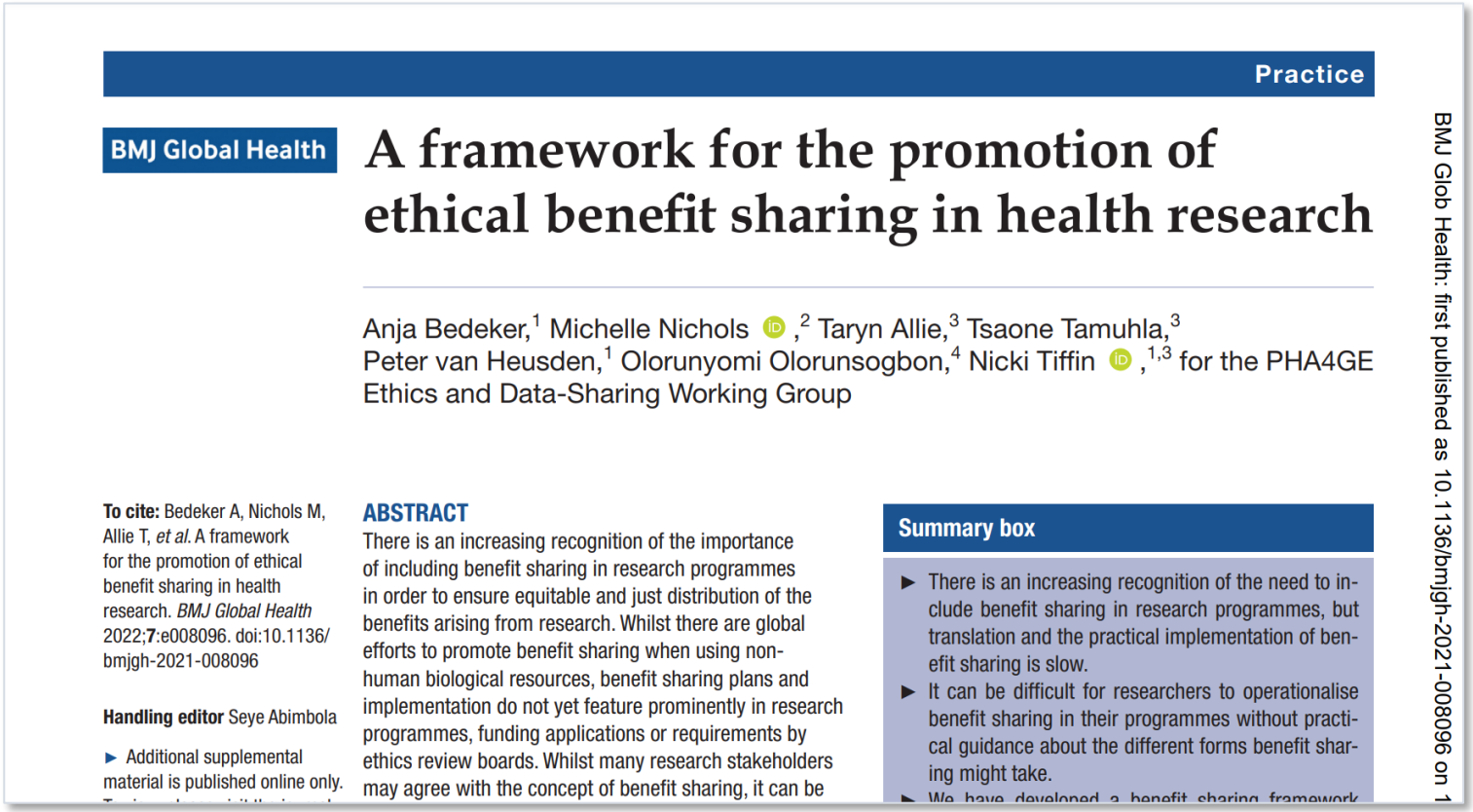
- Discuss and agree how data and biospecimens will be shared

Checklist

- ✓ Does the population under study benefit from this research?
- ✓ Does the population under study face unreasonable risks from this research?
- ✓ Are researchers from the population under study properly involved in leading, conducting AND reporting the research?
- ✓ Do you know what this population thinks about your research plan and findings?
- ✓ What do local people say about risks from this research, especially denigration/ labelling/ stigma/ xenophobia/ ethnic violence/ unwanted attention?
- ✓ Does it add value to identify the specific population in your reporting? Would a regional or national identity be sufficient?

