Get Real

Real-Life Data in Drug Development

GetReal - Project No. 115546

WP1: Deliverable D1.3

Glossary of Definitions of Common Terms (Including Comments & Replies from Consultation Rounds)

Lead Organisation and Investigators:

Zorginstituut Nederland (Wim Goettsch, Amr Makady)

efpia

innovative medicines initiative

The work leading to these results has received support from the Innovative Medicines Initiative Joint Undertaking under grant agreement n° [115546], resources of which are composed of financial contribution from the European Union's Seventh Framework Programme (FP7/2007- 2013) and EFPIA companies' in kind contribution.
www.imi-getreal.eu

Table of Contents

1. Executive Summary

The IMI GetReal consortium has drawn up definitions of key terms, both for the purpose of GetReal, and also with the aim of providing clarity to external stakeholders around these terms. In order to explore the origin and context of some of these terms, the glossary is preceded by an overview of real world data and the origin of the key concepts associated with the term. In addition, this document explores the definitions of both 'efficacy' and 'effectiveness' and attempts to define how efficacy studies can be distinguished from effectiveness (pragmatic) studies.

2. Defining Real World Data, and its origins

Healthcare funding decisions increasingly rely on evidence beyond that which is collected in randomised controlled trials (RCTs), which are required by regulatory authorities for marketing authorisation. There are many reasons for this, including the need to understand health outcomes in routine clinical practice, which it is recognised may differ from those observed under the idealised conditions of a RCTs (for more on this, refer to the section '3. Effectiveness versus Efficacy'). This has long been referred to as the collection of real world data, without consensus on the definition of real world data, or what constitutes a real world study. The evolution of this terminology is probably due to the continual and rapid development of IT infrastructures and capabilities; increasing availability of databases and linked datasets; and methodological refinements in data collection. Fifty years ago researchers used paper patient records to establish a link between smoking and lung cancer. The 1980s brought access to real world data on medicine use at the patient level when administrative information about medicine use was first stored at a significant level.[1] In addition, there is no single

[1] van Staa and Klungel 2013. Real-life data and learning from practice to advance innovation. Background paper commissioned by the WHO (http://www.who.int/medicines/areas/priority_medicines/BP8_4Data.pdf)

legal instrument or guidance specific to real world data collection (in the UK at least; recognised by the ABPI), and levels of support for and practical use of real world evidence varies across markets.[2,3]

Since the evolution of an ISPOR task force dedicated to assisting healthcare decision-makers in dealing with real world data, there appears to be greater consensus and consistency in the terminology used by authoritative bodies about real world data.

The ISPOR task force offers the broad definition of "everything that is not collected in conventional randomised controlled trials".[4] However some comments received during consultation of the report suggested the definition be restricted to primary data collected at patient level by methods that minimise the imposition of artificial constraints (i.e. excludes secondary analyses like systematic reviews and decision analytic models). The ABPI guidance on real world data specify the lack of an intervention: "data obtained by any non-interventional methodology that describes what is happening in normal clinical practice". Or more simply put: "data which describes what is really happening in everyday normal clinical health care practice".[5] The European Forum "Relative Effectiveness" Working group (part of the European Commission) offer a similar explanation of real world data as "a measure in understanding health care data collected under real life practice circumstances".[6] To differentiate between real world 'data' and 'evidence', the ISPOR task force explain that "data conjures the idea of simple factual information, whereas evidence connotes the organization of the information to inform a conclusion or judgment. Evidence is generated according to a research plan and interpreted accordingly, whereas data is but one component of the research plan. Evidence is shaped, while data simply are raw materials and alone are non-informative."

The commonly discussed outcomes investigated in real world evidence are clinical, economic, and HRQoL/PRO. The ABPI guidance additionally refers explicitly to treatment pathways, service models and patient preference/experience/compliance.

