

# The Open Science movement

Maria Panagiotopoulou (ECRIN),  
Birgit Schauffhauser (CHUV)

November 3, 2022  
BOLOGNA, IT



Funded by  
the European Union

eCREAM project  
N. 1010557726



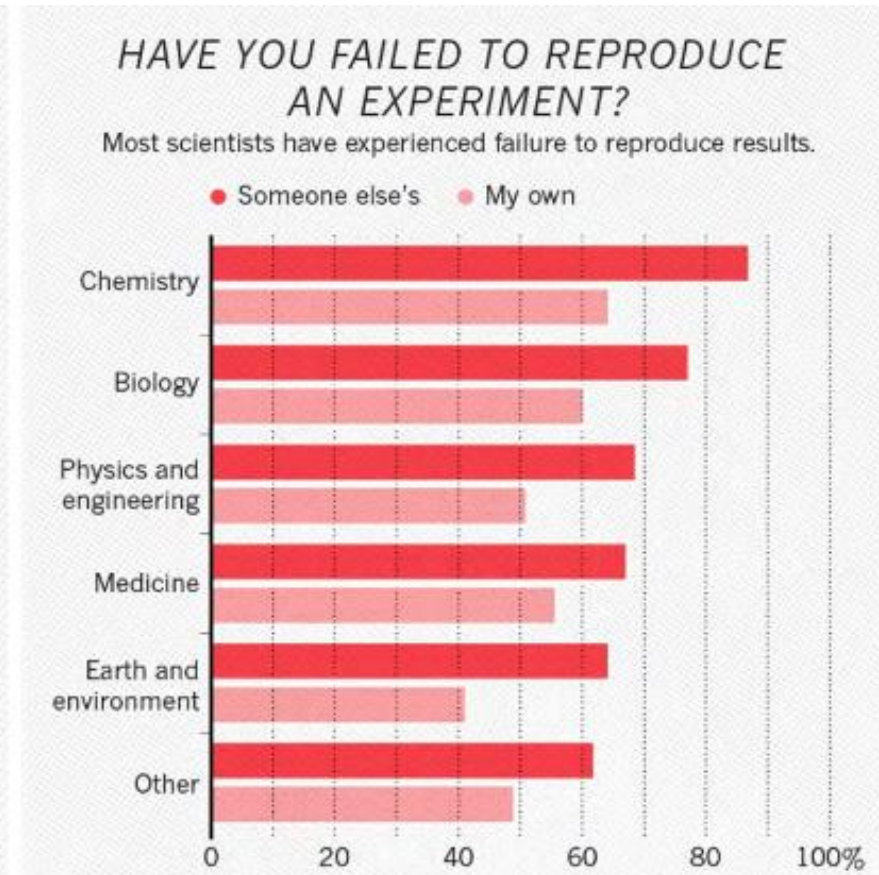
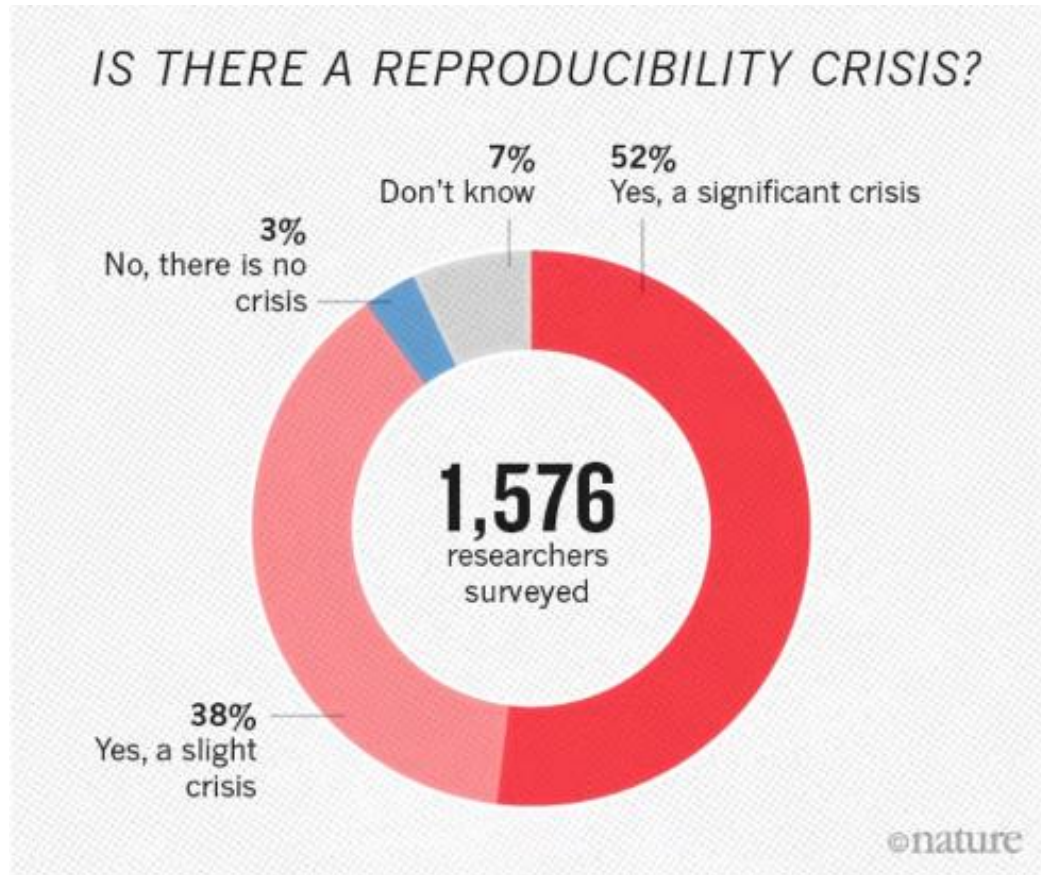
# Open science: what is it?