Benefit sharing

A framework for the promotion of ethical benefit sharing in health research
The PHA4GE Ethics and Data Sharing Working Group
BMJ Global Health 2022;**7**:e008096.



Practice

BMJ Global Health

A framework for the promotion of ethical benefit sharing in health research

Anja Bedeker,[1] Michelle Nichols ,[2] Taryn Allie,[3] Tsaone Tamuhla,[3] Peter van Heusden,[1] Olorunyomi Olorunsogbon,[4] Nicki Tiffin ,[1,3] for the PHA4GE Ethics and Data-Sharing Working Group

To cite: Bedeker A, Nichols M, Allie T, *et al*. A framework for the promotion of ethical benefit sharing in health research. *BMJ Global Health* 2022;**7**:e008096. doi:10.1136/bmjgh-2021-008096

Handling editor Seye Abimbola

► Additional supplemental material is published online only.

ABSTRACT

There is an increasing recognition of the importance of including benefit sharing in research programmes in order to ensure equitable and just distribution of the benefits arising from research. Whilst there are global efforts to promote benefit sharing when using non-human biological resources, benefit sharing plans and implementation do not yet feature prominently in research programmes, funding applications or requirements by ethics review boards. Whilst many research stakeholders may agree with the concept of benefit sharing, it can be

Summary box

- There is an increasing recognition of the need to include benefit sharing in research programmes, but translation and the practical implementation of benefit sharing is slow.
- It can be difficult for researchers to operationalise benefit sharing in their programmes without practical guidance about the different forms benefit sharing might take.
- We have developed a benefit sharing framework

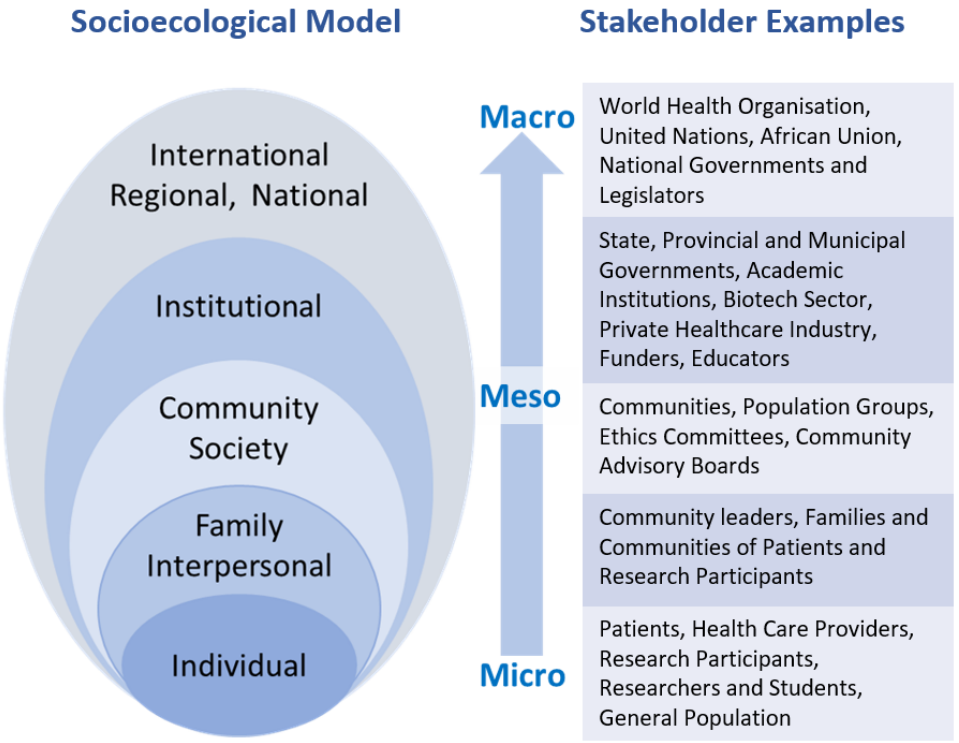
Nicki Tiffin

Benefit sharing

- Build in mechanisms from proposal writing stage to ensure benefit sharing from onward use of resources
- Include benefit-sharing discussions in community engagement processes