As the definitions of real world data are broad, and may be open to interpretation, it might be more relevant to investigate thoughts on how to collect these data, i.e. understand definitions of real world studies. The ISPOR task force defined 6 sources of real world data:

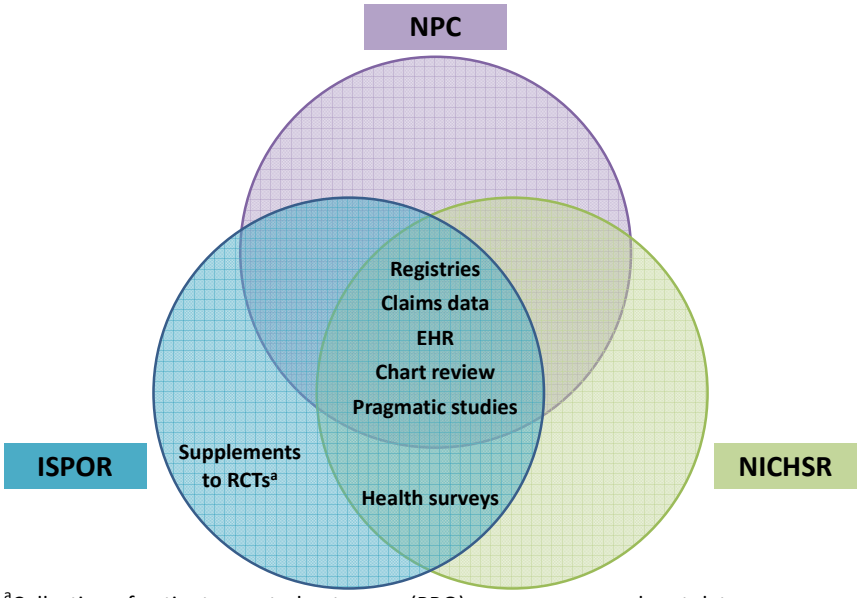
---

[2] Quintiles report 2013: Real World Evidence and Implications for Emerging Technologies (http://www.namcp.org/journals/spring13/Faulkner%20Emerging%20Technologies.pdf)
[3] IMS Consulting report 2013: International comparisons of the impact of Real World Evidence on creating value from medicines (http://www.imsconsultinggroup.com/deployedfiles/consulting/Global/Content/Our%20Latest%20Thinking/Static%20Files/rwe_market_impact_on_medicines.pdf)
[4] Garrison et al. ISPOR task force report (http://www.ispor.org/workpaper/RWD_TF/RWTFManuscript.pdf)
[5] ABPI Guidance 2011: Demonstrating Value with Real World Data (http://www.abpi.org.uk/our-work/library/guidelines/Documents/2011-06-13%20ABPI%20guidance%20-%20Demonstrating%20value%20with%20real%20world%20data.pdf)
[6] http://www.ispor.org/news/articles/oct07/rld.asp

1. **Supplements to traditional registration RCTs:** to collect PRO, HRQoL, resource use and cost data.
2. **Pragmatic clinical trials (also known as large simple trials or practical clinical trials):** involve prospective, randomised assignment but are aimed at larger more diverse real world population. These trials are by design larger than conventional RCTs.
3. **Registry studies:** prospective, observational cohort studies of patients who have a particular disease and/or are receiving a particular treatment or intervention. By definition registries are assembled prospectively, but data on exposure, treatments and events occurring before enrolment may be assembled retrospectively at baseline.
4. **Claims databases/administrative data:** Typically retrospective or real time, if possible. Lend themselves to retrospective longitudinal and cross-sectional analyses of outcomes at patient, group, or population levels.
5. **Health surveys:** to collect descriptions of health status and wellbeing, health care utilisation, treatment patterns, and health care expenditures from patients, providers, or individuals in the general population. Health surveys typically collect information on representative individuals in the target population (patients, physicians or general population).
6. **Electronic health records (EHR) and medical chart reviews:** such as the UK General Practice Research Database (GPRD). These contain more detailed, longitudinal information including disease-specific symptoms at the patient level.

These 6 data sources are echoed in publications from Quintiles and IMS Consulting. Very similar lists are suggested by the National Pharmaceutical Council (NPC) and National Information Center on Health Services Research and Health Care Technology (NICHSR) in the USA (Figure 1).

Figure 1 Venn diagram to illustrate the areas of agreement in the definition of real world studies across national and international organisations (ISPOR, NPC and NICHSR)

