Reducing IrreProducibility in labOratory STudiEs
The RIPOSTE framework

Improving the design and analysis of laboratory based research

Nick Masca, Elizabeth Hensor, Victoria Cornelius, Dawn Teare

&

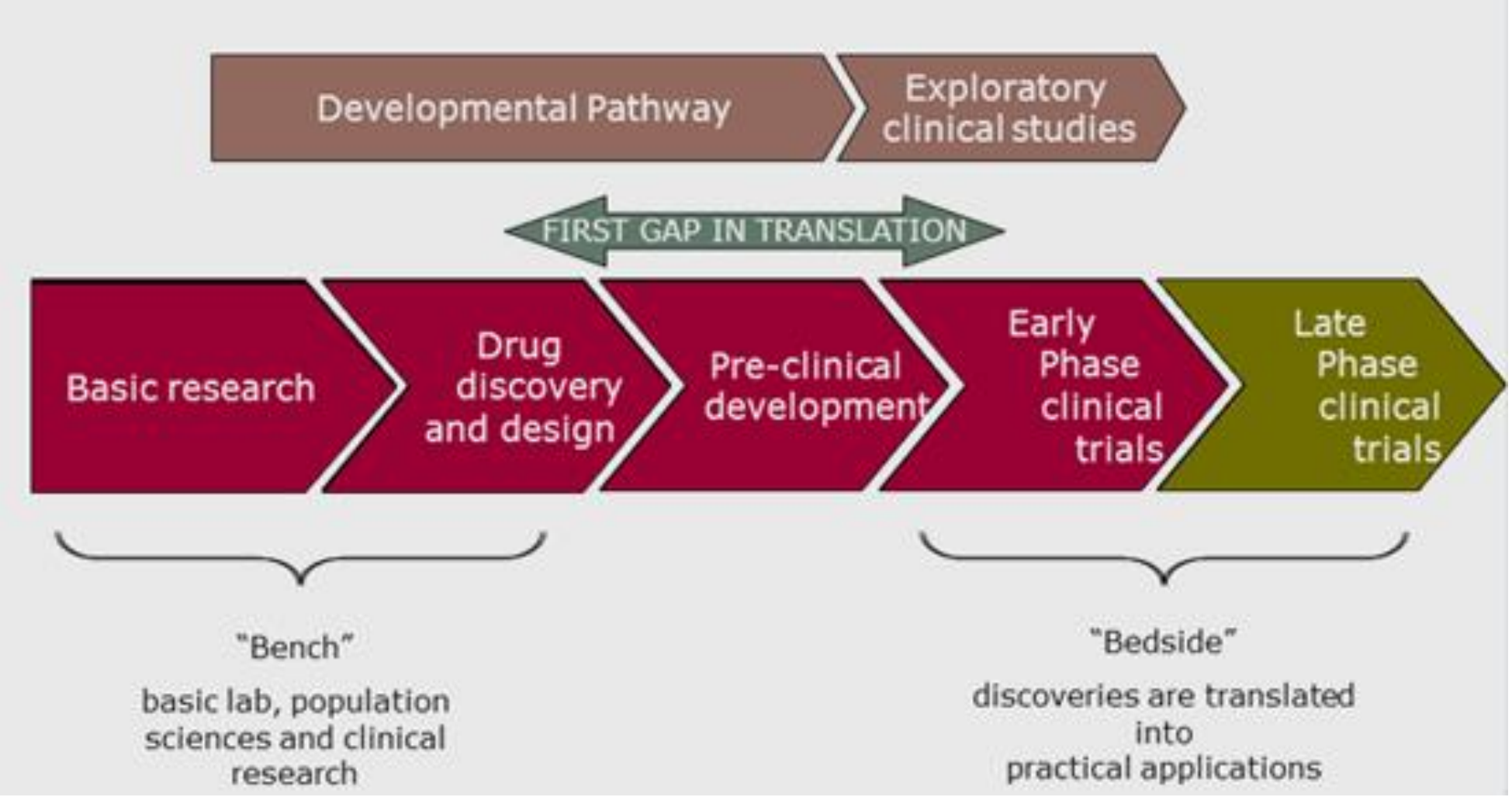
The RIPOSTE consortium

Francesca Buffa, Helen Marriott, James Eales, Michael Messenger, Amy Anderson, Chris Boot, Catey Bunce, Robert Goldin, Jessica Harris, Rod Hinchliffe, Hiba Junaid, Shaun Kingston, Carmen Martin-Ruiz, Christopher P Nelson, Janet Peacock, Paul Seed, Beth Shrinkins, Karl J. Staples, Jamie Toombs, Adam K. A. Wright

What do we mean by “laboratory study”?