Types of benefits

Financial
Health and wellbeing
Infrastructure
Equipment
Skills capacity
Knowledge
Services capacity
Career development
Attribution and recognition



Integration with health service delivery

- Increasing capacity for collection of electronic medical data
- Research partnerships with health services to support data collection and management for health service provision
- With appropriate consent in place, these data can also support the research enterprise
- Avoid building parallel data economies
- Avoid expatriating valuable health data from research, ensure it also benefits patients through the health services
- Avoid the loss of health data once research programmes end

Data Centre Profile: The Provincial Health Data Centre of the Western Cape Province, South Africa. Int J Popul Data Sci. 2019 Nov 20;4(2):1143. doi: 10.23889/ijpds.v4i2.1143. Boulle A, *et al.*

Nicki Tiffin

Scientific Citizenship

- Engaging with public audiences in a relatable way

  Popular press and radio
  Online media
  Social media
  Talks
  School programmes

- Increasing familiarity with the scientific process and benefits of research
- Decreasing public anxiety about sharing of data and biospecimens
- Increasing public awareness about rights and protections around use of data and biospecimens

Career development and mentoring

- It's a form of benefit sharing from research
- Often early- or mid-career researchers can make good use of existing samples and data to start building their research programmes, but need to know how to do this appropriately
- Ensure that researchers in training are exposed to issues around data and biospecimen sharing and governance
- Provide training and workshops to ensure that we all learn how to effectively and ethically share data and biospecimens

Nicki Tiffin

Our publications that might be useful

Tiered informed consent: respecting autonomy, agency and individuality in Africa. Tiffin N BMJ Glob Health. 2018 Dec 17;3(6):e001249. doi: 10.1136/bmjgh-2018-001249,

A framework for tiered informed consent for health genomic research in Africa, Nembaware *et al.* 2019 Nature Genetics, Nov;51(11):1566-1571. doi: 10.1038/s41588-019-0520-x.

Patient-centric research in the time of COVID-19: conducting ethical COVID-19 research in Africa, Nembaware *et* al BMJ Glob Health, 2020 Aug;5(8):e003035. doi: 10.1136/bmjgh-2020-003035.

RedCap eConsent Template for tiered informed consent*:* Build your own electronic participant information and tiered consent process in RedCap using the template. Tamuhla et al. *BMC Med Ethics. 2022 Nov 24;23(1):119. doi: 10.1186/s12910-022-00860-2*

Population-level risks from aggregate genomic data: Potential risks and solutions for sharing genome summary data from African populations. Tiffin N. *BMC Med Genomics. 2019 Nov 4;12(1):152. doi: 10.1186/s12920-019-0604-6.*

Population-level risks from aggregate genomic data: Potential risks and solutions for sharing genome summary data from African populations. Bedeker et al. *BMJ Glob Health. 2022 Feb;7(2):e008096. doi: 10.1136/bmjgh-2021-008096.*

Pillars of data governance

1. Ethics and informed consent
2. Legislation
3. Data Access Controls
    - Procedural
    - Structural
4. Sustainability

Legislation and legal compliance

National and regional legislation

- Protection of personal information Act
- Health Act
- Human rights charters and laws
- Constitutional law

