Starter Question

Poll: www.slido.com #TheDataDialogue

Have you ever:

- ❑ Worked with clinical data?
- ❑ Had clinical data recorded from you?
- ❑ Benefited from clinical data?

Poll: www.slido.com #TheDataDialogue

The processes and benefits of sharing clinical data

P. H. Charlton

Guy's and St Thomas' NHS Foundation Trust

King's College London

Views my own

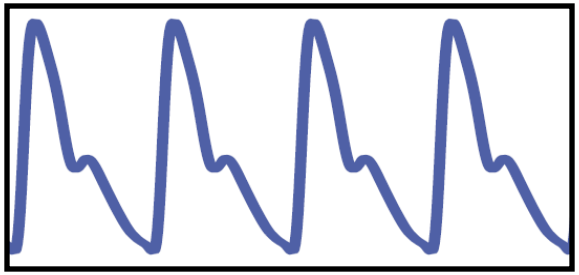
http://peterhcharlton.github.io/RRest/

Poll: www.slido.com #TheDataDialogue

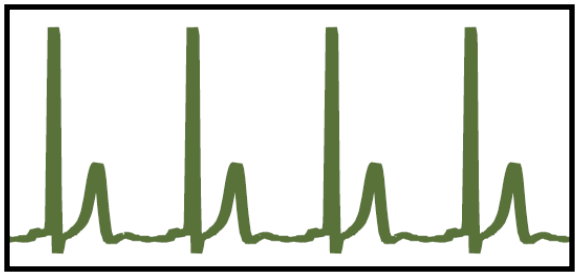
Respiratory Rate Estimation Project

An alternative approach

finger probe

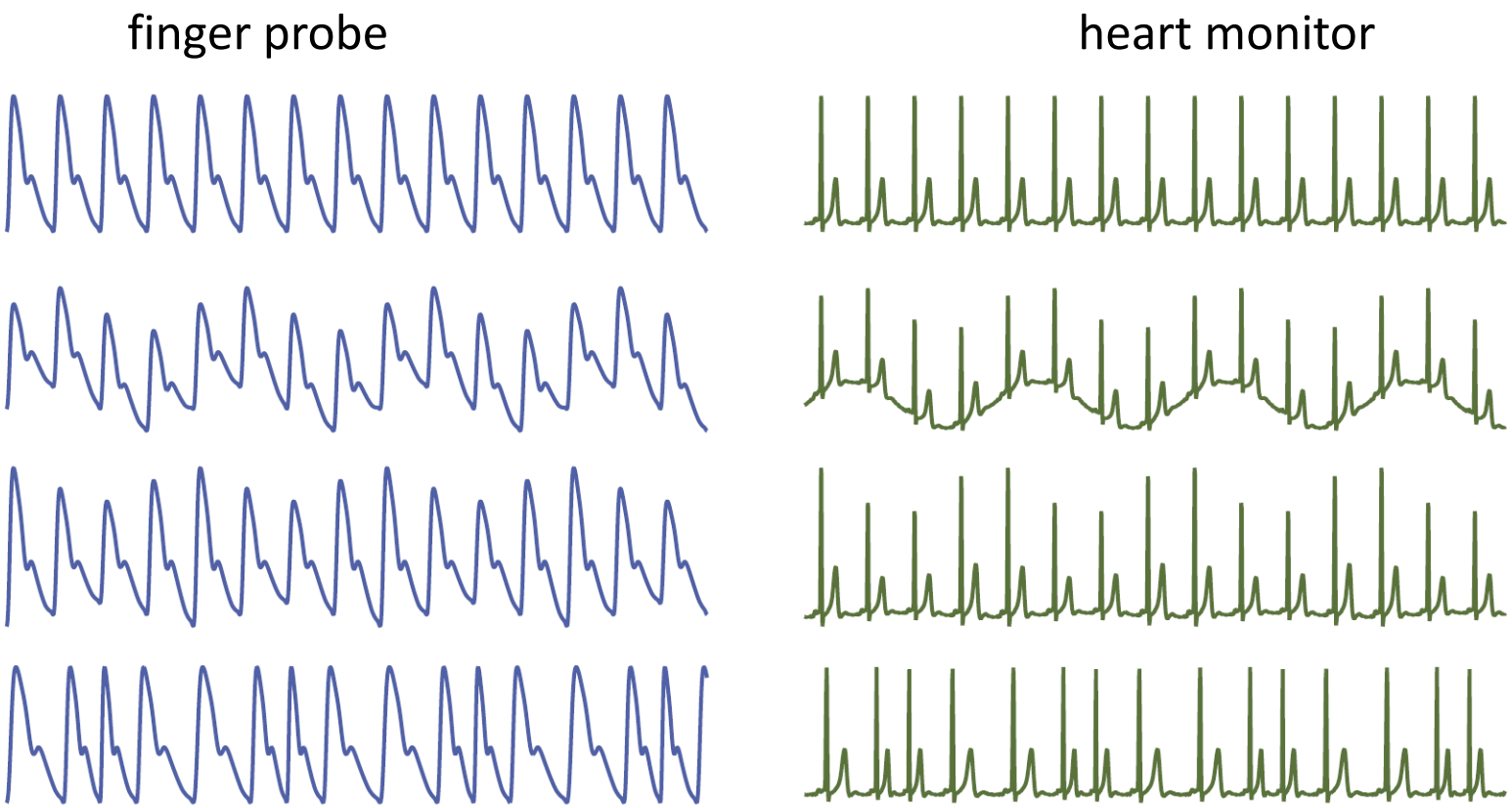


heart monitor



Poll: www.slido.com #TheDataDialogue

Estimating respiratory rate

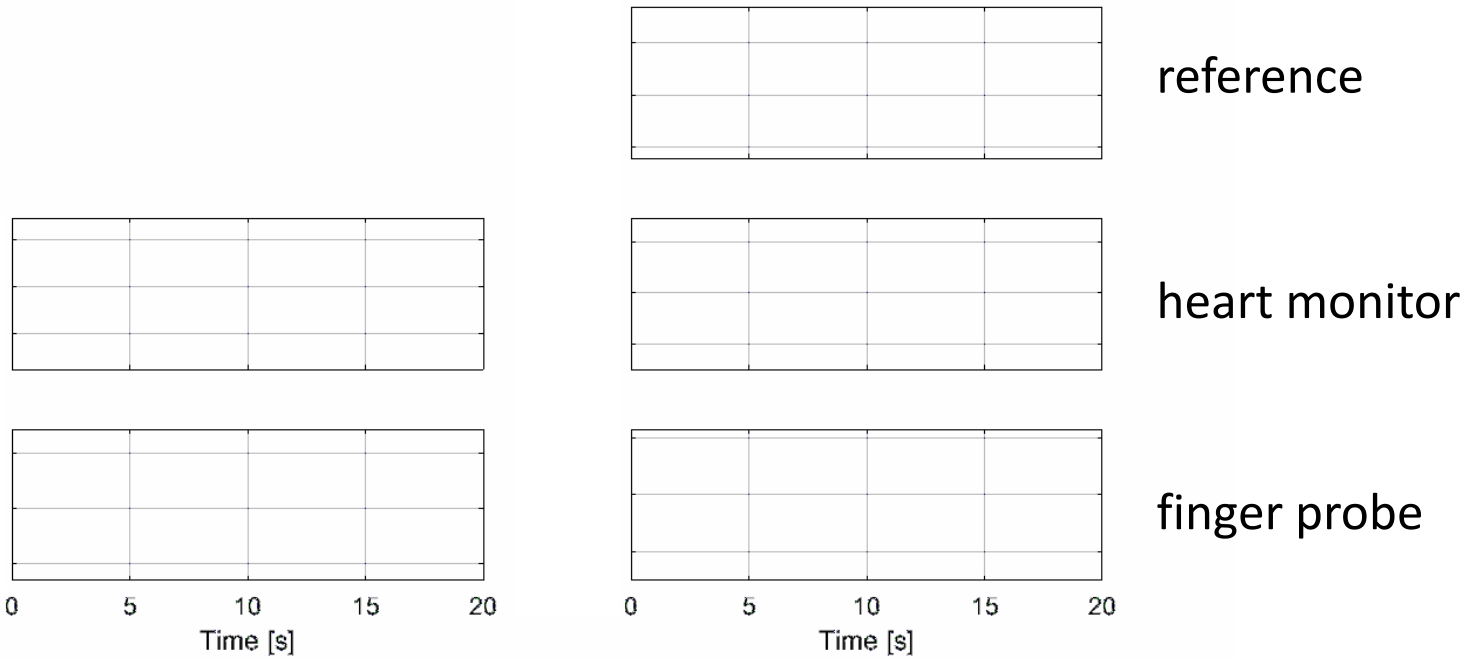


Adapted from:

(1) Addison, P.S. *et al.*: Developing an algorithm for pulse oximetry derived respiratory rate (RR(oxi)): a healthy volunteer study. *Journal of Clinical Monitoring and Computing*, 26(1), 45-51 (2012), DOI: 10.1007/s10877-011-9332-y ; (2) Pimentel, M.A. *et al.*: Probabilistic estimation of respiratory rate from wearable sensors. *in Wearable Electronics Sensors*, Springer International Publishing, 15, 241-262 (2015), DOI: 10.1007/978-3-319-18191-2_10 ; and (3) Charlton P.H. and Bonnici T. *et al.* An assessment of algorithms to estimate respiratory rate from the electrocardiogram and photoplethysmogram, *Physiological Measurement*, 37(4), 2016. DOI: 10.1088/0967-3334/37/4/610

Aim

To assess the performance of algorithms to estimate respiratory rate from routinely monitored signals



Poll: www.slido.com #TheDataDialogue

Secondary aims

Share …

benchmark dataset

standardised implementations of algorithms

… for future research

Poll: www.slido.com #TheDataDialogue

Starter Question

Poll: www.slido.com #TheDataDialogue

Have you ever:

- ❑ Worked with clinical data?
- ❑ Had clinical data recorded from you?
- ❑ Benefited from clinical data?

The processes and benefits of sharing clinical data

"individually identifiable health information" [1]

patient care or a clinical trial [2]

Common law duty of confidentiality

Data Protection Act 1988 [3]

The processes and benefits of sharing clinical data

Processes

Setting up a clinical trial

- Is my study clinical research? ✚

If so, it must:

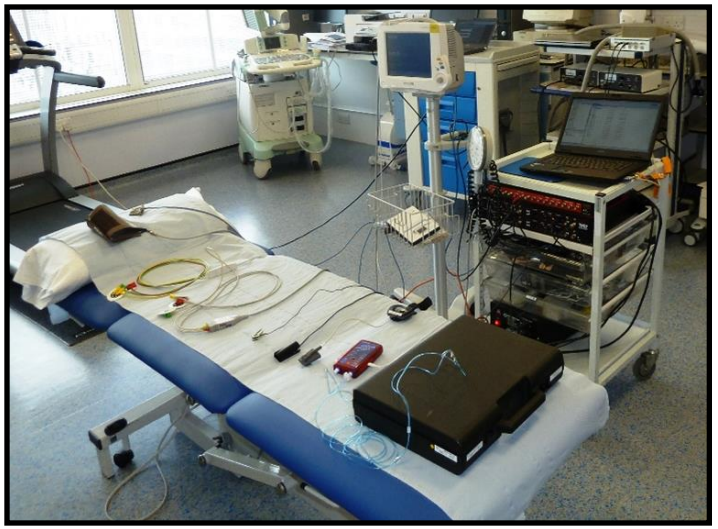
