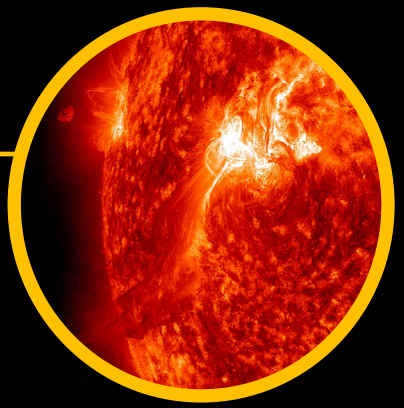
Reproducibility and Openness

Facilitating Scientific Progress and Research Impact

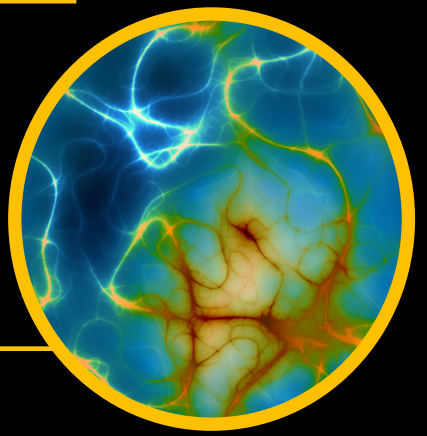
An introduction to Open Science and the changing face of academia

What items have you produced in your [subject name] course so far?

