Are Real World Data the smart way of doing Health Analytics?

Real World Data: la base di una nuova ricerca clinica?

Francesca Ieva

Abstract Real world data (RWD) and real world evidence (RWE) are playing an increasing role in health care. Despite their use may strongly improve healthcare research as well as health related decision making, management and planning, many barriers remain to their use in clinical practice. The aim of this paper is to discuss some issues related to RWD aiming at proposing a new paradigm of healthcare research based on RWE.

Abstract *I dati dal mondo relae e le relative evidenze scientifiche stanno avendo un ruolo di sempe maggiore importanza nello scenario odierno della ricerca in ambito clinico. Nonostante la diffusione di tali dati possa comportare significativi miglioramenti nella ricerca sanitaria e nelle politiche decisionali e di gestione nell'ambito della salute, ancora oggi permangono molte barriere alla diffusione e al consolidamento del loro utilizzo nella prassi clinica. Lo scopo di questo lavoro discutere alcune questioni relative ai dati del mondo reale al fine di proporre un nuovo paradigma di ricerca nell'ambito della salute, basato sulle evidenze generate dai dati del mondo reale.*

Key words: Real World Data, Real World Evidence, Health Analytics

1 Background and setting

The terms "Real World Data" (RWD) and "Real World Evidence" (RWE) refer to the widely heterogeneous amount of data arising from current practice in any area, and to the evidence that may be pointed out applying suitable statistical methods to their analysis. RWD are playing an increasing role in health care research nowadays. The health care community is using RWD to support coverage decisions and to

Francesca Ieva
MOX - Modelling nd Scientific Computing, Deprtment o Mathematics, Politecnico di Milano, via Bonardi 9, 20133, Milano (IT), e-mail: francesca.ieva@polimi.it

develop guidelines and decision support tools for use in clinical practice, to monitor postmarket safety and adverse events and to make regulatory decisions. Medical product developers are using RWD and RWE to support clinical trial designs and observational studies to generate innovative, new treatment approaches.

In other words, healthcare is rapidly transitioning to a new world of patient choice with a focus on outcomes and value, being value

$$VALUE = \frac{CLINICALBENEFITS}{DIRECTCOSTS}$$

Therefore, suitable indicators for "value" and suitable systems for monitoring processes, utilizations and performances in healthcare are more and more needed. Moreover, the use of computers, mobile devices, wearables, and other biosensors to gather and store huge amounts of health-related data has been rapidly accelerating.

Starting from these considerations, the "Real World" term and related data have to be properly defined, in order to give birth to a new definition of clinical evidence, that is real world based. The aim of this paper is to point out a discussion on what RWD are in the actual context, and how they may lead to a convincing Real World Evidence, in order to set a new paradygm for a smart research in clinical biostatistics.

2 Real World Data in practice

There is no univocal definition of what RWD are in practice. It is pretty well known the potential of RWD, especially into the healthcare setting, as well as the issues related to their use. Nevertheless, a clear definition of what RWD are has still to come. This is mainly due to the quick technological evolution which creates new type of data and new potential in data collection day by day.

In general, RWD are the data relating to patient health status and health consumption, as well as the delivery of health care routinely collected from a variety of sources. Among others, the main sources are: Electronic health records (EHRs), claims and billing activities, clinical registries, patient-generated data and data gathered from external sources that can inform on health status, such as mobile devices.

For these reasons, RWD sources generally fall into four categories, which are likely to be expanded in the future:

- Clinical data
- Administrative/claims data
- Patient-generated/reported data
- Non-traditional, health-related digital data sources

Clinical data. Patient-level data pulled from Electronic Medical Records (EMR) and clinical registries that describe how patients are treated. They include lab values, diagnoses, notes and not structured data, and other information from healthcare

visits with physicians.

Administrative/claims data. Patient-level data collected for non-clinical purposes, primarily for billing by providers to insurers and other payors, which can include diagnoses, services provided, costs, and other data required for the reimbursement of healthcare services. In general, the term "administrative" refers to information collected by government departments and other organisations primarily for administrative (i.e., not research) purposes.

Patient-generated/reported data. Individual data describing the patient's experience.

Non-traditional, health-related digital data sources. As digital becomes increasingly prevalent in our lives, new sources of patient-level health data are emerging. These span social media posts that have a rich trove of information, especially health-focused social media sites.

The richness of information contained in such data is undeniable. It can be said that RWD holds potential to allow us to better design and conduct clinical trials and studies in the health care setting to answer questions previously though infeasible. All these statements are true, provided we unlock such potential and properly understand what it may and should be used for, and which methods should be developed and then applied to RWD.

3 Smart use of Real World Data in Healthcare Research

To unlock the potential of RWD, we have to point out i) who would benefit from their use; ii) which barriers their use may encounter, and iii) which are the question they are able to answer at best.