- “The movement to make scientific research, data and dissemination accessible to all levels of an inquiring society.” (FOSTER Taxonomy)
- “The practice of science in such a way that others can collaborate and contribute, where research data, lab notes and other research processes are freely available, under terms that enable reuse, redistribution and reproduction of the research and its underlying data and methods.” (FOSTER)
- “Open Science aims at transforming science through ICT tools, networks and media, to make research more open, global, collaborative, creative and closer to society.” (EC – Open Science Policy)
- “Research simply done properly”

# Open science: why bother?

2016 Nature survey on science reproducibility

1576 researchers surveyed

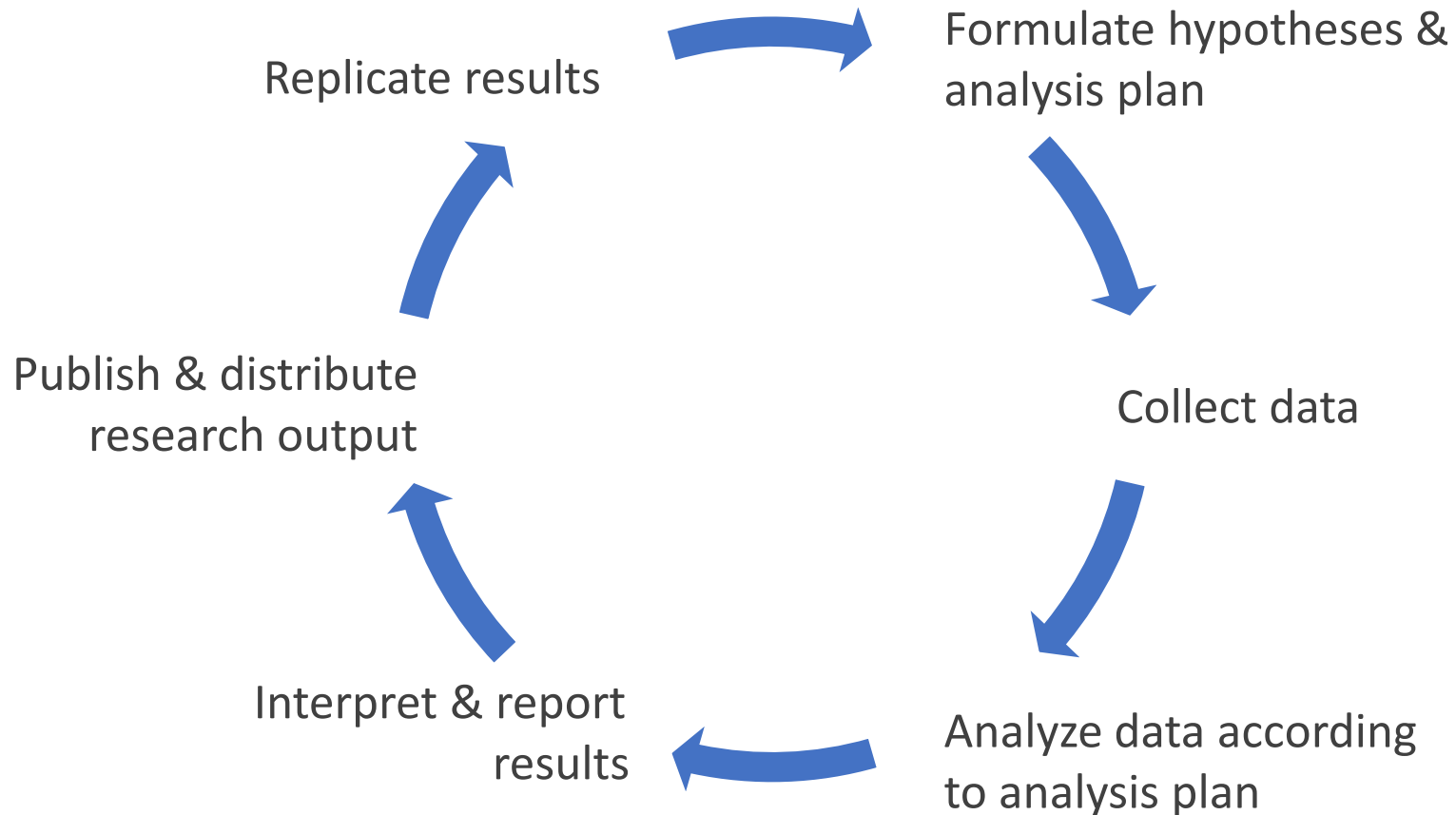


<https://www.nature.com/news/1-500-scientists-lift-the-lid-on-reproducibility-1.19970>