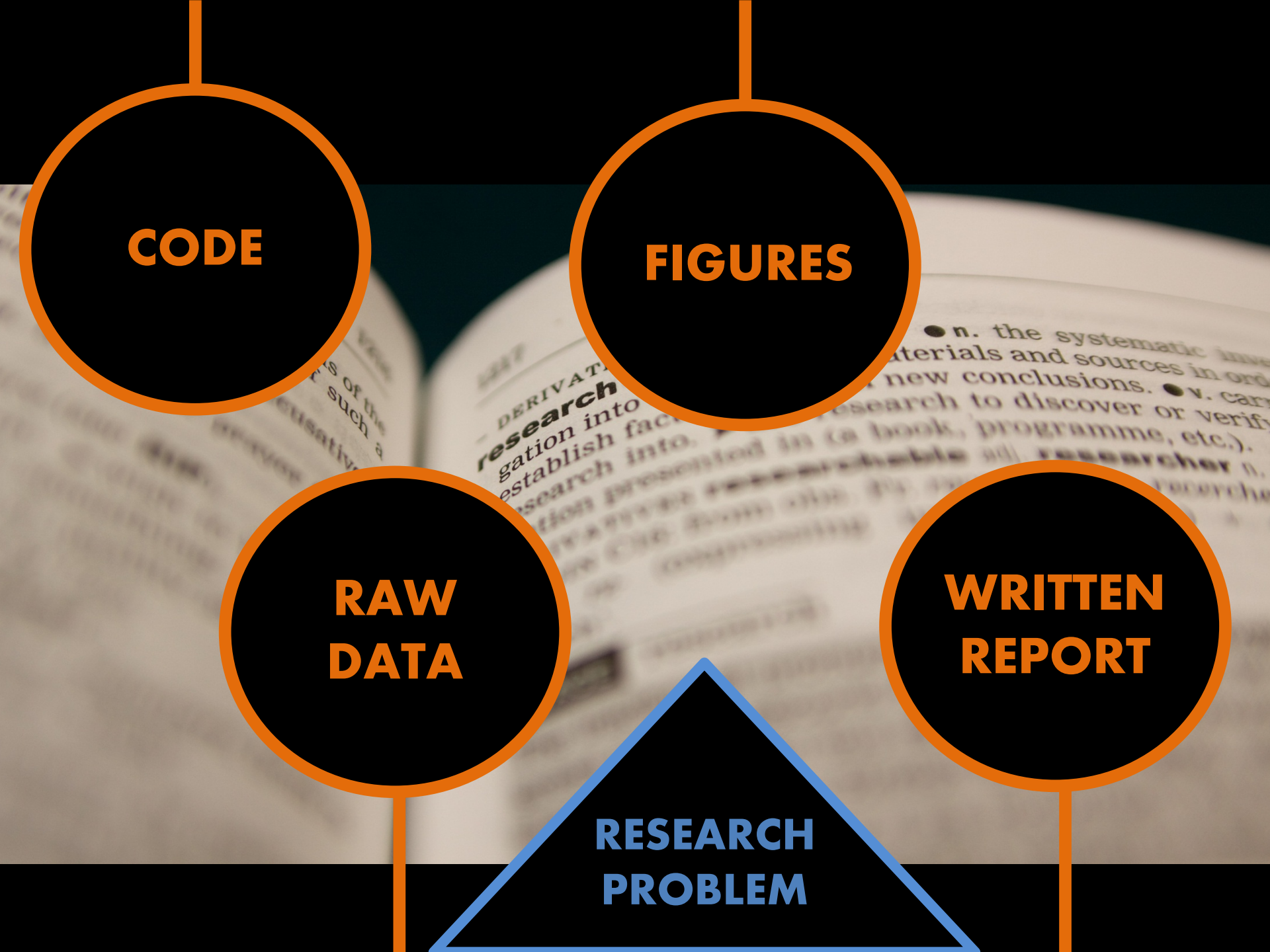
Reproducibility in research



The role of Open Science



Discussion of course assessment


CODE
FIGURES
RAW DATA
WRITTEN REPORT
RESEARCH PROBLEM

Reproducibility

Reproducibility?

Begley & Ellis (2012)
Nature 483, 531-533

COMMENT

AVIAN INFLUENZA Shift expertise to track mutations where they emerge p.534 | EARTH SYSTEMS Past climates give valuable clues to future warming p.537 | HISTORY OF SCIENCE Descartes' lost letter tracked using Google p.540 | OBITUARY Wylie Vale and an elusive stress hormone p.542

Many landmark findings in preclinical oncology research are not reproducible, in part because of inadequate cell lines and animal models.

Raise standards for preclinical cancer research

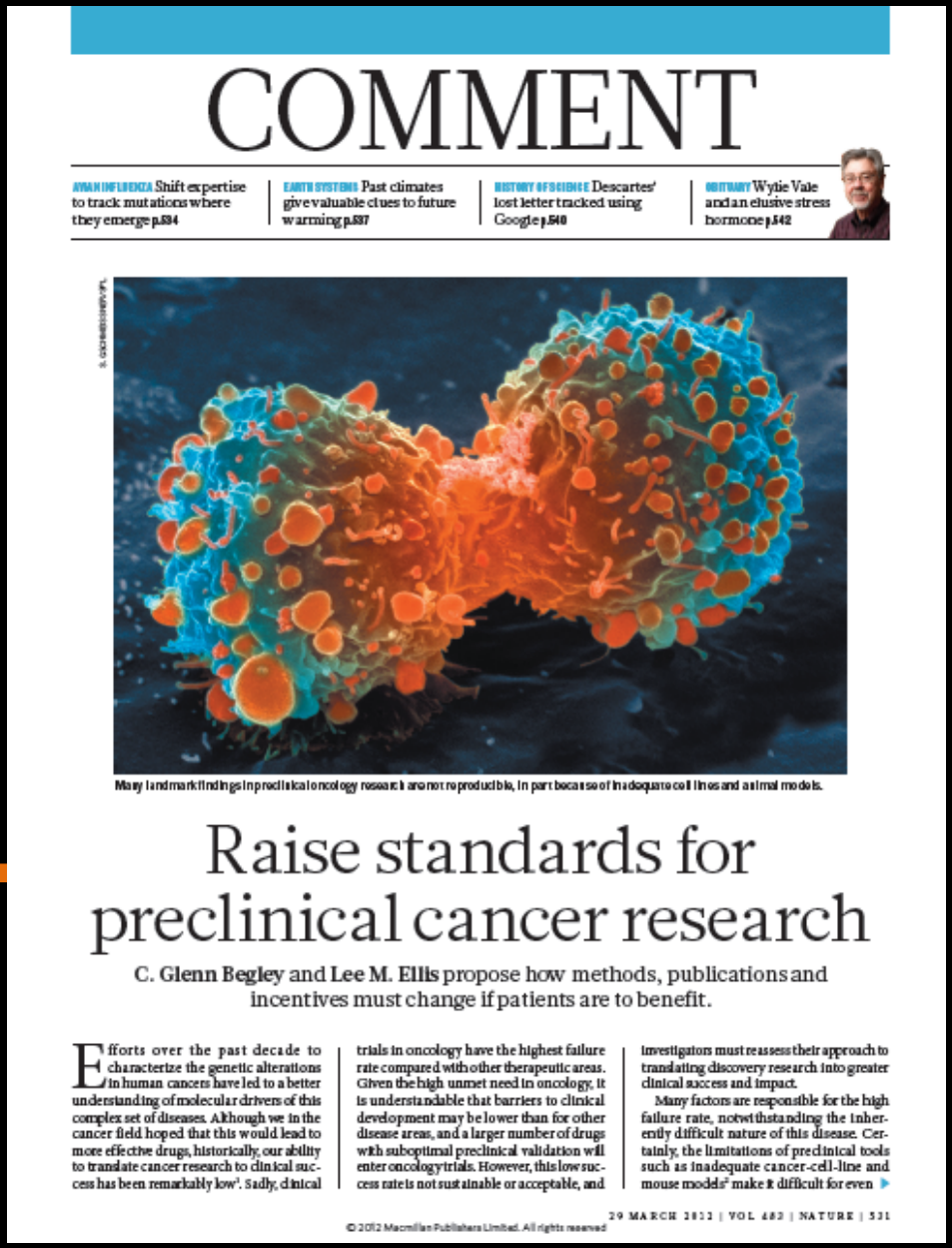
C. Glenn Begley and Lee M. Ellis propose how methods, publications and incentives must change if patients are to benefit.

