The use of real world data throughout an innovative medicine's lifecycle

1. Introduction and objectives

The challenge for health policies is to provide high quality of care for all, within a sustainable health system. Innovations In healthcare such as innovative medicines play a crucial role in improving population's health. The way these medicines are developed, their price and their usage in daily practice can strongly impact on the quality and the sustainability of our health systems. Improved policies are needed to ensure timely patient access to innovative therapies especially in areas of unmet need. Initiatives such as the European medicine's Agency's Adaptive Pathways pilot [i] and PRIority MEdicines (PRIME) [ii] aim at achieving this ambition. However, the generation of evidence for these innovations remains a challenge, especially for rare diseases and for personalised medicine where the patient populations are often small.[iii]

There is an increased interest in the use of real world data (RWD) to support the continuum of evidence generation for innovative medicines.[iv] It is expected for instance that RWD should enable the generation of additional evidence post launch, inform dynamic price-setting in relation to the value of medicines and may optimise appropriate use in daily practice. However, several challenges emerge, such as how to manage expectations about the use of such data, how to better understand their usefulness and their pitfalls throughout an *entire* medicine's lifecycle (and not just post-launch), and how to encourage their optimal use. From the report *Review of current policies/perspectives* from the Innovative Medicine's Initiative (IMI) Get Real initiative, it becomes clear that there is a need for common understanding, reaching consensus on the relevance of RWD, and harmonising the requirements and improved methods and governance.

The purposes of this paper are

a. to discuss the usefulness of RWD throughout the lifecycle of innovative medicines, thereby providing realistic expectations about their possibilities and pointing to their limitations;

b. to list the current issues in the collection, interpretation and implementation of RWD;

c. to propose principles of good practice and necessary actions to improve the use of RWD throughout the lifecycle of innovative medicines.

The emphasis is on the lifecycle application of RWD, hence not only to demonstrate effectiveness of medicines in real practice in the later phases of development. Wherever possible, the paper will also touch upon specific needs, characteristics and challenges in respect to RWD in the field of personalised medicine.

2. Definition of real world data

There is debate about the exact definition of RWD. For instance, RAND Europe defines RWD as an umbrella term referring to "any data not collected in 'conventional randomised controlled trials (RCTs)'. It may include data from existing secondary sources (e.g. databases of national health services) and the collection of new data, both retrospectively and prospectively." [v]

Obviously, there are many types of real world data coming from different sources. In fact, a continuum exists between data collected in strictly controlled conditions on the one hand, and data reflecting what happens in daily practice on the other hand (Figure 1).

Figure 1: the continuum between data collected in strictly controlled conditions and data reflecting what happens in daily practice

On the far-left side of the continuum, patients are heavily selected in order to avoid confounding effects, and data are collected while following strict protocol instructions and mandatory actions. On the far-right side of the continuum, data are collected based on observing daily life practice in patients, for instance based on registries or existing databases. In between we find, for instance, pragmatic trials whereby there is still 'control', due to patient selection and protocol instructions, but where attempts are made to better mimic real practice.

It is clear that RWD can come in different formats (e.g. registries) and may have different types of content (e.g. resource use or patient reported outcomes). Outlined below is a non-exhaustive list.

a. Data obtained from Electronic Patient Records (EPRs)

When connected to a central database, the use of EPRs allows a thorough insight into real life practice, since they combine diagnostic information and medical practice data in a longitudinal manner. Hence, patients can be tracked over time. The use of this kind of data can be situated across the lifecycle of medicines, for instance during the development process to identify patients for inclusion in clinical trials, during the market access phase (where evidence needs to be provided to inform pricing and reimbursement decisions) to help better understand the current practice and standard of care, and during the market usage phase (post launch) to assess the performance of new medicines in real life.

b. Data obtained from claims databases

These data, typically from health insurers, provide insights into medical resource use and offer similar possibilities as EPR based data. As is the case with EPRs, claims data are typical types of routinely collected data. However, these claims databases often lack diagnostic information.

c. Genetic and biomarker data

These are datasets holding rich information about genetic characteristics (for instance of patients or of cancer cells) and other biomarkers. These 'bundles of information' have a special use in the field of personalised medicine and can be very useful in the development phase of a medicine to better understand the relationship between genetic characteristics and treatment outcome.