# Open science: why bother?

*Lack of transparency*

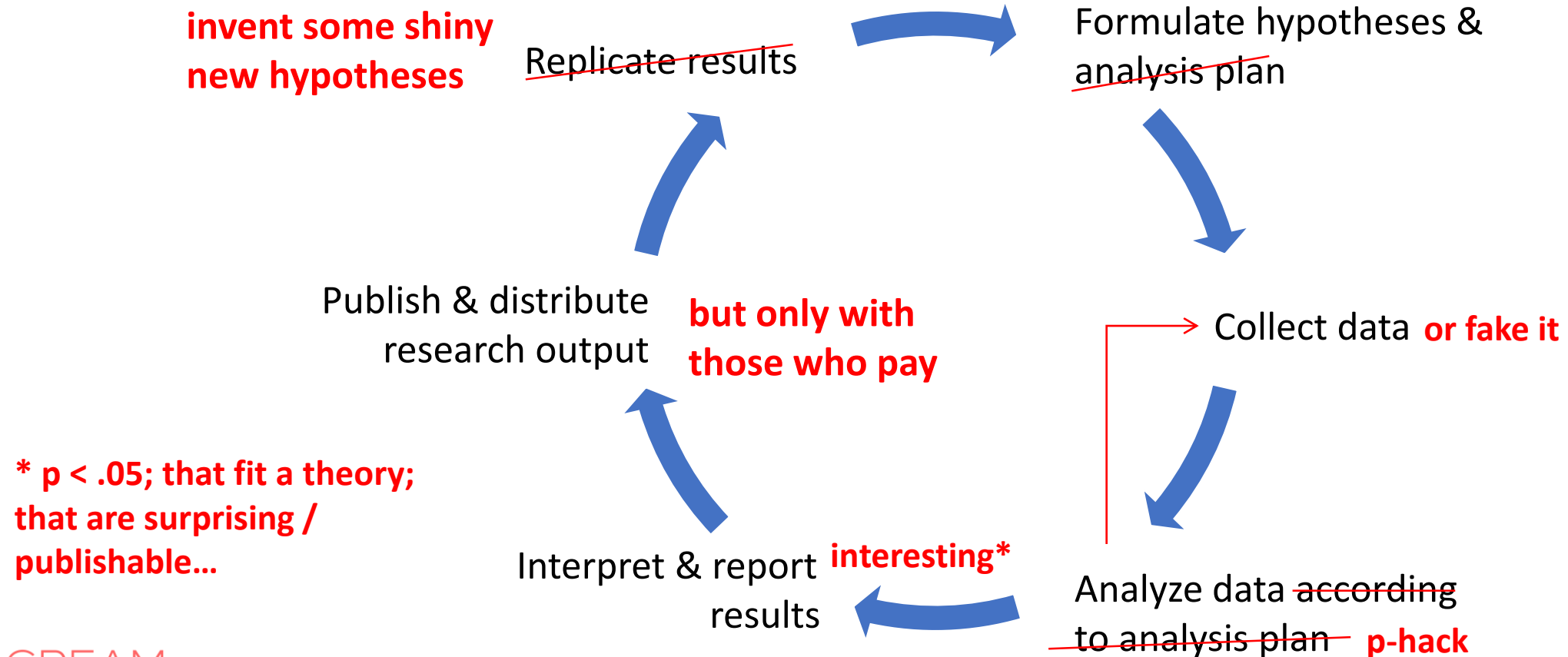
## The Confirmatory Research Process



# Open science: why bother?