- Comply with the Declaration of Helsinki [1] …
- … by following Good Clinical Practice [2]

Setting up a clinical trial

- Is my study clinical research? ✚

If so, it must:

- Comply with the Declaration of Helsinki [1] …
- … by following Good Clinical Practice [2]



This trial:

- changed patient care and generated generalisable findings
- was reviewed by ethics committee
- required informed consent and publication of trial design
- did not disclose participants' identities

Preparation to share data

1. Plan in funding applications [3]

2. Plan to ask subjects for consent ✚
   - Include details in the information sheet
   - Statement on consent form

(Consent may not always be required: ✚)

De-identification

Clinical trial data must usually be de-identified before sharing: [7]

- It must not identify an individual
- There must be no reasonable basis to believe it can be used to identify an individual (multiple datasets?)

"The data holder is ultimately responsible for ethical and legal obligations" [4]

De-identification

Clinical trial data must usually be de-identified before sharing: [7]

- It must not identify an individual
- There must be no reasonable basis to believe it can be used to identify an individual (multiple datasets?)

	Expert Determination	Safe Harbor
Methodology	Apply statistical or scientific principles	Removal of 18 types of identifiers
Result	Very small risk that anticipated recipient could identify individual	No actual knowledge residual information can identify individual

Adapted from [7]

De-identification

Clinical trial data must usually be de-identified before sharing: [7]

- It must not identify an individual
- There must be no reasonable basis to believe it can be used to identify an individual (multiple datasets?)

e.g. names, dates (inc. age)

	Expert Determination	Safe Harbor
Methodology	Apply statistical or scientific principles	Removal of 18 types of identifiers
Result	Very small risk that anticipated recipient could identify individual	No actual knowledge residual information can identify individual

Adapted from: [7]

De-identification

Clinical trial data must usually be de-identified before sharing: [7]

- It must not identify an individual
- There must be no reasonable basis to believe it can be used to identify an individual (multiple datasets?)