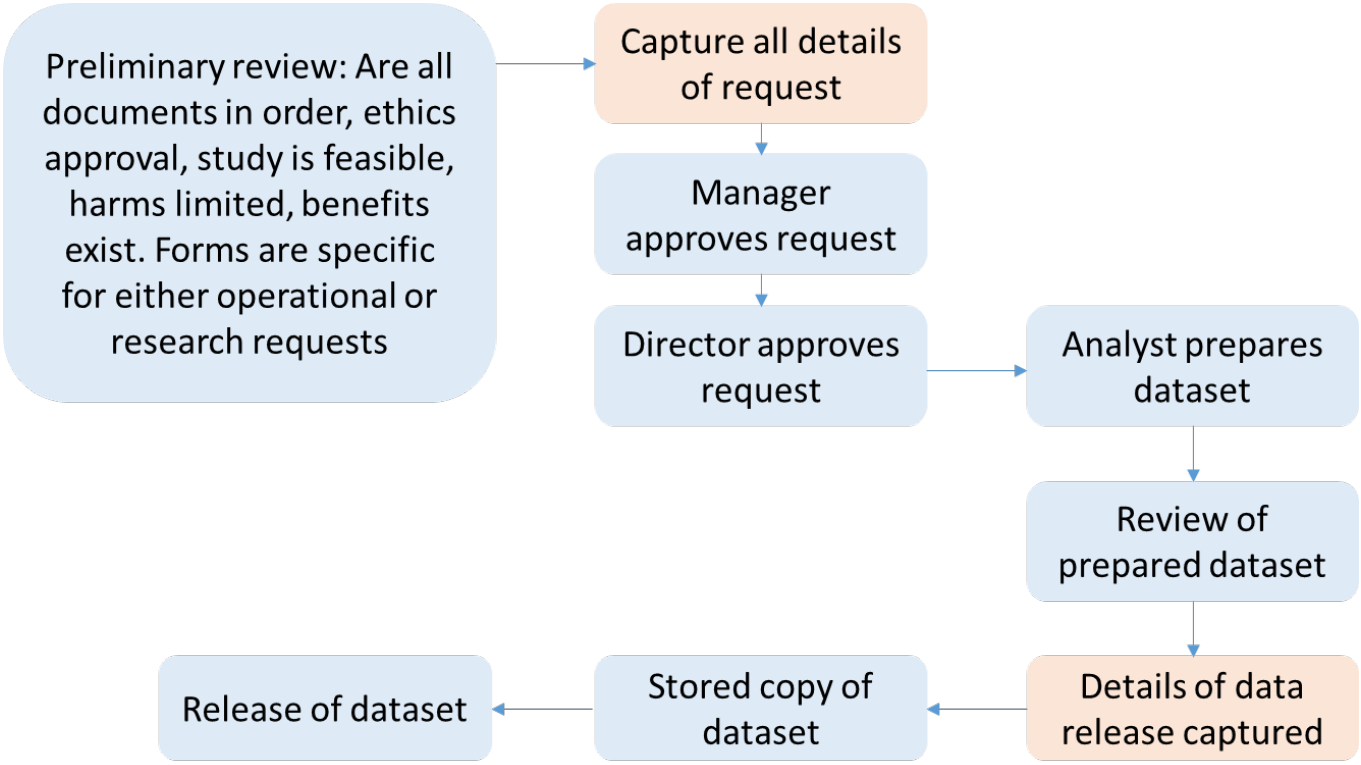
Pillars of data governance

1. Ethics and informed consent
2. Legislation
3. Data Access Controls
    - Procedural
    - Structural
4. Sustainability

Access control - procedural

- Systematic processing of data access requests.
- Detailed SOPs governing data access, use and management.
- Logging of all identifiers used in all studies
- Logging all queries and use of data
- Systematic decisions who can view and use which data

Example: research data request process



Data sharing infrastructure

- Protocols and SOPs describe data sharing agreements, collaborative agreements, benefit-sharing agreements

- We need ways to make these processes accessible and manageable for researchers sharing data and biospecimens

- We need to address barriers to data sharing:

- Difficulty accessing guidelines and ‘how to’ resources
- Resistance to centralising resources
- Difficulty conforming to data formats and specimen prep
- Flexibility needed depending on consent and legal restrictions

Pillars of data governance

1. Ethics and informed consent
2. Legislation
3. Data Access Controls
    - Procedural
    - Structural
4. Sustainability