## *Lack of transparency*

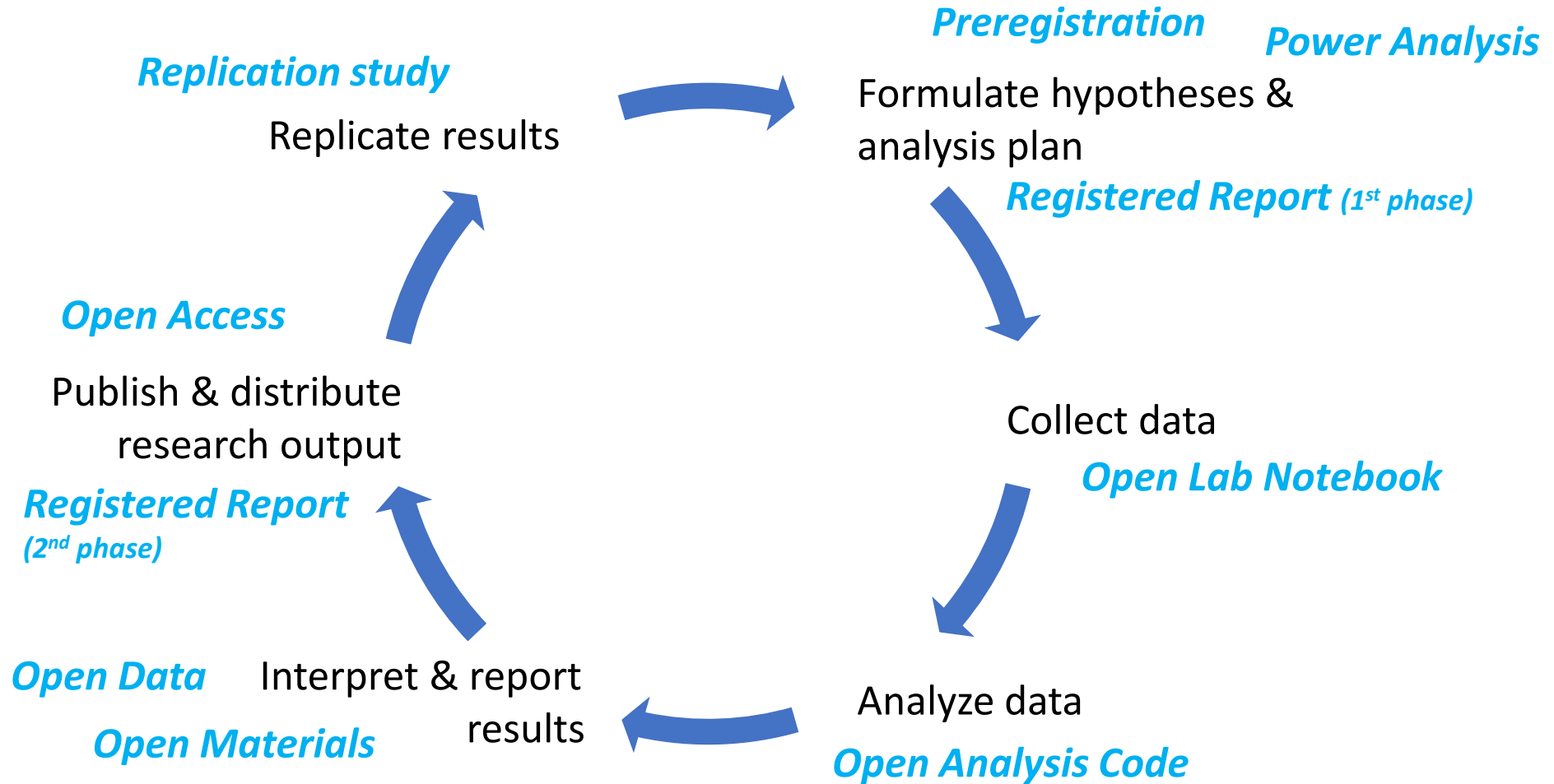
How can you know that it does not look like this?