Starting from the latter point, RWD should be used to monitor the real use of healthcare (in terms of treatments, drugs, practices, healthcare paths, ...) pursued by a given population of interest. Such monitoring is needed to verify the applicability of guidelines arising from traditional clinical trials in a real world setting. In other words, RWD enables the evaluation of clinical practices and, in doing so, they represent a concrete support to decision and policy making.

The reasons why their use is still not intensive in healthcare research are many. Among the most significant barriers to expanding use of real-world data is the consensus that randomized controlled trials (RCT) remain the gold standard for demonstrating the efficacy and safety of medical products and treatments. This consensus, shared by physicians, patients, payors and regulators, creates significant hurdles to using RWD and to set a RWE-based decision making, even though there is a growing recognition that RCT alone cannot provide sufficient data for informed healthcare decision making in some situations. Limitations can also be attributed to a lack of

common technologies used across institutions that collect the initial patient data, and then to the uneven quality of such data. Moreover, the accessibility of RWD, which still represent one of the main issues for using them as basis in healthcare research, differs consistently from region to region and from country to country. Last but not least, the lack of standardization of RWE analytics makes stakeholders doubtful about the use of RWE in healthcare research.

This last point represent a strong limitation, since leads patients and clinicians not to understand that they would be the first and main beneficiaries of the use of RWD in clinical practice and healthcare research. In fact, the use of RWD in regulations and decision making is something that, in the end, would improve the quality of healthcare offered to patients and the efficiency of the system the clinicians work in.

4 What's next?

Expanding the use of RWE requires a multi-stakeholder action aimed at:

Increasing understanding and communication of RWE value. RWE analytics delivers valuable information, which healthcare researchers and data scientists are responsible for getting into the hands of payors, healthcare providers, and regulators to improve their healthcare decision making.

Shaping an integrated, adaptive partner ecosystem. Collaborations among academy, reference administrations and companies are essential to ensure credibility and trust in analyses, as well as gain access to novel data sources. Partnerships with database owners will be required in the short term to use data in restricted access databases, especially large, government-funded databases of public health systems in Europe which typically restrict access to this data.

Building platforms at scale to manage and analyze data in a rapid, low-cost fashion. RWE would incredibly benefit of platforms incorporating standardized methodology, which can be applied across all studies, improving the robustness and credibility of outputs.

5 Conclusions

Modern healthcare is undergoing a huge shift in how key stakeholders are approaching and evaluating patient data. An indicator for this shift is the growing access to and use of RWD, such as de-identified data collected from registries, electronic health records, wearable devices, and administrative and healthcare claims databases. RWE could significantly improve healthcare decisions across the health

system and ultimately improve patient care, provided we unlock the potential of RWD.

To achieve this goal, data integration and appropriate statistical techniques are essential ingredient, since the challenge of getting reliable evidence by RWD passes through:

- achieving medical-level accuracy;
- integrating data from multiple sources;
- collaborating with numerous stakeholders holding the data;
- removing confounders;
- developing suitable statistical methods;

complying the regulatory requirements of the context the data are generated from.

Making progress on this will establish the kind of culture where RWE is needed for making a new paradigm of healthcare research feasible.

References

1. Armitage, P., Berry, G., Matthews, J.N.: Statistical Methods in Medical Research, 4th Edition, Wiley-Blackwell (2001) ISBN: 978-0-632-05257-8
2. Corrao G., Mancia G.: Generating Evidence From Computerized Healthcare Utilization Databases. Hypertension, **65**: 490–498 (2015).
3. Ieva, F., Gasperoni, F.: Discussion of the paper "Statistical challenges of administrative and transaction data" by David J. Hand. Journal of the Royal Statistical Society - Series A (2018) doi: 10.1111/rssa.1231
4. Hand, D.J.: Statistical Challenges of administrative and transaction data. Journal of the Royal Statistical Society - Series A (2018) doi: https://doi.org/10.1111/rssa.12315
5. Hess L.M., Raebel M.A., Conner D.A., Malone D.C., Measurement of adherence in pharmacy administrative databases: a proposal for standard definitions and preferred measures. Annals of Pharmacotherapy, **40**: 1280–88 (2006)
6. Karve S., Cleve M.A., Helm M. et al., Prospective Validation of Eight Different Adherence Measures for Use with Administrative Claims Data among Patients with Schizophrenia. Value In Health **12**(6) (2009).
7. Mazzali, C. et al., Methodological issues on the use of administrative data in healthcare research: the case of heart failure hospitalizations in Lombardy region, 2000 to 2012. BMC Health Service Research, **16** (234) (2016).
8. Velentgas, P., Dreyer, N.A., Nourjah, P., Smith, S.R., Torchia, M.M., eds. Developing a Protocol for Observational Comparative Effectiveness Research: A Users Guide. AHRQ Publication No. 12(13)-EHC099. Rockville, MD: Agency for Healthcare Research and Quality; January 2013. www.effectivehealthcare.ahrq.gov/Methods-OCER.cfm.