"it is not possible to ensure that the probability of re-identification is zero" [8]

De-identification

Data collection:

- Subject key

Subject ID	Participant
RRest001	Mark Antony
RRest002	Marcus Brutus
RRest003	Julius Caesar
RRest004	Octavius Caesar

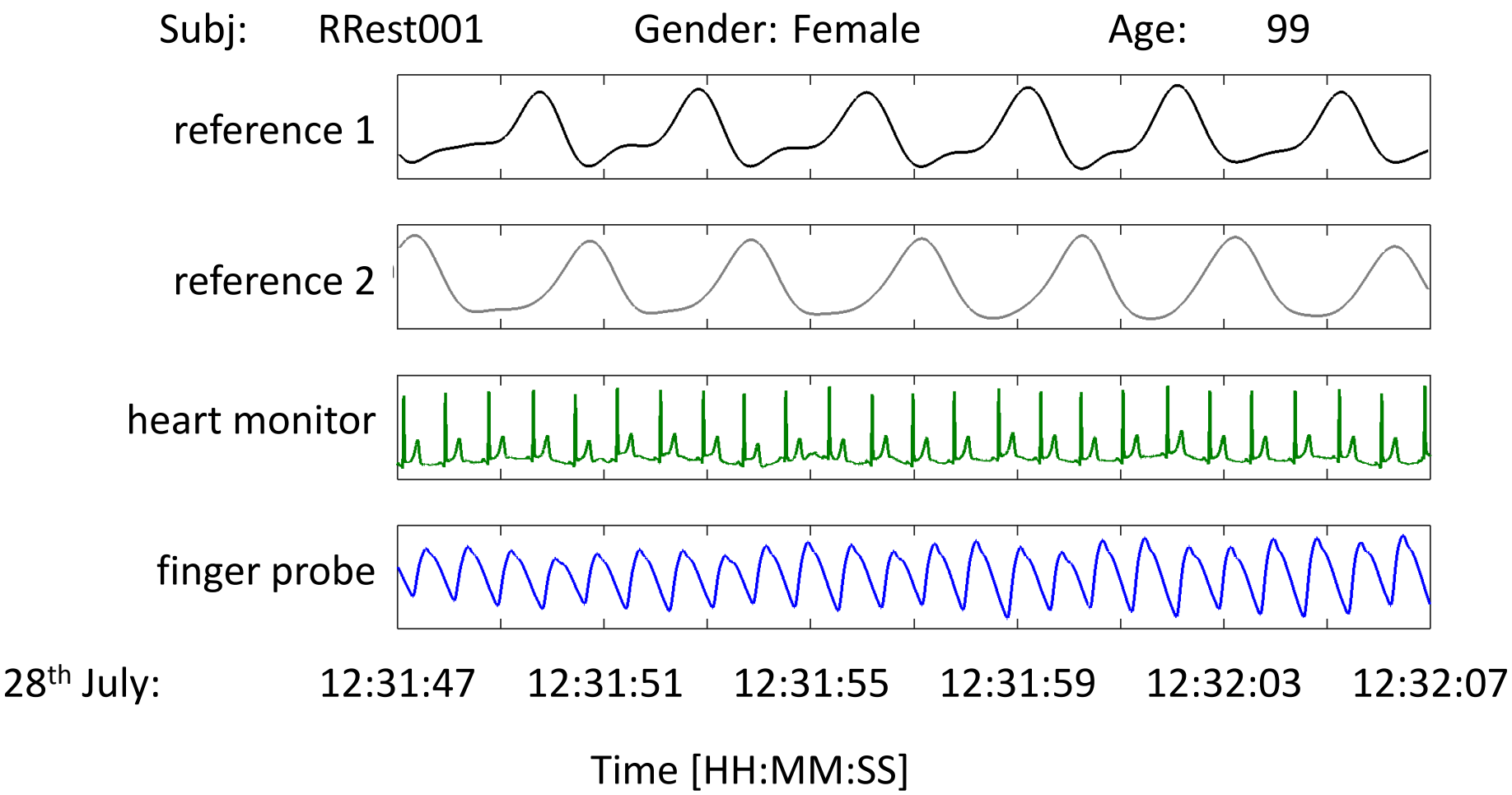
- Filenames

RRest001_finger_probe.csv

RRest001_heart_monitor.csv

RRest001_demographics.csv

De-identification

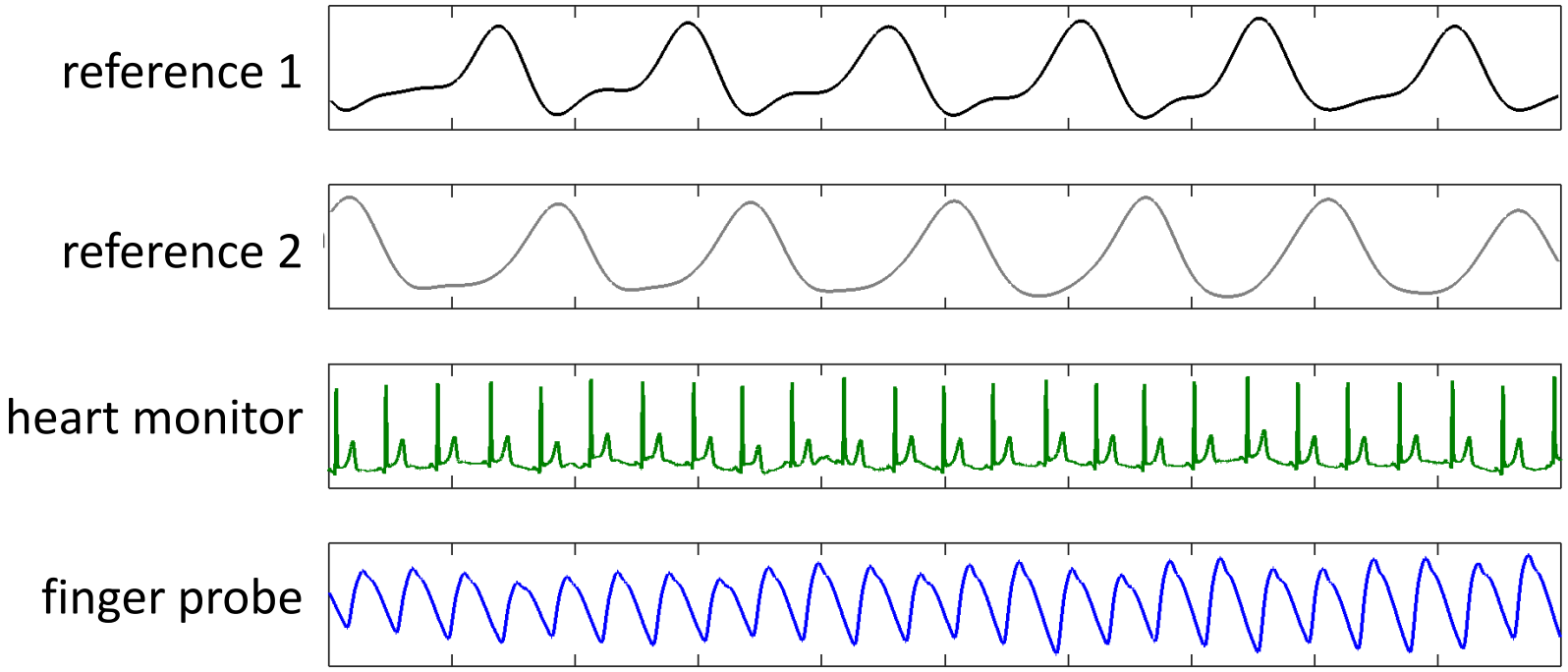


De-identification

Pseudonymous

Age > 90

Subj: **RRest001** Gender: Female Age: **99**

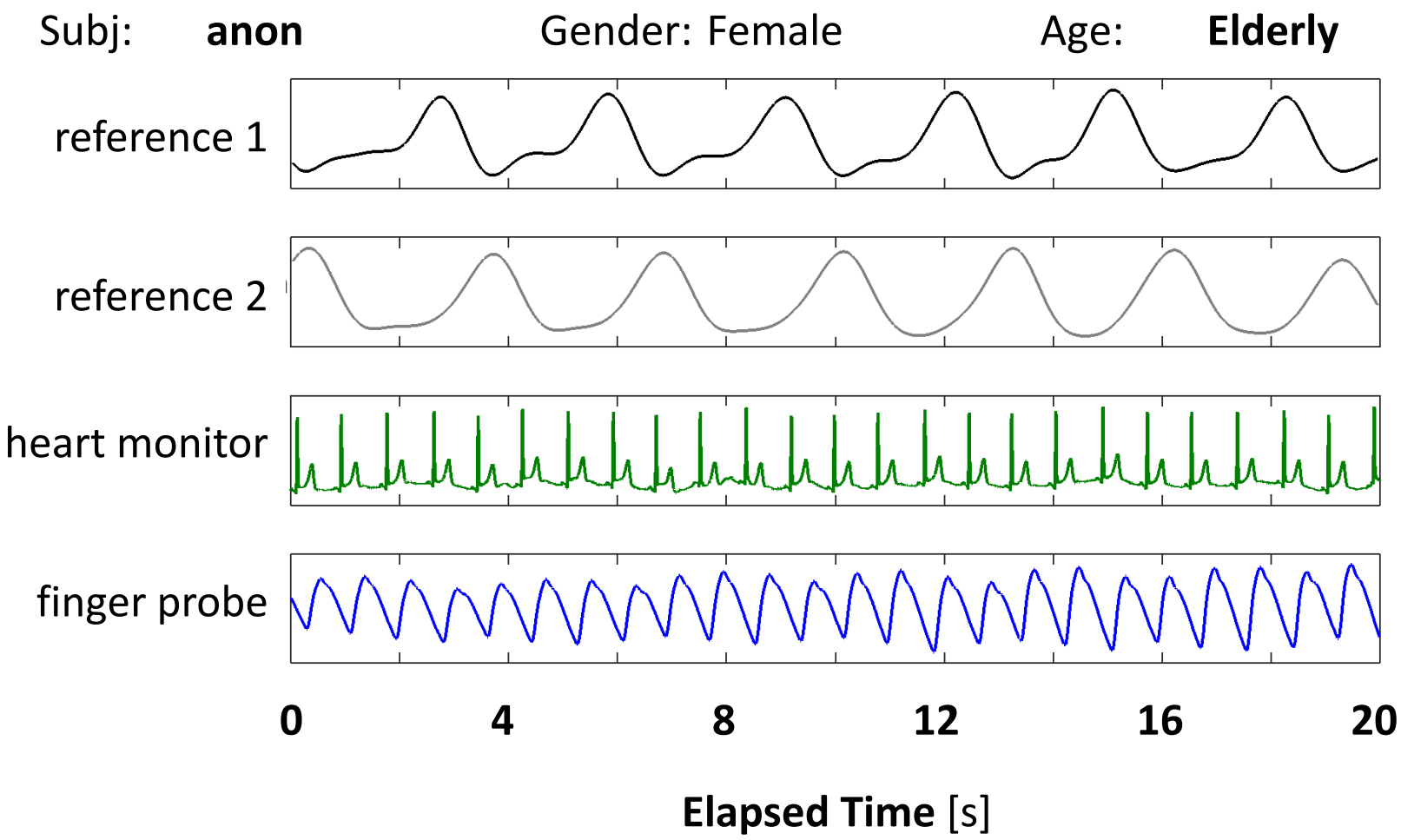


28th July: **12:31:47** **12:31:51** **12:31:55** **12:31:59** **12:32:03** **12:32:07**

Dates

Time [HH:MM:SS]

De-identification



Data preparation

Data prepared for analysis:

- to reduce workload and domain-specific knowledge requirements
- whilst retaining all potentially useful information (usually not raw data [9])

Data preparation

Data prepared for analysis:

- to reduce workload and domain-specific knowledge requirements
- whilst retaining all potentially useful information (usually not raw data [9])

This trial:

- re-format

```
Milliseconds since 01.01.1970;SpO2-O2(%);Perf-
REL(-);Pulse-Pulse(bpm);NBP-MEAN(mmHg);RR-
RR(rpm);NBP-SYS(mmHg);NBP-DIA(mmHg);PVC-CNT(bpm)
4102444800000;;;56;;18;;;0
4102444801025;98.6;2.1;57;;18;;;0
4102444802050;98.4;2.0;58;;18;;;0
4102444803075;98.2;2.0;57;;18;;;0
4102444804100;98.3;2.0;56;;18;;;0
4102444805125;98.2;1.9;56;;19;;;0
```