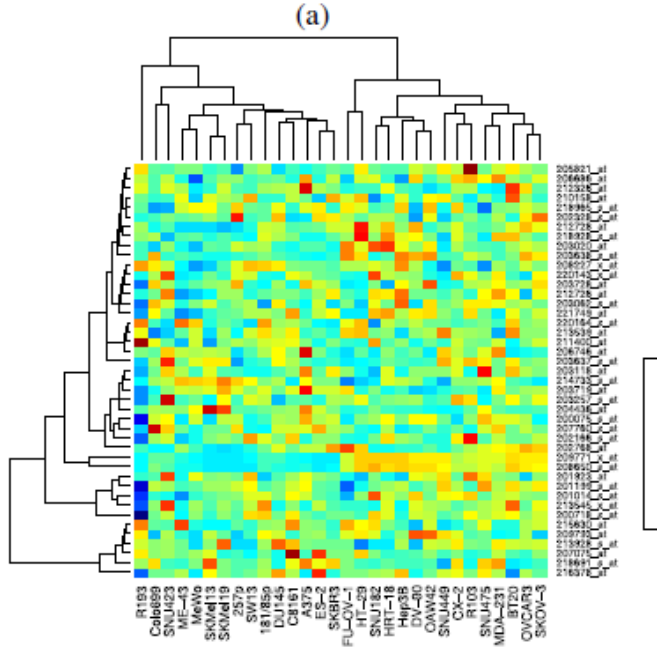
Laboratory studies encompass a diverse range of techniques and span a broad range of biomedical fields: investigations of biological pathways underpinning drug response; microbial pathogenesis; discovery of biomarkers; safety and efficacy of drug interventions

- **A study in which any aspect of the procedure or analysis is carried out in a research facility/laboratory**
- **May be in vivo (e.g. imaging) or in vitro (e.g. cell culture)**
- **Includes both experimental and observational studies**
- **May involve hypothesis generation or hypothesis testing/ confirmation**
- **Can be small (e.g. within a single lab) or large scale (e.g. multi-centre genome-wide association studies)**



- Play a major role in preclinical research
- Pivotal to informing the translational pathway
- Promising results lead to decisions to undertake clinical trials

Example: data processing



Heat map: expression submatrix for the signature genes across 30 cell lines

- Biological microarrays can be used to answer questions about how diseases operate, and promise us personalize therapy
- Not all cancer patients respond to first line drug therapy
- An array-based test indicating whether a patient is expected to respond would be highly valuable

Example: data processing

The Annals of Applied Statistics
2009, Vol. 3, No. 4, 1309–1334
DOI: 10.1214/09-AOAS291

DERIVING CHEMOSENSITIVITY FROM CELL LINES: FORENSIC BIOINFORMATICS AND REPRODUCIBLE RESEARCH IN HIGH-THROUGHPUT BIOLOGY

BY KEITH A. BAGGERLY[1] AND KEVIN R. COOMBES[2]

- Related publications using microarray-based signatures of drug sensitivity derived from cell lines to predict patient response
- On the basis of the results found in the included studies, patients were being allocated to treatment arms in clinical trials
- Set out to reproduce the results
- **raw data**, **written data processing & analysis methods** and publication **results**

Example: data processing

- Found five publications that the incorporate several simple errors that had the potential to put patients at risk (1 since retracted)
- Data processing errors included:
  - mixing up gene labels, mixing up group labels (resistant and sensitive)
  - Inclusion of duplicate samples from patients
  - Exclusion on a large number of processed samples
- ‘Poor documentation can shift from an inconvenience to an active danger when it obscures not just methods but errors. This can lead to scenarios where well-meaning investigators argue in good faith for treating patients with apparently promising drugs that are in fact ineffective or even contraindicated’

Irreproducibility

Many studies published articles have been withdrawn

Studies showing promising results are never replicated

Wastes resource

Hinders development of new treatments

Potentially puts patients at risk

Irreproducibility

- Irreproducibility of preclinical studies is a acknowledged problem
- Journal Editors and researchers are implementing reporting guidelines - to allow researchers the ability to be able to replicate

What contributes to irreproducibility?