d. Data from registries (existing or newly set up)

In several countries, registries in different disease areas have been established which are maintained and updated with new information and patients. Registries can also be newly created for a given purpose. Typically, they play a role during the market access phase to understand current practice, but even more during the market usage phase (post launch) to assess the performance of new medicines in real life.

e. Data collected in pragmatic trials

Pragmatic trials aim to provide evidence regarding the relative effectiveness of treatment options in a setting that tries to mimic real life. Yet, one should be mindful that prospectively collected data are not truly RWD (hence not at the far right side of the above continuum) because the actual fact of participating in such a prospective study may already introduce a bias (the Hawthorne effect), and lead to behaviour and outcomes that differ from the real world setting.

f. Data from social media, blogs, chat rooms, patient communities

Large quantities of unorganised data can nowadays be analysed in order to identify patterns that explain the course of diseases and treatments. Although the usefulness of such data is still a matter of debate they represent a trend in the generation of data that cannot be ignored.

3. Why real world data?

The collection of RWD has many benefits that cannot be sufficiently offered by conventional RCTs.

i. Outcomes can be collected in a variety of typical practice settings.

ii. It better helps identify patients that benefit most from treatments as well as those who will not benefit.

iii. It provides more opportunities to compare multiple alternative interventions.

iv. It can lead to estimates of longer-term clinical benefits and harm.

v. A broader range and variety of clinical and economic outcomes can be collected.

vi. It offers a better opportunity to evaluate potential rare patient characteristics and rare harm from an intervention.

vii. It allows collection of more representative data on resource use and costs in daily practice.

viii. It provides information on how a treatment is dosed and applied in daily practice.

ix. It provides information on levels of adherence (consisting of persistence and compliance) to therapy.

The above list may suggest that RWD should be considered as an alternative to conventional RCT. This is however not the case and should not be the ambition of RWD. It is more correct to state that the availability of RWD before, in parallel to, and after RCTs broadens the options to collect relevant data and has different purposes. RWD should not be seen as being in competition with RCT for the hierarchy of evidence, but as complementary.[vi] It should be recognized moreover that RWD hold some drawbacks and limitations, such as variable data quality, incomplete data, difficulties in obtaining comparative evidence and absence of clear standards for analysis. On the other hand, it should also be recognized that in exceptional cases it is simply not possible to conduct RCTs. For instance, in the field of personalised medicine, many strata of patients with a given disease exist and sometimes it is not feasible to organise RCTs in all these strata, especially given the need for long-term follow-up in many disease areas. Typically, other relevant examples are in the fields of geriatrics and paediatrics where it would be cumbersome to repeat RCTs in these patient segments.

4. Possible use of RWD throughout the lifecycle of a medicine

As outlined above, the use of RWD must be seen as complementary to the use of RCTs. RWD can indeed prove to be very useful during the different phases of a medicine's lifecycle, i.e. the development phase, the market access phase and the market usage phase. The challenge is to define when and how much data are needed and for which purpose. Outlined below are examples of typical applications of RWD.

a. During the development phase, RWD can help:

   i. to better characterize diseases, patient populations, and to understand current needs; for instance, RWD can show how many patients with a given disease are insufficiently controlled or whose treatment is inadequate and what their characteristics are.

   ii. to better identify patients for participation in research programmes, to speed up the recruitment process and to make the recruitment more efficient. For instance, well managed databases based on EPRs allow queries leading to fast identification of patients meeting the recruitment criteria of an RCT.

   iii. as input to make the design of RCTs more 'pragmatic' (i.e. moving slightly more to the right of the continuum thus better reflecting real life). For instance, claims databases can show what the routinely number of follow up visits and investigations are in daily practice and this practice can be mimicked in the pragmatic trial.

   iv. As regards personalised medicine, RWD containing genetic and biomarker information can permit a swifter, efficient analytical and clinical validation of biomarkers and change the architecture of clinical development programmes (from 1 protocol for 1 population with 1 drug to multiple combinations).

It is clear from the four above-mentioned examples that RWD *can enable more effective and efficient research and development processes.*

b. During the market access phase, RWD allow:

   i. a better understanding of current patient management, and modalities of the current standard of care, for the sake of comparison with the new medicine; typically, in health economic evaluations the new medicine is compared to the standard of care. It is thereby indispensable that the standard of care is described as accurately as possible.

   ii. a better understanding of the real life outcomes related to the current standard of care, such as the number of complications, adverse events, disease progression, resource use and costs.