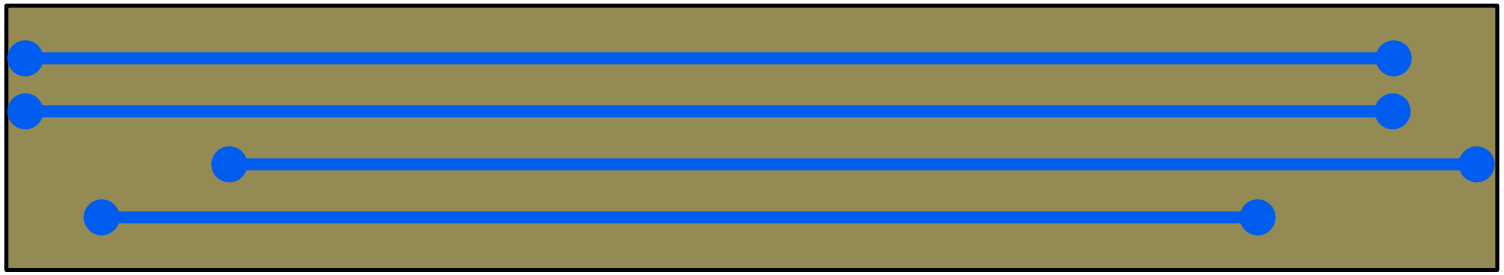
Data preparation

Data prepared for analysis:

- to reduce workload and domain-specific knowledge requirements
- whilst retaining all potentially useful information

This trial:

- re-format
- time-alignment



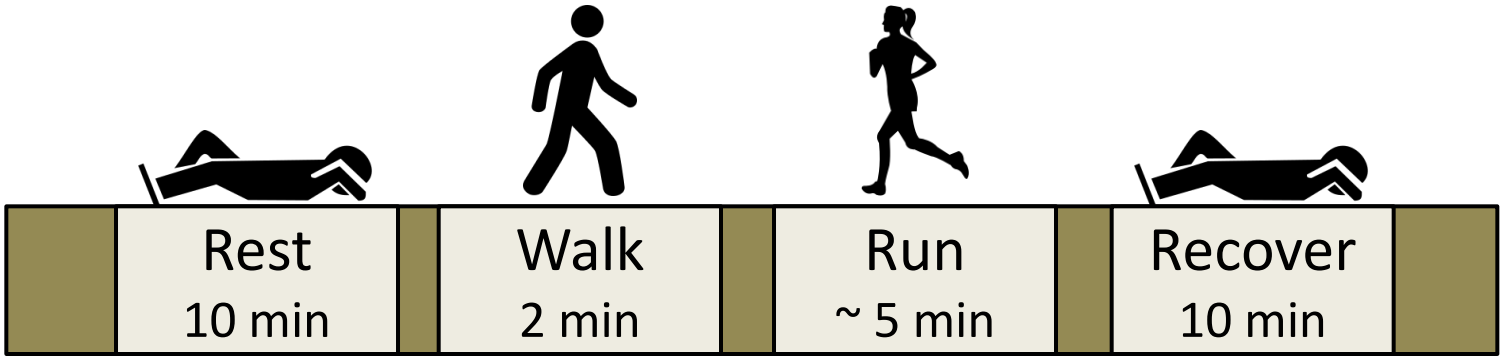
Data preparation

Data prepared for analysis:

- to reduce workload and domain-specific knowledge requirements
- whilst retaining all potentially useful information

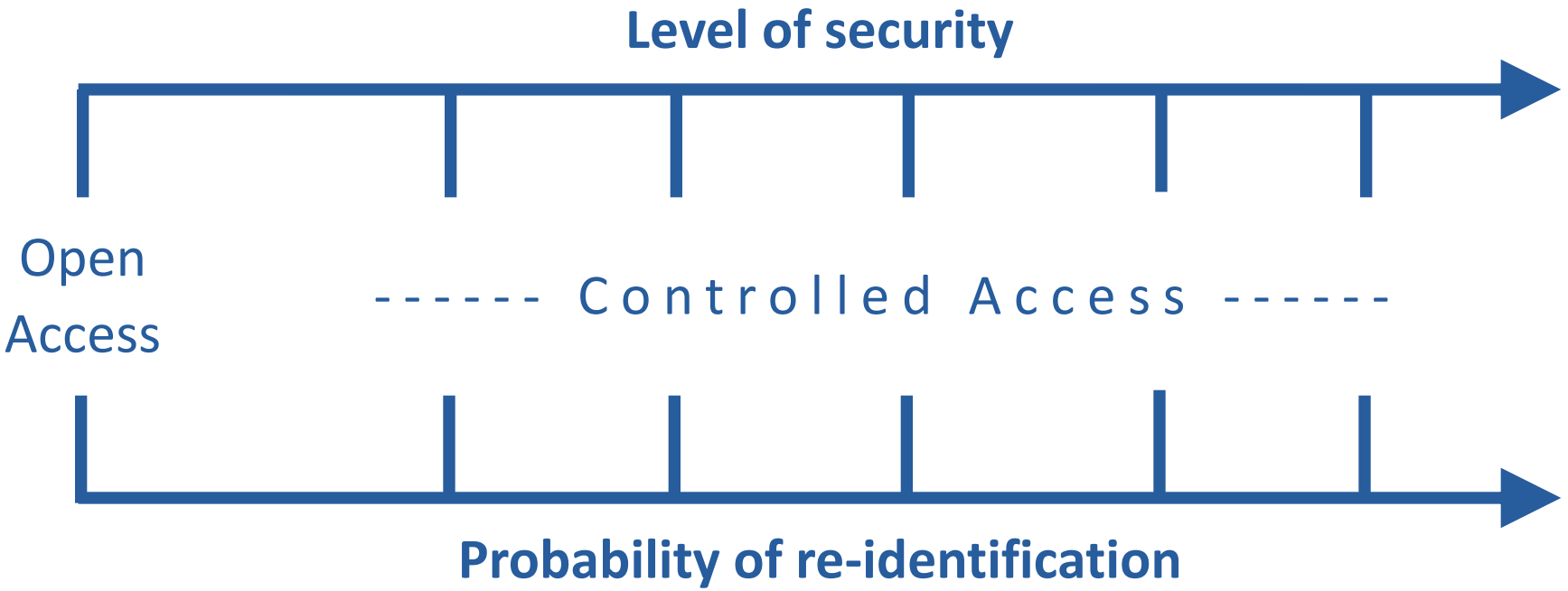
This trial:

- re-format
- time-alignment
- extraction of relevant periods



Sharing data

Method: [5],[8]