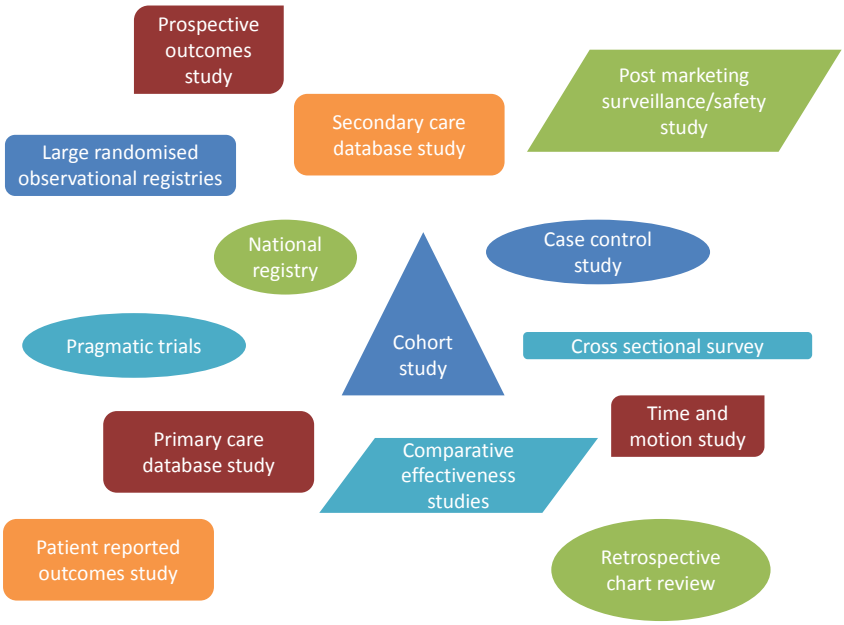


[a]Collection of patient reported outcomes (PRO), resource use and cost data

EHR, Electronic Heath Records; ISPOR, International Society For Pharmacoeconomics and Outcomes Research; NICHSR, National Information Center on Health Services Research and Health Care Technology (USA); NPC, National Pharmaceutical Council (USA); RCTs, Randomised Controlled Trials

The ABPI guidance on real world data provides a much longer list of observational data sources, but these can essentially all be grouped into 1 of the 6 study types provided by the ISPOR task force. A contradiction in terms exists in the ABPI guidance, in the suggestion of “large randomised observational registries”. By definition (according to ISPOR), a registry does not involve randomisation. This goes to show that some confusion still exists in the industry.

Figure 2 ABPI examples of real world studies



ABPI Guidance 2011: Demonstrating Value with Real World Data (http://www.abpi.org.uk/our-work/library/guidelines/Documents/2011-06-13%20ABPI%20guidance%20-%20Demonstrating%20value%20with%20real%20world%20data.pdf)

Very few bodies appear to employ a narrower definition of real world studies than that from the ISPOR task force. The European Commission include the concept of no randomisation in their definition; although this is in reference to observational research, as opposed to real world data specifically (these terms are often used interchangeably). By deduction, this would exclude pragmatic studies from their definition. ISPOR's definition of observational research is more vague on this point, stating that care is not *typically* a result of randomisation (or other forms of patient assignment), presumably to allow the inclusion of pragmatic studies. The classification of pragmatic studies as a method of real world data collection could potentially be a point of contention. The ISPOR task force acknowledge that whether they are strictly real world studies is open to debate. However, many national and international bodies group include them in their definition of real world studies:

- National Pharmaceutical Council (NPC), USA
- National Information Center on Health Services Research and Health Care Technology (NICHSR[7]), USA

[7] The NICHSR are part of the US National Institutes of Health (NIH) and were set up at the National Library of Medicine to improve "..the collection, storage, analysis, retrieval, and dissemination of information on health services research, clinical practice guidelines, and on health care technology, including the assessment of such technology."

- The NICHSR appear to differentiate between large simple trials and pragmatic clinical trials, unlike other organisations which used the terms interchangeably. But they explain that some large simple trials are also pragmatic trials.
- Large simple trials: retain the methodological strengths of prospective, randomised design, but use large numbers of patients, more flexible patient entry criteria and multiple study sites to generate effectiveness data and improve external validity. Fewer types of data may be collected for each patient, easing participation by patients and clinicians. Prominent examples of include the GISSI trials of thrombolytic treatment of acute myocardial infarction (AMI) (Maggioni 1990), the ISIS trials of alternative therapies for suspected AMI (Fourth International Study of Infarct Survival 1991), and the CATIE trial of therapies for schizophrenia (Stroup 2003).
- Pragmatic trials are a related group of trial designs whose main attributes include: comparison of clinically relevant alternative interventions, a diverse population of study participants, participants recruited from heterogeneous practice settings, and data collection on a broad range of health outcomes.