# Open science: why bother?

## Open science in the research process



# Open science: why bother?

*Avoid research retraction and duplication of funding...*

## Hydroxychloroquine or chloroquine with or without a macrolide for treatment of COVID-19: a multinational registry analysis

Mandeep R Mehra, Sapan S Desai, Frank Ruschitzka, Anne N Patel

### Summary

**Background** Hydroxychloroquine or chloroquine, often in combination with a second-generation macrolide, are widely used for treatment of COVID-19, despite no conclusive evidence of their benefit. Although generally safe when used for approved indications such as autoimmune disease or malaria, the safety and benefit of these treatment regimens are poorly evaluated in COVID-19.

**Methods** We did a multinational registry analysis of the use of hydroxychloroquine or chloroquine with or without a macrolide for treatment of COVID-19. The registry comprised data from 671 hospitals in 34 countries. We included patients hospitalised between Dec 20, 2019, and April 14, 2020, with a positive laboratory finding for SARS-CoV-2. Patients who received one of the treatments of interest within 48 h of diagnosis were included in one of four treatment groups (chloroquine alone, chloroquine with a macrolide, hydroxychloroquine alone, or hydroxychloroquine with a macrolide), and patients who received none of these treatments formed the control group. Patients for whom one of the treatments of interest was initiated more than 48 h after diagnosis or while they were on mechanical ventilation, as well as patients who received remdesivir, were excluded. The main outcome of interest was in-hospital mortality and the occurrence of de-novo ventricular arrhythmias (as defined as sustained ventricular tachycardia or ventricular fibrillation).

**Findings** 96032 patients (mean age 53·8 years, 46·3% women) with COVID-19 were hospitalised during the study period and met the inclusion criteria. Of these, 10 698 patients were in the treatment groups (3868 received chloroquine, 3783 received chloroquine with a macrolide, 3016 received hydroxychloroquine, and 6221 received hydroxychloroquine with a macrolide) and 85334 were in the control group. 10 698 (11·1%) patients died in hospital. After controlling for multiple confounding factors, including age, sex, race or ethnicity, body-mass index, underlying cardiovascular disease and its risk factors, diabetes, existing lung disease, smoking, immunosuppressed condition, and baseline disease severity, we compared in-hospital mortality in the control group (9·3%), hydroxychloroquine (18·0%; hazard ratio 1·335, 95% CI 1·236–1·447), hydroxychloroquine with a macrolide (23·8%; 1·447, 1·368–1·531), chloroquine (16·4%; 1·365, 1·218–1·531), and chloroquine with a macrolide (22·2%; 1·368, 1·273–1·469) were each independently associated with an increased risk of in-hospital mortality. Compared with the control group (0·3%), hydroxychloroquine (6·0%; 2·56–14·935–2·904), hydroxychloroquine with a macrolide (8·1%; 3·196, 4·104–5·983), chloroquine (4·3%; 1·81, 1·324–4·596), and chloroquine with a macrolide (6·5%; 4·011, 3·344–4·812) were independently associated with an increased risk of de-novo ventricular arrhythmias during hospitalisation.

**Interpretation** We found no evidence to confirm a benefit of hydroxychloroquine or chloroquine, when used alone or with a macrolide, on in-hospital outcomes for COVID-19. Each of these drug regimens was associated with decreased in-hospital mortality and increased frequency of ventricular arrhythmias when used for treatment of COVID-19.

**Funding** William Grey Distinguished Chair in Advanced Cardiovascular Medicine at Brigham and Women's Hospital.

Copyright © 2020 Elsevier Ltd. All rights reserved.

### Introduction

The absence of an effective treatment against severe antiviral properties as well as immunomodulatory



nature

Explore content

About the journal

Publish with us

[nature](#) > [news](#) > article

Published: 19 December 2013

## Scientists losing data at a rapid rate

Elizabeth Gibney & Richard Van Noorden

[Nature](#) (2013) | [Cite this article](#)

656 Accesses | 9 Citations | 595 Altmetric | [Metrics](#)

Decline can mean 80% of data are unavailable after 20 years.

[Scientists losing data at a rapid rate. Nature \(2013\). https://doi.org/10.1038/nature.2013.14416](#)



The EC estimates that **2.6 billion euros per year** could be saved if open science practices would be broadly implemented.

[Cost of not having FAIR research data](#), Published: 2019-01-16.

Column: How a retracted research paper contaminated global coronavirus research

# Sharing data in medical sciences

- In 2016, the International Committee of Medical Journal **Editors** (ICMJE) published an editorial stating that '*it is an ethical obligation to responsibly share data generated by interventional clinical trials because participants have put themselves at risk*'.
- The **general public** is positive towards health data sharing for research (~70-90%). Several international studies report this *high willingness which is motivated by a so called 'reciprocity' position: patients who benefit from medical research should contribute to research themselves*.

