Matthew McIntyre, PhD
Senior Scientist, Data Collection
23andMe

Informed consent for passive data collection

Informed consent for passive data collection

- Data may flow without the participant being aware.
- Data may be collected via a third party and/or for non-research purposes.
- Lots of potential secondary data, metadata, or “paradata”.

Passive data: Awareness

- Once consent for research and/or authorization of transfer to researcher occur, data may flow without the participant being aware.
- Analogous to running extra assays on a biobanked sample?
- What level of information should participants be given about what or how much data will be collected in a study?
- Consent forms provide few specifics about data, as such.

Passive data: Non-research uses

- Existing policy requires a distinction between data initially collected for research vs. non-research purposes.
- Non-research uses of passive data may contribute to research goals, such as participant engagement.
- Existing policy focuses on de-identification of data collected for non-research purposes (i.e. privacy and security focus)
- What if participant concerns are more focused on the level of information being provided about the uses of data?

Passive data: Metadata and paradata

- When, on what device, with what device settings, rawer form of data, etc.
- Passive data source may generate a lot of this.
- Can make true de-identification difficult.
- Can be valuable for quality control and validation, so data minimization might not be the best solution.

Consent for passive data: Policy needs

- Single clinical trials with narrow scope and short duration can probably treat passive data similarly to other kinds of data.
- Taking full advantage of the opportunity for incorporation of passive data in clinical trials and other studies may require:
  - New policies for mixed uses and sources of data.
  - Dynamic ways to inform participants about data collection.
  - Dynamic ways to seek consent for research uses of data.