- Patient-centred outcomes research institute (PCORI), USA
  - PCORI implies that pragmatic studies are categorised as real world studies through its announcement of a new research funding initiative for pragmatic trials ("More Support for Bigger, Longer Real-World Trials").[8]
- Medicines and Healthcare products Regulatory Agency (MHRA), UK
  - It could be inferred from the MHRA business strategies that pragmatic studies (utilising EHR) are considered to produce real world data.[9]
- Farr Institute[10], UK
  - It could be inferred from the programme for their industry forum meeting (in collaboration with the ABPI) that pragmatic trials are included in their consideration of real world data collection.[11]

To revisit the concept of observational research - it seems the main defining feature of observational research, common across publications, is that care is not dictated or mandated. That is, the

[8] http://www.pcori.org/blog/introducing-new-pcori-research-funding-initiative-large-pragmatic-clinical-trials
[9] http://www.mhra.gov.uk/home/groups/comms-ic/documents/publication/con261796.pdf
[10] The Farr Institute was set up by a 10-founder consortium as well as funds from the Medical Research Council (MRC).
[11] http://www.farrinstitute.org/events/49/2014-12-16/farr-institute-abpi-industry-forum-2014.html

investigator does not interfere with choice of the prescribed health intervention such that interventions are prescribed in the usual manner in accordance with the terms of the marketing authorisation.[12,13,14,15] But it is not that simple, as there are inconsistencies internally regarding the definition of an intervention. Depending on local regulations, for example, blood tests and patient questionnaires may be considered interventional in some countries (particularly in Europe) but not in others.

3. Effectiveness versus Efficacy

Distinguishing efficacy from effectiveness and emphasising its importance to decision making dates back to at least 1978 (Office of Technology Assessment 1978), but confusion still exists.

It is widely thought that efficacy is the extent to which a healthcare intervention produces a therapeutic effect *as compared to a placebo* under ideal conditions (i.e. the highly-controlled conditions of *RCTs*).[16,17,18,19,20,21] Because RCTs use randomisation and other features to minimise bias, they can prove a causal relationship between an intervention and an outcome. On the other hand, the effectiveness of an intervention refers to its health benefits in *routine clinical practice* (that is, in *real world studies*; according to the ISPOR task force, and online glossaries published by the INAHTA and Cochrane Collaboration) where multiple variables that might influence patient outcomes (such as concomitant medication and comorbidities) are introduced. Effectiveness research generally involves at least two active comparators (i.e. compares an intervention to standard practice, rather than placebo).

Disagreement about the terminology across Europe was highlighted at the High Level Pharmaceutical Forum.[22] It was concluded that there is no clear consensus as to whether clinical

[12] ISPOR Taxonomy of Patient Registries 2013 (ISBN: 978-0-9743289-4-2).
[13] HTA glossary (http://htaglossary.net/HomePage)
[14] Cochrane collaboration glossary (http://www.cochrane.org/glossary/)
[15] European Commission (2001). DIRECTIVE 2001/20/EC of the European Parliament and of the Council. Official Journal of the European Communities. Cited in ABPI Guidance.
[16] ISPOR Taxonomy of Patient Registries 2013 (ISBN: 978-0-9743289-4-2).
[17] Cochrane collaboration glossary (http://www.cochrane.org/glossary/)
[18] HTA glossary (http://htaglossary.net/HomePage)
[19] Holve, E. and P. Pittman, A First Look at the Volume and Cost of Comparative Effectiveness Research in the United States. AcademyHealth. June 2009
[20] European Monitoring Centre for Drugs and Drug Addiction (http://www.emcdda.europa.eu/publications/glossary#cochraneCollaboration)
[21] The International Working Group for HTA Advancement: Milbank Q. Jun 2010; 88(2): 256–276 (http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2980346/)
[22] High Level Pharma Forum: http://eunethta.fedimbo.belgium.be/sites/5026.fedimbo.belgium.be/files/Final%20version%20of%20Backgroun%20Review%20on%20Relative%20Effectiveness%20Assessment%2Bappendix.pdf

trials yield efficacy or effectiveness information. While some EU Member States use effectiveness to describe what is actually happening in real life, others use it exclusively to "describe clinical trials that are as far as possible to the effectiveness side of the spectrum".