Whitepaper: Practical challenges for researchers in data sharing; 2018. @springernature

Figure 11: Problems in sharing data in the medical sciences (n=2,683)

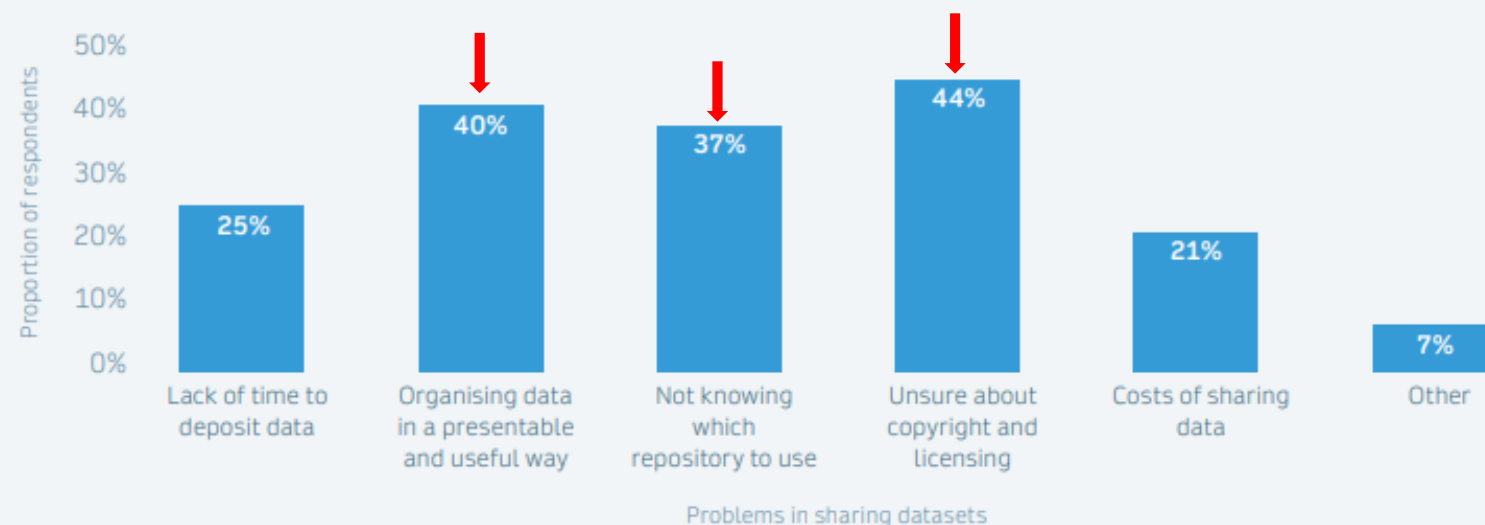
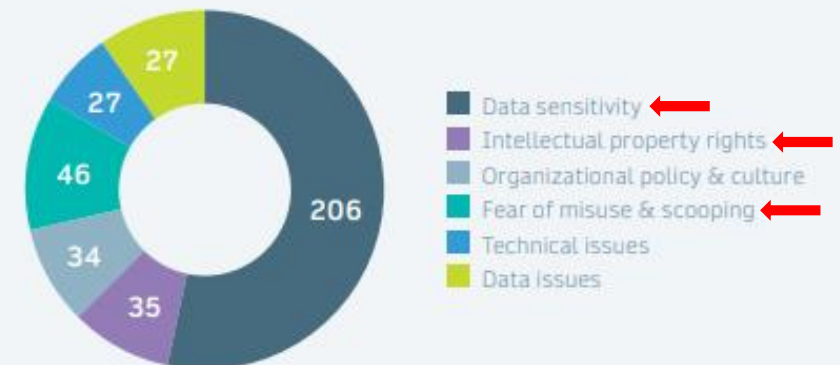


Figure 4: 'Other' problems mentioned in sharing datasets mentioned (n=385)

Number of mentions







# The General Data Protection Regulation (GDPR)

- EU-wide data protection regulation that came into force on **25 May 2018**
- Enables data subjects to have greater control over their **personal data**, whilst **modernising** and **unifying** European data protection rules
- **Permits EU Member States – in certain areas – to make specific domestic provisions for particular aspects of the GDPR**
- Applies to **personal data** and data of **living persons**
- Applies to:
  - ✓ **data controller / data processor** in the EU who collects personal data about a data subject of any country, anywhere in the world
  - ✓ data controller or data processor based outside the EU but collects personal data on EU citizens