The above two examples show that during the market access phase, RWD allow a *better understanding of the current situation*. This, in combination with data from comparative trials, leads to better estimates of the benefits of new medicines.

c. Finally, during the market usage phase, RWD allow:

   i. the provision of evidence on the real world usage of innovative medicines: for example, in which patients and according to which modalities (dosage, duration etc.)

   ii. the assessment of the outcomes of innovative medicines in real life practice, which may serve as input in outcomes-based Managed Entry Agreements (MEAs). In such agreements, a point of verification can be set, which allows to assess whether the predicted benefits of a medicine can be confirmed.

   iii. the creation of clinical decision support systems

The first two examples are typical applications of RWD during the market usage phase whereby initial estimates on the use and benefits of a new treatment need to be confirmed. Combined with the initial data from RCTs, RWD can offer the necessary information post launch to confirm (or not) the benefits of a medicine. One can even envisage a dynamic reassessment of the value of a medicine, based on RWD.

Finally, (the 3rd example), RWD can be used to build and improve clinical decision support systems that allow physicians and hospitals to make better, and correct, use of medicines, thereby also allowing a more stratified approach. RWD help to provide insights as to why patients respond or not to treatments and how different patients should be managed. In other words, RWD can help better understand healthcare outcomes and how to get better value for money. As such, RWD can also contribute to health systems performance.

5. Real life challenges related to the use of real world data.

The above shows many applications of RWD. Yet, the combination of overly optimistic expectations about their use and the presence of immature data may lead to a situation of

"surreal world data": RWD used for the wrong purpose, at the wrong time, with poor quality and with wrong analyses. It should be recognised that there are many current issues related to the collection, interpretation and application of RWD. These issues are related to:

a. the quality and integrity of the data. For example, quality issues can be related to missing data, wrong inputs, and issues with definitions. Data integrity refers to maintaining and assuring the accuracy and consistency of data;

b. the lack of agreement between different involved parties regarding what data are needed, at which point in time, and for which purpose;

c. the difference in structure, setup and content of different databases, leading to significant challenges in conducting pan European use of RWD.

d. the differences in medical management between and within Member States, leading to wide heterogeneity in the data.

e. the scientific rigour of the data management and analysis. Without a good data management and analysis plan, the data risks being incorrectly analysed, which results in poor evidence. Even with good analysis, it should be recognised that real world data (RWD) cannot achieve the same internal validity as that of RCTs. For instance, there is evidence from literature that RWD may wrongly estimate comparative effectiveness;

f. lack of access to, and availability of, data. Excellent data sources may exist in a given jurisdiction, but due to rules and restrictions regarding data sharing, the data may be difficult, and sometimes even impossible, to access;

g. the time to collect, analyse and report the data may take too long so that the evidence based on RWD comes in too late to serve their purpose;

h. the cost of collecting the data. In some circumstances, access is possible but at a high cost for the party that needs the data for a given purpose;

i. issues with privacy and data security;

j. lack of governance: for example, no, or poor standards for collaboration, and lack of incentives for data sharing.

6. Principles for good RWD practice in Europe

Given the above, it is clear that RWD are not a panacea and do not replace conventional RCT data. RWD provide a source of data that is complementary to data from RCTs and can be applied throughout the lifecycle of a medicine, in different formats, with different content and for a wide spectrum of purposes, not only directly related to the development and assessment of innovations but also to improve health systems. In order to achieve these purposes, some principles of good practice are proposed.

Principle 1: Clarify the purpose of RWD use

The purpose of RWD use should be clear from the start to all involved parties. The approach should not be 'we have data, what should we do with it?', but 'we have well-defined questions, which data do we need and by when to answer those questions?' It should be recognised that certain designs are not possible for certain purposes.

Principle 2: Make a trade-off

A trade-off always needs to be made regarding the amount of data needed, the purpose, and the cost. Indeed, sometimes the use of RWD is not effective nor cost-effective. Duplication of efforts should be avoided as much as possible.