It has been suggested that any confusion that exists might be attributed to the interchangeable use of the terms efficacy and effectiveness in the FDA's legislation and regulations; effectiveness is used when efficacy is intended. This misapplication of terms been used by other organisations as well. For example, the Drug Effectiveness Review Project (DERP)'s stated mission is to "obtain the best available evidence on effectiveness and safety comparisons between drugs in the same class, and to apply the information to public policy and related activities", yet DERP relies exclusively on evaluations based on RCTs. Similarly, the Cochrane Collaboration describes its reviews as exploring "the evidence for and against the effectiveness and appropriateness of treatments … in specific circumstances", however they also demonstrate an almost complete reliance on RCT literature.[23]

RCTs that assess effectiveness are sometimes called pragmatic or management trials, which denotes a grey area (as discussed earlier, in the discussion of the classification of pragmatic trials). To allow for more fluid definitions in the face of the evolving methodology of clinical trials, some publications refer to a spectrum of trial designs and conduct.[24] Some groups have attempted to develop a tool to distinguish efficacy trials from effectiveness (pragmatic) trials (

[23] The International Working Group for HTA Advancement: Milbank Q. Jun 2010; 88(2): 256–276 (http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2980346/)
[24] The International Working Group for HTA Advancement: Milbank Q. Jun 2010; 88(2): 256–276 (http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2980346/)

Figure 3 and

Figure 4).

Figure 3 Tool to distinguish efficacy from effectiveness studies (Gartlenarner et al., prepared for the Agency for Healthcare Research and Quality [AHRQ][25])

1) Populations in primary care 2) Less stringent eligibility criteria 3) Health Outcomes as principal outcomes (e.g., functional capacity, quality of life, mortality; as opposed to objective or subjective outcomes) 4) Long study duration 5) Assessment of AEs 6) Adequate sample size to assess a minimally important difference from a patient perspective 7) Intention to treat analysis * The tool uses yes/no answers; where yes denotes effectiveness studies (despite acknowledgment from the authors that the 2 study types exist on a continuum).

Figure 4 The 'PRECIS' tool to distinguish efficacy from effectiveness studies (Thorpe et al.[26])

A pragmatic trial across the 10 domains of the PRECIS tool would fulfil the following criteria: • There are no inclusion or exclusion criteria • Practitioners are not constricted by guidelines on how apply the experimental intervention • The experimental intervention is applied by all practitioners, thus covering the full spectrum of clinical settings • The best alternative treatments are used for comparison with no restrictions on their application • The comparative treatment is applied by all practitioners, covering the full spectrum of clinical settings • No formal follow-up sections • The primary outcome is a clinical meaningful one that does not require extensive training to assess • There are no plans to improve or alter compliance for the experimental or the comparative treatment • No special strategy to motivate practitioner's adherence to the trial's protocol • The analysis includes all participants in an intention-to-treat fashion

[25] Gartleharner G. and Hansen RA. A Simple and Valid Tool for Distinguishing Efficacy from Effectiveness Studies. Journal of Clinical Epidemiology 2006. 59(10):1040-8.
[26] Thorpe KE., Zwarenstein M., Oxman AD., et al. A pragmatic-explanatory continuum indicator summary (PRECIS): a tool to help trial designers. J Clin Epidemiol. 2009;62:464–475

4. IMI-GetReal Glossary of Key Terms

The table below contains definitions of key terms relevant to the GetReal consortium. The GetReal glossary working group acknowledges there is a debate around many of these terms, and in some cases no clear consensus has been reached in the international community. However, for the purpose of GetReal these are the definitions proposed for the project.

Please note: the definitions presented here are the result of several rounds of consultation. More specifically, definitions in previous versions of the glossary have been subjected to: an internal GetReal consultation round for consortium members in 2015, a public consultation round for external stakeholders in 2015, as well as second internal GetReal consultation round in 2016. For readers interested in the comments received during these rounds as well as authors' replies, please see section 5 below.

Table 1 Terms of key relevance to GetReal.