Efforts over the past decade to characterize the genetic alterations in human cancers have led to a better understanding of molecular drivers of this complex set of diseases. Although we in the cancer field hoped that this would lead to more effective drugs, historically, our ability to translate cancer research to clinical success has been remarkably low[1]. Sadly, clinical trials in oncology have the highest failure rate compared with other therapeutic areas. Given the high unmet need in oncology, it is understandable that barriers to clinical development may be lower than for other disease areas, and a larger number of drugs with suboptimal preclinical validation will enter oncology trials. However, this low success rate is not sustainable or acceptable, and investigators must reassess their approach to translating discovery research into greater clinical success and impact.

Many factors are responsible for the high failure rate, notwithstanding the inherently difficult nature of this disease. Certainly, the limitations of preclinical tools such as inadequate cancer-cell-line and mouse models[2] make it difficult for even ▶

"Some non-reproducible clinical papers have spawned an entire field, with hundreds of secondary publications that expanded on elements of the original observation, but did not actually seek to confirm or falsify its fundamental basis".

Image shown is from front page of Begley & Ellis (2012), produced by the Nature Publishing Group



COMMENT

AVIAN INFLUENZA Shift expertise to track mutations where they emerge p.534 | EARTH SYSTEMS Past climates give valuable clues to future warming p.537 | HISTORY OF SCIENCE Descartes' lost letter tracked using Google p.540 | OBITUARY Wylie Vale and an elusive stress hormone p.542

Many landmark findings in preclinical oncology research are not reproducible, in part because of inadequate cell lines and animal models.

Raise standards for preclinical cancer research

C. Glenn Begley and **Lee M. Ellis** propose how methods, publications and incentives must change if patients are to benefit.