# Defining some terms...

**Personal data:** any information relating to an identified or identifiable natural person ('data subject')

- *General: Name, gender, age, date of birth, civil status, nationality, languages, IP addresses*
- *Organisational: work/home addresses, phone, e-mail, id number*
- ***“Special categories”:** Race, Religion, sexual orientation, **health**, sex life, criminal record, biometric data*

**Pseudonymised data:** the personal data can no longer be attributed to a specific data subject without the use of additional information ([in the scope of the GDPR](#))

**Anonymised Data:** data that cannot identify individuals in any way - anonymisation irreversibly destroys any way of identifying the data subject ([out of the scope of the GDPR](#))



# Principles for processing personal data

## 1. Process lawfully, fairly and transparently

*Inform the participant of what will be done with the data, process accordingly*

## 2. Keep to the original purpose

*Collect data for specified, explicit and legitimate purposes; do not process further in a manner incompatible with those purposes*

## 3. Minimise data size

*Personal data collected should be adequate, relevant and limited to what is necessary*

## 4. Uphold accuracy

*Personal data should be accurate and kept up to date*

## 5. Remove data which are not used

## 6. Ensure data integrity and confidentiality

*Protection against unauthorised or unlawful processing, accidental loss, destruction or damage, using appropriate technical or organizational measures*



# Medical data cannot be “open”

## “Open” vs FAIR

- **Findable:** “Metadata and data should be easy to find for both humans and computers. Machine-readable metadata are essential for automatic discovery of datasets and services.”
- **Accessible:** “Once the user finds the required data, she/he needs to know how can they be accessed, possibly including authentication and authorisation.”
- **Interoperable:** “The data usually need to be integrated with other data. In addition, the data need to interoperate with applications or workflows for analysis, storage, and processing.”
- **Reusable:** “The ultimate goal of FAIR is to optimise the reuse of data. To achieve this, metadata and data should be well-described so that they can be replicated and/or combined in different settings.”

## TO BE FINDABLE:

F1. (meta)data are assigned a globally unique and eternally persistent identifier.

F2. data are described with rich metadata.

F3. (meta)data are registered or indexed in a searchable resource.

F4. metadata specify the data identifier.

## TO BE ACCESSIBLE:

A1 (meta)data are retrievable by their identifier using a standardized communications protocol.

A1.1 the protocol is open, free, and universally implementable.

A1.2 the protocol allows for an authentication and authorization procedure, where necessary.

A2 metadata are accessible, even when the data are no longer available.

## TO BE INTEROPERABLE:

I1. (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.

I2. (meta)data use vocabularies that follow FAIR principles.

I3. (meta)data include qualified references to other (meta)data.

## TO BE RE-USABLE:

R1. meta(data) have a plurality of accurate and relevant attributes.

R1.1. (meta)data are released with a clear and accessible data usage license.

R1.2. (meta)data are associated with their provenance.

R1.3. (meta)data meet domain-relevant community standards.

Wilkinson, M., Dumontier, M., Aalbersberg, I. *et al.* The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* **3**, 160018 (2016).  
<https://doi.org/10.1038/sdata.2016.18>

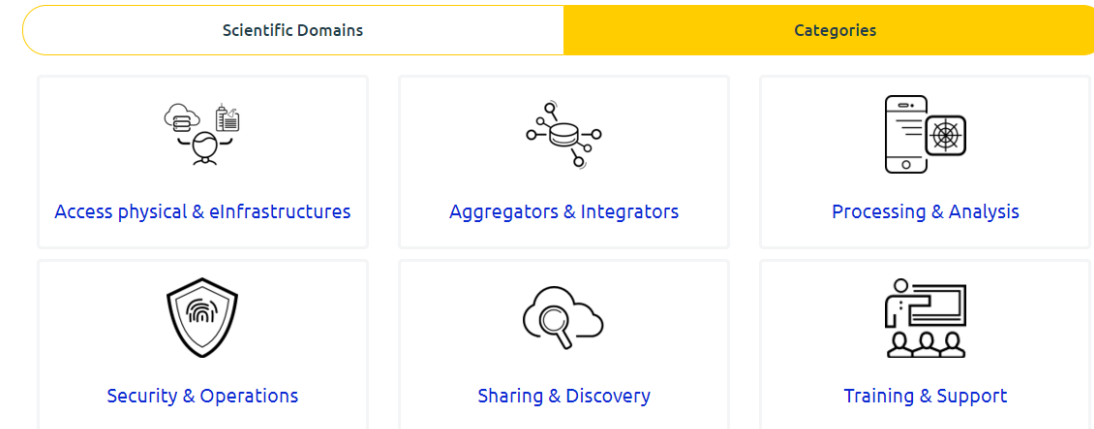
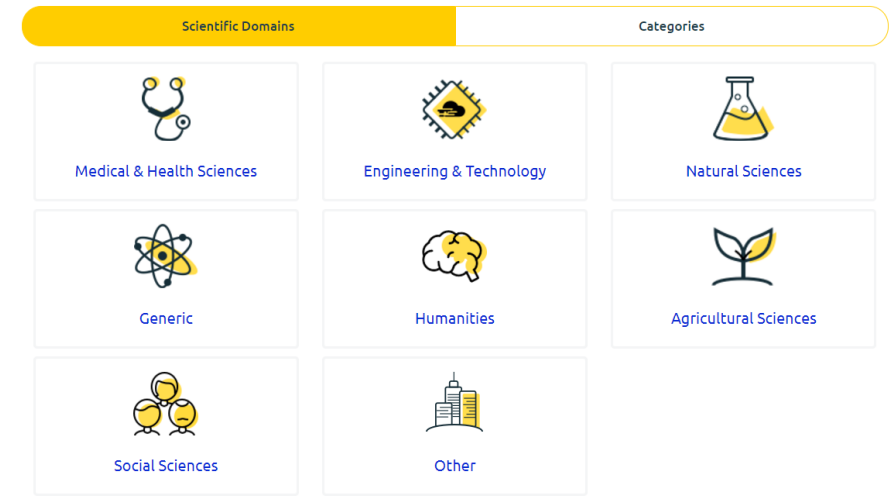