Access control, structural protection

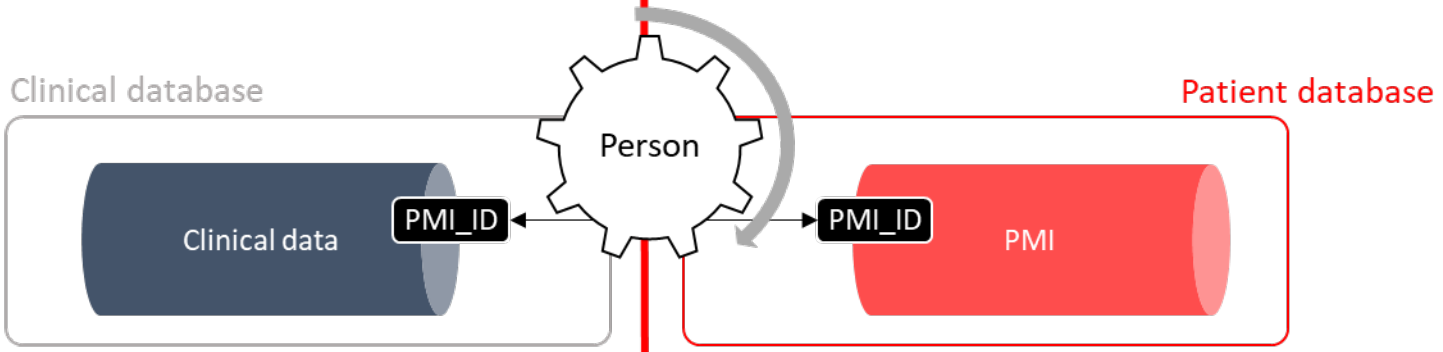
- Database- and field- specific access restrictions
- Firewalls, passwords
- Data transfer protocols
  - Use purposed platforms e.g. liftserver
  - No emailing of sensitive information
  - Password protection
  - Encryption (e.g. bitlocker, 7zip)
  - Separate file sends for clinical and identifying data
- Separation of identifying and attribute data

Nicki Tiffin

Separating identifying and sensitive data

Analytic uses: No need for identifying data

Linking / matching: No need for clinical data



Analysis

Anonymous Identifier_A

Clinical data

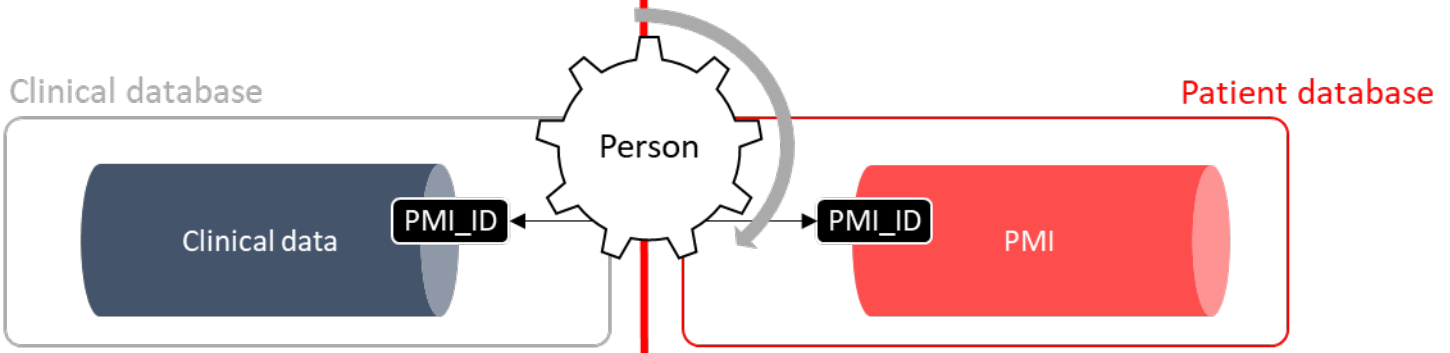
(visit,lab,drug etc)
Date-time
Facility
Evidence
Visit type
Diagnosis

Nicki Tiffin

Separating identifying and sensitive data

Analytic uses: No need for identifying data

Linking / matching: No need for clinical data



Age, sex, possibly suburb

Anonymous Identifier_B

Names, IDs, Folder#, DOB, Address, Contacts

Linkage to patient

Name, date of birth, date of death, sex, address, phone no.

Nicki Tiffin

Pillars of data governance

1. Ethics and informed consent
2. Legislation
3. Data Access Controls
    - Procedural
    - Structural
4. Sustainability

Sustainability

- **Documentation**
- Database structure
- Data processing protocols
- Changes to data structure and content
- SOPs (Data access process, data handling, etc)
- Knowledgebase not reliant on individuals
- Deployment-ready for implementation elsewhere

Nicki Tiffin

Sustainability

- **Standardisation**
- Use data standards
- Requires upkeep of fields and meta-data (code books)
- Increases accessibility
- Ensures secondary use is feasible
- Keeps data management person-agnostic

Oversight and enforcement

Oversight by

- Data access committees
- Ethics committees

Enforcement by

- Legal infrastructure

  e.g. POPI regulator (SA)

  Data Protection Officers (EU)

How can we do data sharing?

How can we do data sharing?