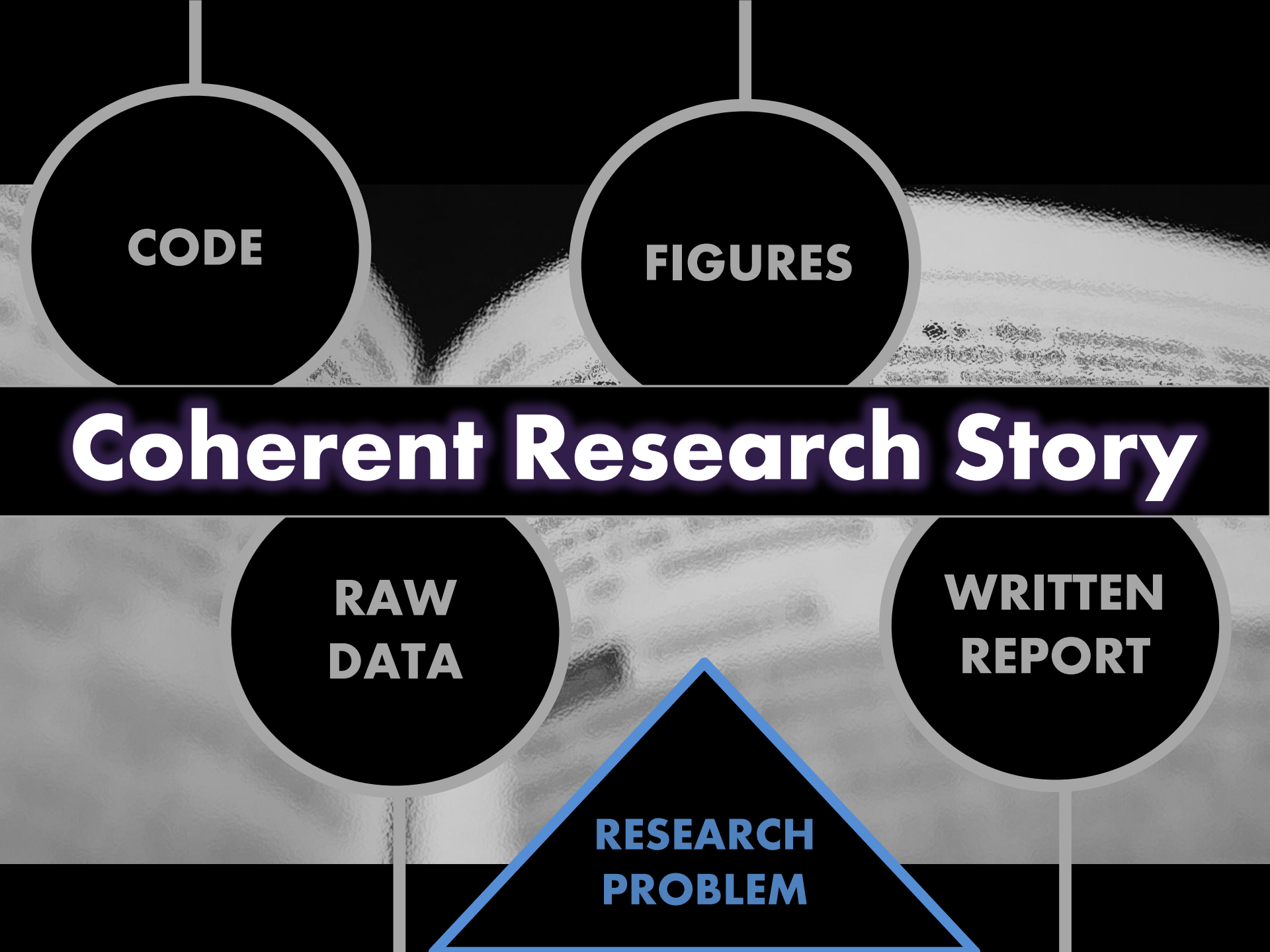
Efforts over the past decade to characterize the genetic alterations in human cancers have led to a better understanding of molecular drivers of this complex set of diseases. Although we in the cancer field hoped that this would lead to more effective drugs, historically, our ability to translate cancer research to clinical success has been remarkably low[1]. Sadly, clinical trials in oncology have the highest failure rate compared with other therapeutic areas. Given the high unmet need in oncology, it is understandable that barriers to clinical development may be lower than for other disease areas, and a larger number of drugs with suboptimal preclinical validation will enter oncology trials. However, this low success rate is not sustainable or acceptable, and investigators must reassess their approach to translating discovery research into greater clinical success and impact.

