Tim Friede (Göttingen University)

Reproducibility, open data and evidence synthesis in clinical research

Topic 1 – Bringing in information from where we can get it

Keywords: data repositories, electronic health records, clinical trials, meta-analysis

Introduction

Quantitative empirical sciences are among the pillars of modern medicine and biomedical research in particular. The ultimate aim is to improve patient care by basing decision-making in clinical practice and in health care provision in general on sound, reliable evidence gained through well designed, conducted and analysed research studies. To this end the evidence-based medicine approach has developed a "hierarchy of evidence" with meta-analyses of randomised controlled trials at the top followed by individual randomized controlled trials and non-randomised studies down to case series and expert opinion at the bottom.

The evidence on a particular research question available at a certain timepoint is summarized in so-called systematic reviews which can include meta-analyses formally integrating data from various studies on a particular research topic.

Methods / Problem statement

There is a crisis in medical research including basic sciences as well as clinical research in that it is perceived as wasteful (or at least inefficient) and in that results are often not reproducible.

Results / Proposed solution

A number of approaches have been suggested to make research more reproducible including the publication of data. In the context of publishing in biostatistics, for instance, a number of journals have implemented reproducible research policies which include the publication of software code and data (see e.g. Hothorn et al, 2009).

In the context of clinical trials the pressure have significantly increased over the past years to provide access to trial data which led to the establishment of data repositories (see e.g. Sudlow et al, 2016). These enable meta-analyses of the available evidence which is of particular interest in situations where evidence is scarce, e.g. in small populations and rare diseases (Friede et al, 2016). Also routinely collected data are more readily available for research purposes, for instance in the form of electronic health records, which again can be of great value in particular in rare diseases (Chataway and Friede, 2016). These developments have the potential of making research more efficient.

At the same time, however, sharing and reusing data carries some risks in terms of protecting the privacy of individuals and informed consent of patients, both key concepts in enabling clinical research. This in term makes it necessary to develop and implement new concepts to protect the interests of patients (see e.g. Wegscheider and Friede (2016) for a discussion).

Conclusions

The advance in data capturing and storage combined with increasing analytical ability carries some promise to make clinical science more efficient and to strengthen the evidence base. At the same time these

developments require a new form a governance in terms of use and access of data to protect patients' interests.