Publicly available

Terms of Use

Data Analysis Plan

Full Contract

Sharing data

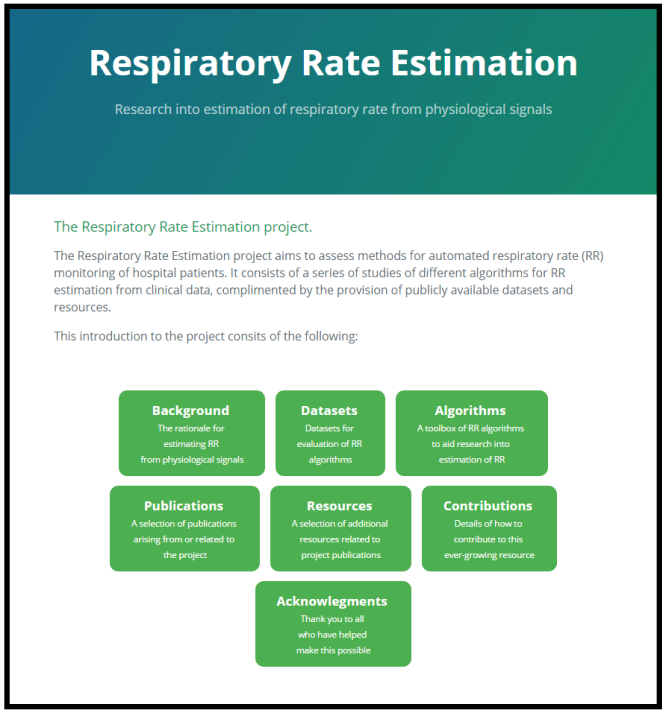
The following should be shared: [6]

- Analytic Dataset
- Metadata
- Protocol
- Study Analysis Plan
- Analysis code

Sharing data

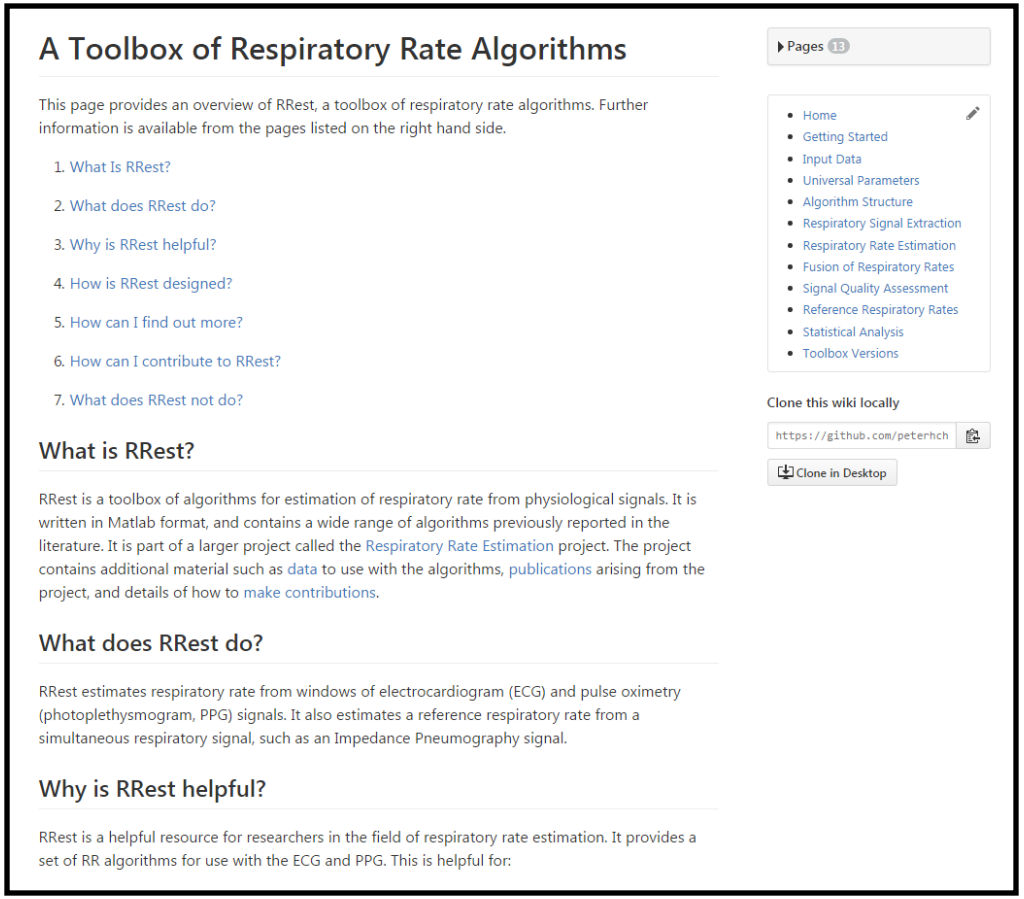
This trial:

- Data: Open access, accessible format
- Algorithms: GitHub respository



Additional Materials

- User Manual
  - updated as Qs arise



A Toolbox of Respiratory Rate Algorithms

This page provides an overview of RRest, a toolbox of respiratory rate algorithms. Further information is available from the pages listed on the right hand side.

1. What Is RRest?
2. What does RRest do?
3. Why is RRest helpful?
4. How is RRest designed?
5. How can I find out more?
6. How can I contribute to RRest?
7. What does RRest not do?

What is RRest?

RRest is a toolbox of algorithms for estimation of respiratory rate from physiological signals. It is written in Matlab format, and contains a wide range of algorithms previously reported in the literature. It is part of a larger project called the Respiratory Rate Estimation project. The project contains additional material such as data to use with the algorithms, publications arising from the project, and details of how to make contributions.

What does RRest do?

RRest estimates respiratory rate from windows of electrocardiogram (ECG) and pulse oximetry (photoplethysmogram, PPG) signals. It also estimates a reference respiratory rate from a simultaneous respiratory signal, such as an Impedance Pneumography signal.

Why is RRest helpful?