# The European Open Science Cloud (EOSC)

EC and member state initiative started taking shape in 2015 as a vision of a large infrastructure to support and develop **open science** and **open innovation** in Europe and beyond.

Realized in 2021 and will be Europe's virtual environment for all researchers to **store**, **manage**, **analyse** and **re-use data** for research, innovation and educational purposes.

*“to offer to the 1.7 million European researchers and 70 million professionals in science and technology a virtual environment with free at the point of use, open and seamless services for storage, management, analysis and re-use of research data, across borders and scientific disciplines”*



# The European Health Data Space (EHDS)

**Strengthens the rights of individuals in relation to greater control over their electronic health data:**

*Access, share health data with health professionals nationally or cross-border, add information, rectify errors, restrict access, know what health professional accessed data, issue and accept health data in a common European format, strengthen interoperability.*



**Rules for  
electronic health  
record systems  
(EHR systems)**

**Rules and  
mechanisms  
supporting the  
secondary use of  
electronic health  
data**

**Mandatory cross-border  
infrastructures for  
primary and secondary  
use of health data**

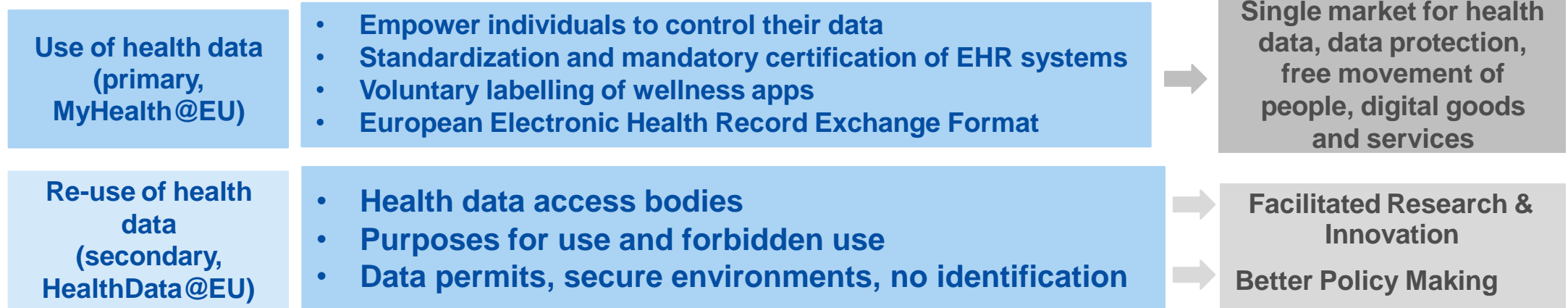
- MyHealth@EU
- HealthData@EU

# The European Health Data Space (EHDS)

## OBJECTIVES

Effective use of health data

## SCOPE & EXPECTED IMPACT



## MEANS





**THANK YOU**  
**FOR YOUR**  
**ATTENTION**

Maria Panagiotopoulou  
[Maria.panagiotopoulou@ecrin.org](mailto:Maria.panagiotopoulou@ecrin.org)