Principle 3: Alignment about the relevance of RWD

Building further on the first principle, there needs to be alignment, at the least between researchers, regulators, payers, and HTA bodies on what is relevant data at which point in time and for which purpose.

Principle 4: Create a joint activity around RWD

All parties involved (researchers, regulators, payers, HTA bodies, industry, patients etc.) should engage in a dialogue regarding the timely collection, correct analysis and adequate reporting of RWD. The use, assessment, and interpretation of the data should be a joint activity. Capacity building is necessary to achieve this.

Principle 5: Invest in data quality, access and analytical excellence.

RWD can only become real world *evidence* if the data are of high quality, if they can be accessed and are analysed correctly to answer the well-formulated questions. European Reference Networks should play an increasing role to achieve this.

Principle 6: Develop standards for RWD

There is a need for common standards in the EU regarding the systems for data input and data organisation of registries and routinely collected health data. This requires a minimum set of common data, and interoperability of different applied systems, to be facilitated by DG Connect. We refer also to the EMA initiative on registries. [vii]

Principle 7: Facilitate access to RWD

In the interest of patients, there is a need for clear and consistent rules about access and cost of access to routinely collected health data. Public Private Partnerships should be encouraged for this purpose. It should thereby be made clear from the start who should pay for the systems to collect RWD, especially considering long-term infrastructure.

Principle 8: Guarantee data security and privacy

Also in the interest of patients, there must be clear communication about the measures that policy makers take to guarantee data security and privacy. In this regard, there is a need for trust-building institutions. More research is needed to better understand the trade-offs that citizens want to make between keeping their privacy on the one hand, and on the other hand sharing personal data that can benefit themselves as well as others. The format and content of a so-called 'one time consent' needs to be worked out to enable people to indicate what they want to share, and when.

Principle 9: RWD as catalyst for Health Systems Improvement

It should be recognised by all stakeholders that RWD provide a unique opportunity to assess health systems performance, by identifying bad or poor practices and allowing to take and follow-up on measures for improving the efficiency and equity of systems. As such, RWD can contribute to the continuous learning of health care systems and are therefore of importance to guide health systems reform.

We hope that these principles, as well as the necessary actions that should follow, will improve the use of RWD throughout the lifecycle of innovative medicines, help to get valuable and affordable innovations faster to all EU patients and improve health system performance overall.

This paper has been commissioned by the Belgian Institut National d'Assurance Maladie-Invalidite (INAMI)/ RIZIV) and authored by Professor Lieven Annemans (University of Ghent). The following individuals and stakeholder representatives from patient groups, healthcare professional organisations, government bodies and industry have participated in a series of roundtable discussions, facilitated and chaired by FIPRA and sponsored by the European Federation of Pharmaceutical Associations (EFPIA), Roche, Amgen, and Pfizer, which have contributed to the development of the paper.
European Patient's Forum (EPF)
EURORDIS, Rare Diseases Europe
European Organisation for Research and Treatment of Cancer (EORTC)
Dr Thomas Mueller, Head of Pharmacueticals & Dr Florian Janschak, Scientfic Officer, Gemeinsamer Bundesausschuss (G-BA), Germany
Ministry of Health, Luxembourg
National Institute of Clinical Excellent (NICE), UK
Zorginstitute Nederlands (ZIN), Netherlands
Pascal Garel, Chief Executive, European Hospital and Healthcare Federation (HOPE)
Luca Pani, CHMP Member, European Medicines Agency (EMA), former Director General, AIFA

[i] http://www.ema.europa.eu/ema/index.jsp?curl=pages/regulation/general/general_content_000601.jsp
[ii] http://www.ema.europa.eu/ema/index.jsp?curl=pages/regulation/general/general_content_000660.jsp&mid=WC0b01ac058096f643
[iii] Council Conclusions on Personalised Medicine for Patients, December 2015
[iv] https://www.imi-getreal.eu/
[v] RAND Europe http://www.rand.org/randeurope/research/projects/real-world-data-policy-landscape.html
[vi] Harbour R and Miller J. A new system for grading recommendations in evidence based guidelines. BMJ. 2001 Aug 11; 323(7308): 334–336
[vii] http://www.ema.europa.eu/ema/index.jsp?curl=pages/regulation/general/general_content_000658.jsp&mid=WC0b01ac0580961211