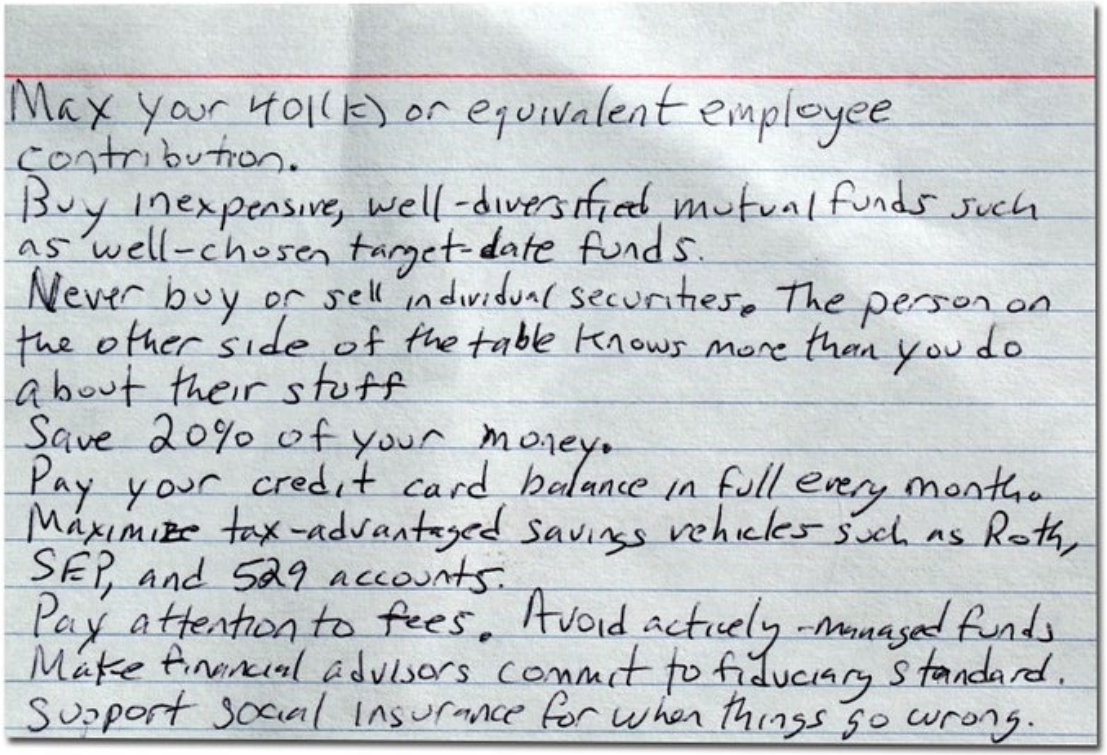
Responsibly!

Nicki Tiffin

How can we do data sharing?

Responsibly!

- An index card for data governance
- Riffing off the work of Harold Pollack



Nicki Tiffin

KEY POINTS FOR GOOD DATA GOVERNANCE

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

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

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 (e.g. in SA: POPI, Health Act)**

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 (e.g. in SA: POPI, Health Act)
4. All datasets and drives are password protected and encrypted

Easy software solutions

Examples:

7zip

- Password protection and AES encryption for all data files

Excel (RedCAP or MySQL is better!)