- Poor Reporting; scientists unable to reproduce findings
- Rewarded for novel research and significant results over attempts to replicate finding
- Technical issues
- Poor study design introducing bias
- Small sample size, testing many hypotheses without adjusting
- Inappropriate analysis methods used

Existing guidance

- Control of Substances Hazardous to Health Regulations (COSHH)
- Good Laboratory Practice (GLP)
- International Committee on Harmonisation of Good Clinical Practice (ICH GCP)
- Human Tissues Act 2004
- Some work requires accreditation from UKAS & CPA, ISO/BSI standards

However none of these regulations specifically address study design, and there is often no formal requirement to produce a study specific protocol or analysis plan

Influencing the design of Laboratory studies

First NIHR Laboratory Statistics meeting in April 2013

- Improve the communication between scientists and statisticians
- Break down technical barriers

Influencing the design of Laboratory studies

First NIHR Laboratory Statistics meeting in April 2013

- Improve the communication between scientists and statisticians
- Break down technical barriers

Draft framework developed

- RIPOSTE Consortium formed
- 12 statisticians & 12 Scientists

Influencing the design of Laboratory studies

First NIHR Laboratory Statistics meeting in April 2013

- Improve the communication between scientists and statisticians
- Break down technical barriers

Draft framework developed

- RIPOSTE Consortium formed
- 12 statisticians & 12 Scientists

Feb 2014 RIPOSTE Consortium

- Framework tested and further developed

Influencing the design of Laboratory studies

First NIHR Laboratory Statistics meeting in April 2013

- Improve the communication between scientists and statisticians
- Break down technical barriers

Draft framework developed

- RIPOSTE Consortium formed
- 12 statisticians & 12 Scientists

Feb 2014 RIPOSTE Consortium

- Framework tested and further developed

Two ‘Delphi’ surveys

- Final framework and publication

The RIPOSTE framework

- To support early discussion of study design and analysis between scientists and statisticians with an aim to improve the design, analysis and reporting of laboratory based studies
- Consist a series of **Items**, followed by **prompts/considerations** and then list of **Details** to facilitate discussion
- ‘laboratory studies’ coves a wide range of study types- some aspects of the framework will not be relevant for discussion

1) Research aims, objectives, specific outcomes and hypotheses

Define the key aims of the study
Identify the variables and quantities/ qualities of interest that will be measured (these may be different for each hypothesis)
List the research question(s) that will be addressed and/or any hypotheses that you would like to test