RRest is a helpful resource for researchers in the field of respiratory rate estimation. It provides a set of RR algorithms for use with the ECG and PPG. This is helpful for:

Pages 13

- Home
- Getting Started
- Input Data
- Universal Parameters
- Algorithm Structure
- Respiratory Signal Extraction
- Respiratory Rate Estimation
- Fusion of Respiratory Rates
- Signal Quality Assessment
- Reference Respiratory Rates
- Statistical Analysis
- Toolbox Versions

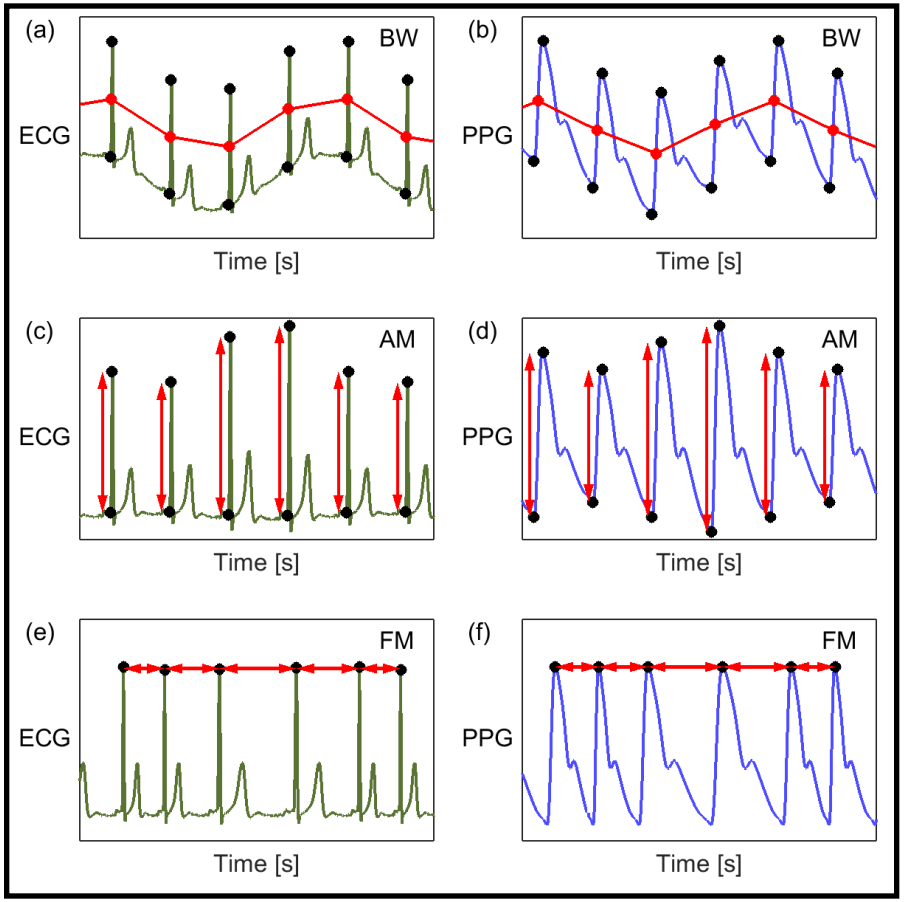
Clone this wiki locally

https://github.com/peterhch

Clone in Desktop

Additional Materials

- User Manual
  - updated as Qs arise
- Tutorial



Adapted from: [16] which is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (http://creativecommons.org/licenses/by-nc/4.0/)

Additional Materials

- User Manual
  - updated as Qs arise
- Tutorial
- Instructions for replicating analyses

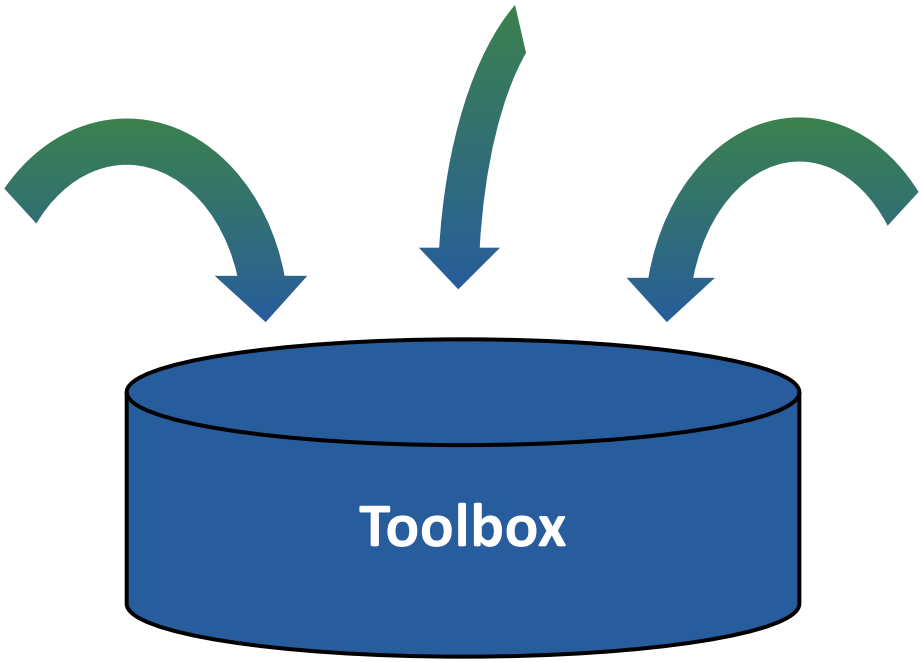
Replicating this Publication

The work presented in this case study can be replicated as follows:

- Download data from the MIMIC II dataset using the script provided here.
- Use Version 1 of the toolbox of algorithms. To perform the analysis call the main script using the following command: *RRest('mimicii')*

Compatability

- Other datasets take a variety of formats
- They can be imported using the scripts provided



Benefits

This project

- Transparency
- Reproducibility
- Internal peer review
- Ongoing peer review
- Required by some journals [14] and funding providers

Future benefits

- Build on our work
- More accessible to non-specialists
- Multiple dataset studies
- Promoting collaboration
- Increase speed of research
- New research questions
- Decreased burden on research subjects
- Education of students

Conclusions

- Considered the processes for collection and sharing of clinical trial data, using the Respiratory Rate Estimation project as a case study.
- Looked briefly at the benefits of sharing clinical data
- Links are provided to references and additional resources

This presentation is part of the **Respiratory Rate Estimation Project** at:

http://peterhcharlton.github.io/RRest/

Acknowledgments

Clinical	Engineering	Funders
Dr Tim Bonnici	Prof David Clifton	EPSRC
Prof Richard Beale	Prof Lionel Tarassenko	NIHR
Dr Peter Watkinson	Dr Mauricio Villarroel	Wellcome Trust
John Brooks	Dr Marco Pimentel	Royal Academey of Engineering
Isabelle Schelcher	Dr Christina Orphanidou	
Ricky Yang		
Katie Lei		
John Smith		

The views expressed are those of the author and not necessarily those of Guy's and St Thomas' NHS Foundation Trust, King's College London, the EPSRC, NHS, NIHR, Department of Health, Wellcome Trust, or Royal Academy of Engineering.