- Spreadsheets with password protection, encryption

Bitlocker

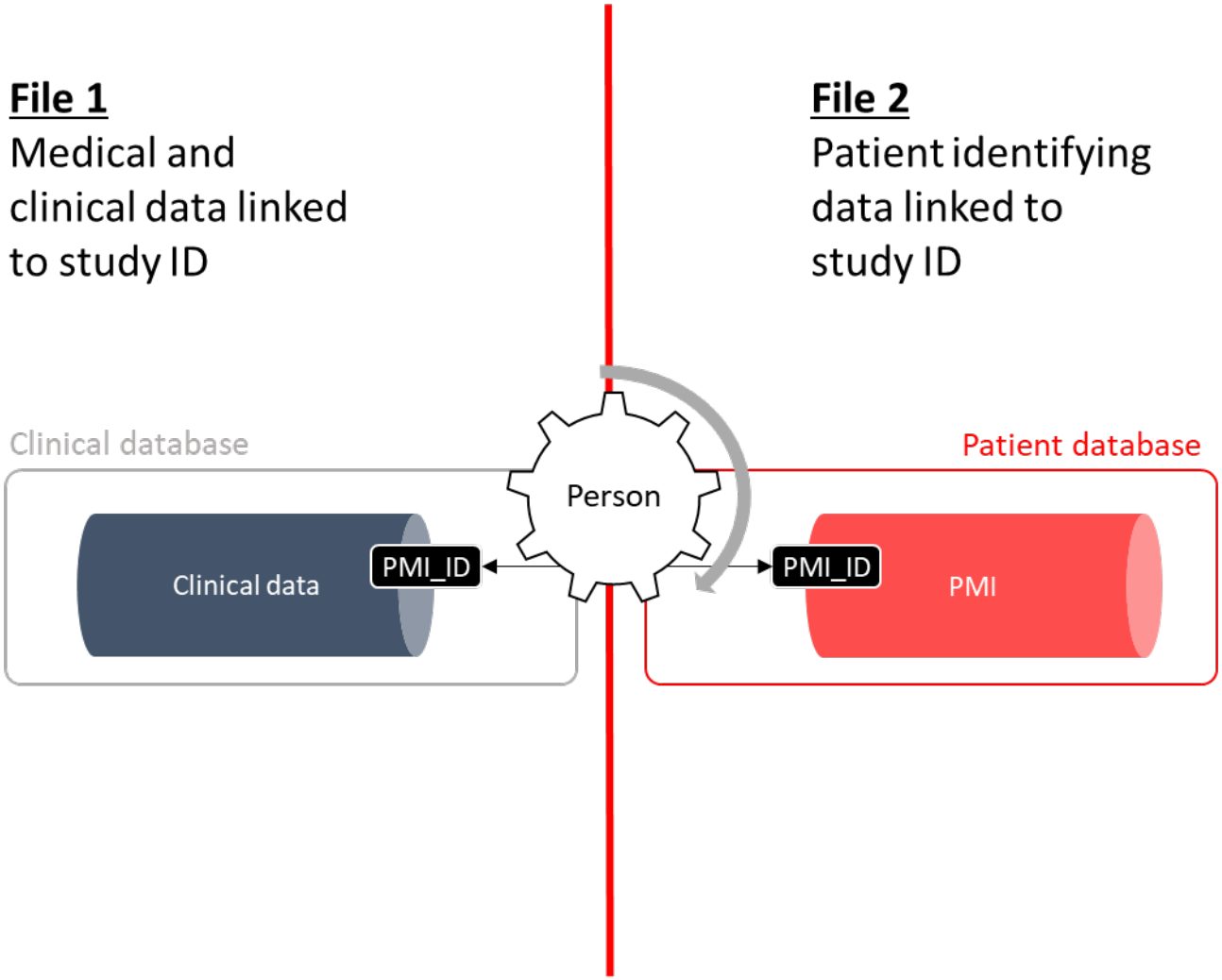
- Encrypt all drives where data are stored, including flash drives for data transfer
- Ask your departmental admin to file codes securely

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 (e.g. in SA: POPI, Health Act)
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**

Nicki Tiffin

Separation of sensitive and identifying data



Nicki Tiffin

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 (e.g. in SA: POPI, Health Act)
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**

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 (e.g. in SA: POPI, Health Act)
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).**

Nicki Tiffin

Passwords and secure transfer platforms

Never email passwords

- Use sms or telephone call

Never send passwords together with data

- Just don’t.

Use secure data transfer platforms

Examples: Filesender, Liftserver.

- Do not use external resources such as Google Drive and Dropbox for sensitive data

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 (e.g. in SA: POPI, Health Act)
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.**

Nicki Tiffin

Thank you

MLHD Organising Team, ICTS hosts
Rahul: thank you for the invitation to participate

Team members and Collaborators:

Tsaone Tamuhla
Eddie Lulamba
Taryn Allie
Vicky Nembaware

PHA4GE Ethics Working Group

ICDA Global Equity Working Group

ICDA Ethics and Policy Working Group

PHDC, Western Cape Dept of Health.

AAS Data and Biospecimen Governance Committee

Funders:

The Bill and Melinda Gates Foundation, Calestous Juma Fellowship: The African Health Information Exchange: OPP1164272

UKRI/MRC, A virtual genotyped comorbidity cohort. MC_PC_22007