Many factors are responsible for the high failure rate, notwithstanding the inherently difficult nature of this disease. Certainly, the limitations of preclinical tools such as inadequate cancer-cell-line and mouse models[2] make it difficult for even ▶

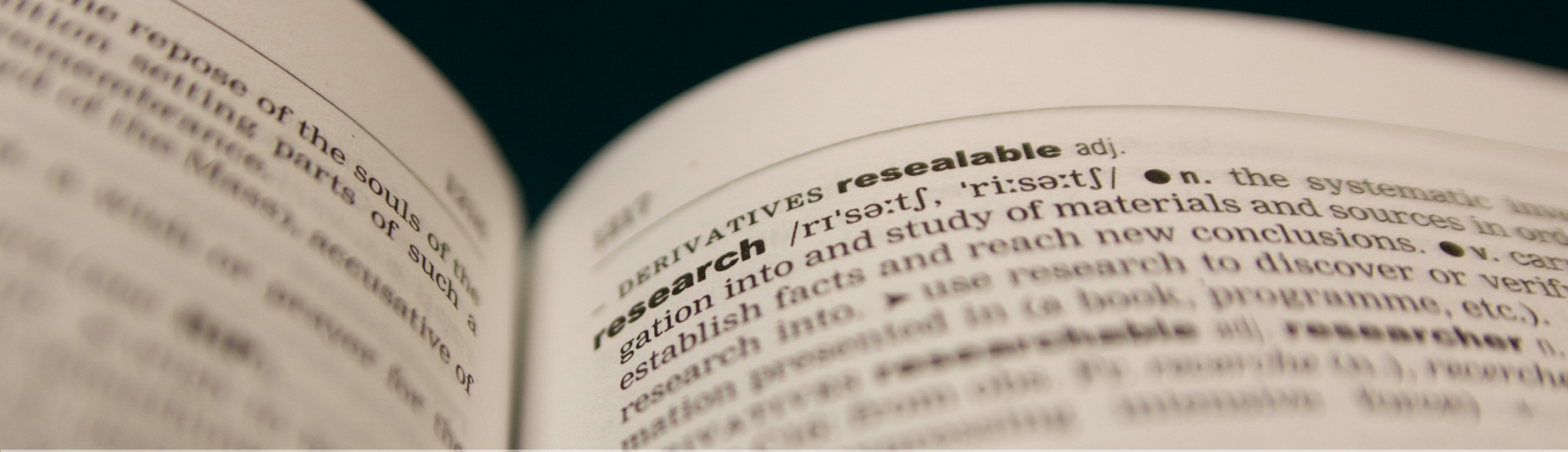
Scientific Publication or Advertisement?

"It lies in our hands – the community of editors and reviewers – to insist that publications are accompanied by the source code and data required to allow their results and conclusions to be reproduced..."

Landrum & Stiefl (2012)
Future Med. Chem.
4(15), 1885-1887


CODE
FIGURES
Coherent Research Story
RAW DATA
WRITTEN REPORT
RESEARCH PROBLEM

The Role of Open Science



"A piece of content or data is open if anyone is free to use, reuse, and redistribute it — subject only, at most, to the requirement to attribute and/or share-alike."

What is openness?

In learning about open science, you'll learn how to effectively and safely share your research outputs, through licensing, code repositories and publication, and in doing so, to foster reproducibility in your work and enhance your research impact.

What is openness?

Users of these slides will need to embed the video from Science Commons here, available on youTube at the following address:
http://www.youtube.com/watch?v=hZAcTNFzF-s

Big Data, Networked Research

And all sorts of other approaches...

- Data Mining
- Crowdsourcing
- Understanding Negative Results
- Web-based Lab Books

Open Science facilitates discovery by supporting the dissemination and availability of information, and enabling its reuse through provision of legal licensing.

Examples of Data-Enabled Research

Søren Brunak, Denmark Technical University

- molecular level sys bio + healthcare sector data.
- Mining electronic health records: towards better research applications and clinical care'

http://www.ncbi.nlm.nih.gov/pubmed/22549152

Chas Bountra,
Structural Genomics Consortium

- accelerate identification of candidate targets for drug discovery by generating freely available novel reagents

- They crystallize 5–10 structures per month!
- SGC is responsible for 25 - 50% of all structures deposited into the Protein Data Bank on human parasites + biomed proteins

UK Research Councils

"...BBSRC will require a copy of any resulting published journal article...to be deposited, at the earliest opportunity, in an appropriate e-print repository..."

BBSRC
bioscience for the future

RESEARCH COUNCILS UK

"Free and open access to the outputs of publicly-funded research offers significant social and economic benefits..."

"All published research articles arising from EPSRC-sponsored research...must become available on an Open Access basis through any appropriate route"

EPSRC
Engineering and Physical Sciences Research Council

Further Reading

Michael Nielsen's book "Reinventing Discovery"