Additional Acknowledgments

Thanks also to:

- Jason Long for Cayman Theme which inspired this presentation template
- Open Clipart for some of the images in this presentation

References

Additional details of the Vortal Study, which formed part of the case study in this presentation, are available in the following publication:

Charlton P.H. and Bonnici T. *et al.* **An assessment of algorithms to estimate respiratory rate from the electrocardiogram and photoplethysmogram**, *Physiological Measurement*, 37(4), 2016. DOI: 10.1088/0967-3334/37/4/610

The toolbox of algorithms is described in the following publication:

[16] Charlton P.H. *et al.* **Waveform Analysis to Estimate Respiratory Rate**, in Secondary Analysis of Electronic Health Records, Springer International Publishing, pp. 377–390. DOI: 10.1007/978-3-319-43742-2_26

References

[1] World Medical Association, 2013. World Medical Association Declaration of Helsinki: Ethical Principles for Medical Research Involving Human Subjects. *JAMA*, 310(20), pp.E1 – E4. DOI: 10.1001/jama.2013.281053

Foundational principles for clinical research

[2] ICH Expert Working Group, 1996. Guideline for Good Clinical Practice. Link

A standard for ensuring the ethical quality of trials involving human subjects. Consistent with the principles in the Declaration of Helsinki.

[3] Veerle Van den Eynden *et al.*, 2011. Managing and Sharing Data - Best Practice For Researchers. *UK Data Archive*, pp.1–40. Link

Helpful 'how to' for managing and sharing data, including: writing a data management and sharing plan (p.6)

References

[4] Tucker, K. et al., 2016. Protecting patient privacy when sharing patient-level data from clinical trials. *BMC Medical Research Methodology*, 16(S1), p.77. DOI: 10.1186/s12874-016-0169-4

Recommends data anonymisation and controlled access to shared data, with suggestions of practical methods for each process.

[5] Smith, C. et al., 2015. Good practice principles for sharing individual participant data from publicly funded clinical trials. *Trials*, 16(Suppl 2). DOI: 10.1186/1745-6215-16-S2-O1

Helpful advice on sharing data from clinical trials, including: data sharing models (p.9); handling data requests (p.13); text for use in a consent form (p.16); an example Data Use Agreement (p.22).

[6] Institute of Medicine, 2015. *Sharing Clinical Trial Data: Maximizing Benefits, Minimizing Risk*, Washington, D.C.: National Academies Press. DOI: 10.17226/18998

Essential reading – perhaps the most useful resource to date on sharing clinical trial data.

References

[7] HHS (U.S. Department of Health and Human Services), 2012. Guidance regarding methods for de-identification of protected health information in accordance with the Health Insurance Portability and Accountability Act (HIPAA) Privacy Rule. Link – accessed on 26th July 2016.

Detailed guidance on exactly how to de-identify data. A helpful introduction is provided in [4].

[8] Emam, K. El *et al.*, 2015. Anonymising and sharing individual patient data. *The BMJ*, 350, p.h1139. DOI: 10.1136/bmj.h1139

Brief overview of "key concepts and principles for anonymising health data while ensuring it remains suitable for meaningful analysis".

[9] Lo, B., 2015. Sharing Clinical Trial Data. *JAMA*, 313(8), p.793. DOI: 10.1001/jama.2015.292.

A commentary by the Chair of the committee who contributed to the report in [6]. Considers the question of whether to share raw data.

References

[10] Rosenblatt, M., Jain, S.H. & Cahill, M., 2015. Sharing of Clinical Trial Data: Benefits, Risks, and Uniform Principles. *Annals of Internal Medicine*, 162(4), p.306. DOI: 10.7326/M14-1299

[11] Goodman, S.N., 2015. Clinical trial data sharing: What do we do now? *Annals of Internal Medicine*, 162(4), pp.308–309. DOI: 10.7326/M15-0021

[12] Mello, M.M. et al., 2013. Preparing for Responsible Sharing of Clinical Trial Data. *New England Journal of Medicine*, 369(17), pp.1651–1658. DOI: 10.1056/NEJMhle1309073

[13] Drazen, J.M., 2015. Sharing Individual Patient Data from Clinical Trials. *New England Journal of Medicine*, 372(3), pp.201–202. DOI: 10.1056/NEJMp1415160

[14] Taichman, D.B. et al., 2016. Sharing Clinical Trial Data — A Proposal from the International Committee of Medical Journal Editors. *New England Journal of Medicine*, 374(4), pp.384–386. DOI: 10.1056/NEJMe1515172

[15] Sandercock, P.A. et al., 2011. The International Stroke Trial database. *Trials*, 12, p.101. DOI: 10.1186/1745-6215-12-101

Additional Resources

Is my study clinical research?

A study is clinical research if it involves:

- Randomised treatments, or
- Changing treatment / patient care from accepted standards, or
- Generalisable findings

For further information, and to check whether your study is classed as clinical research, see the tool at:

http://www.hra-decisiontools.org.uk/research/

Obtaining consent to share data

Include information in the patient information sheet, and a statement in the informed consent form, such as the following text (as suggested by the Health Research Authority):

"I understand that the information collected about me will be used to support other research in the future, and may be shared anonymously with other researchers."

These details are taken from [5], p.16

Is consent required?

"The best way … is to obtain consent" [5]

"Many jurisdictions … do not designate anonymised health data as personal information. Therefore, such data would no longer be covered by privacy laws." [8]

"A lack of consent for sharing should not prohibit sharing … anonymised data" [5]

"Sharing of data without specific participant consent may be ethically acceptable and legally permitted in certain instances." [6]

Usually anonymisation is a must. Even then, check with local authorities (*e.g.* Healthcare provider, Ethics Committee)

Is consent required? A case study

The International Stroke Trial Database reported in did not require consent specifically for data sharing:

When creating this database, “consent for publication of raw data was not obtained from participants”. Whilst consent for participation in the trial was obtained, the patients “were treated 15-20 years ago, and many have died. The dataset is fully anonymous … In our view, publication of the dataset clearly presents no material risk to confidentiality of study participants.” [15]

How to share data

The UK Data Archive have suggested several ways of sharing data [3]:

- Data repository
- Supplementary material in a journal publication
- Intitutional repository
- Project or institutional website
- Informal sharing with peers

For further details see p.4 of [3]

Benefits of sharing data

Several benefits of sharing clinical data are mentioned in the literature, such as:

- Providing new insights [10]
- Clarifying the effectiveness and safety of medicines [10]
- Preventing repetition of data collection and research [10]
- Improving transparency [10]
- Improving public trust in research [10]
- Improved meta-analyses [11]
- Real-world examples for teaching [11]
- Speed up innovation [12]
- Reward risk of trial participants [13]

Several concerns are also mentioned in the literature, such as:

- Potential for misleading analyses of trial data [10]
- De-identification of trial participants [10]