2) Study Planning

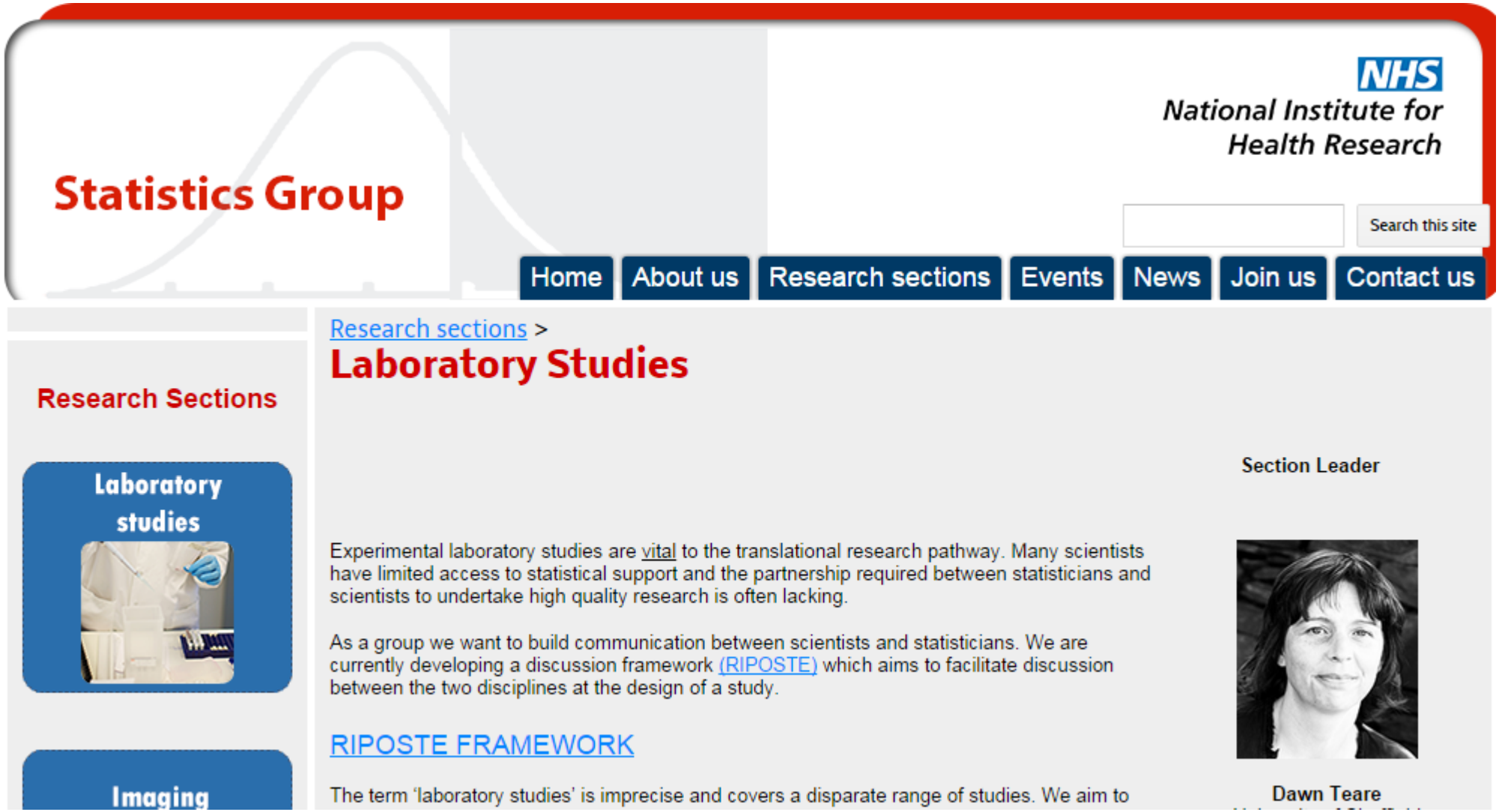
Logistical considerations	Ethical approval
	Statistical support
	Data collection & management
Materials and techniques	Laboratory equipment & methods
	Configuration and standardisation of materials and methods
	Software
	What constraints/ limits are there to the available resources?

3) Study Design

Design	Experimental/ sampling units
	Randomisation
	Blinding (Masking)
	Groups, treatments, and other predictors of interest
	Use of analytical controls
	Other potential biases, confounders and sources of variability
	Sample size considerations

4) Planned Analysis

Data assessment and preparation	Quality control criteria
	Data verification
	Data normalisation/ correction
	Outliers
Statistical methods	Describe the different analyses to be performed
	Missing data
	Multiple testing
	Interim analysis
	Replication and/or validation



http://www.statistics-group.nihr.ac.uk/

Any questions?

Workshop

Split into 4 groups: Janet Peacock, Paul Seed, Fiona Reed, Victoria Cornelius

Scientist/ researchers:

- What improvements to undertaking and publishing research do you think could be introduced to improve the problem of irreproducibility? E.g. Publications of protocols for laboratory experiments
- How could these be achieved?

Statisticians:

- When providing statistical advice for laboratory studies - what are there common issues that arise? For example, due to:
  - Your own familiarity with the techniques involved (if these are particularly specialised)?
  - Clarity of the description of hypotheses/methods

All groups:

- With the aim or promoting robust design, good reporting and multidiscipline team working: what aspect do you think the laboratory research group should tackle next?
- Would you like to be involved?