Term	Definition	References
Adaptive clinical trial	A clinical trial that evaluates patients' response and reaction to an intervention at pre-determined intervals, beginning with evaluation at an early stage in the clinical trials and subsequently modifying the trial according to findings generated by interim analyses. By means of an adaptive design, researchers thus have the opportunity to modify the trial procedures at different stages on the basis of analysing interim data from study subjects. Such modifications may include, but are not limited to: selected drug dosages, sample size, and patient selection criteria. (see also: "clinical trial") (Adapted from CHMP 2007 and FDA, 2010).	Committee for Medicinal Products for Human Use. (2007). Reflection paper on methodological issues in confirmatory clinical trials planned with an adaptive design. London: EMEA; FDA (2010) *Guidance for industry. Adaptive design clinical trials for drugs and biologics.* Obtained on 5 March 2014. URL: http://www.fda.gov/downloads/Drugs/.../Guidances/ucm201790.pdf
Administrative claims data	Data arising from a person's use of the healthcare system and reimbursement of healthcare providers for that care. (Strom 2005)	Strom, 2005; FDA GUIDANCE ON PHARMACOEPIDEMIOLOGIC SAFETY STUDIES (MAY 2013)
Adverse event	Any untoward medical occurrence in a patient or clinical investigation subject administered a health intervention, and which does not necessarily have a causal relationship with this treatment. (ICH, 1994)	ICH (1994). Clinical Safety Data Management: defnitions and standards for expedited reporting. http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E2A/Step4/E2A_Guideline.pdf
Aggregate (Study) Data	Summary data of the results of a study (e.g. on subjects in a trial), as opposed to Individual Patient Data which represents the raw data of each study subject. Summary data is derived from individual patient data using a (statistical) method of aggregation. (Adapted from Lyman, 2005)	Lyman, G. H., & Kuderer, N. M. (2005). The strengths and limitations of meta-analyses based on aggregate data. *BMC medical research methodology*, *5*(1), 14.
Aggregate Data Drug Information System (ADDIS)	Software designed to aid evidence-based decision-making in the healthcare setting. This software does so by offering on-demand features such as: pair-wise-network analysis, meta-analysis, and multiple-criteria decision analysis. (van Valkenhoef, 2013)	Gert van Valkenhoef, Tommi Tervonen, Tijs Zwinkels, Bert de Brock, Hans Hillege, ADDIS: A decision support system for evidence-based medicine, Decision Support Systems, Volume 55, Issue 2, May 2013, Pages 459-475, ISSN 0167-9236, http://dx.doi.org/10.1016/j.dss.2012.10.005.
Alternative study design	This refers to all scientific studies investigating health interventions whose design does not follow the design of a randomised controlled clinical trial (see also:"randomised controlled clinical trial (RCT)") (Freemantle, 2010)	Nick Freemantle, Thomas Strack, Real-world effectiveness of new medicines should be evaluated by appropriately designed clinical trials, Journal of Clinical Epidemiology, Volume 63, Issue 10, October 2010, Pages 1053-1058, ISSN 0895-4356, http://dx.doi.org/10.1016/j.jclinepi.2009.07.013.
Bayesian methods	Statistical methods based upon Bayes' Theorem, which shows how to update prior knowledge in the light of new data (i.e. posterior probability $\propto$ likelihood x prior probability). Prior knowledge is defined in terms of probability distributions and can be based on subjective opinion, or on objective evidence, such as the results of previous research, or both, and is explicitly included in consequent calculations. Statistical inference is then based on suitable summaries from the posterior probability	Rothman, K.J., Greenland S., Lash T.L. (2008) Modern Epidemiology. Lippincott Williams & Wilkins. ISBN: 978-0-7817-5564-1; HTA Glossary - Bayesian Methods URL: http://htaglossary.net/Bayesian+analysis&highlight=real-world%20evidence

	distribution. (Adapted from Rothman, 2008 and HTA glossary)	
Bias	Systematic (non-random) errors in values of parameters that are the object of study. Errors in estimations can result from, for example, improper study design or analysis, and implicitly affect the internal validity and generalisaibility of study results. There are three main categories of bias: selection bias, information bias and confounding bias. (Adapted from Rothman, 2008 and Delgado-Rodriguez, 2004)	Rothman, K.J., Greenland S., Lash T.L. (2008) Modern Epidemiology. Lippincott Williams & Wilkins. ISBN: 978-0-7817-5564-1; Delgado-Rodríguez, M., & Llorca, J. (2004). Bias. Journal of epidemiology and community health, 58(8), 635-641.
Bridging study	A study, supplemental to a randomised controlled clinical trial, designed to provide additional clinical data on the safety, efficacy, dose and regimen, thus allowing for the extrapolation of external trial data to a new subjects population with (possibly) different population characteristics. An ethnicity briding study is one example of such a supplemental study.(Adapted from ICH, 1998)	ICH (1998) ICH Topic E5 (R1) Ethnic Factors in the Acceptability of Foreign Clinical Data. http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E5_R1/Step4/E5_R1__Guideline.pdf
Case-control study	A study in which the exposure of a group that have experienced a certain outcome (cases) is compared to the exposure of a group who have not (yet) experienced the outcome (controls). Case-control studies can be retrospective or prospective. (Adapted from Rothman, 2008 and Grobbee & Hoes, 2015)	Rothman, K.J. Greenland, S. Lash, T.L. (2008). Modern Epidemiology. Lippincott Williams & Wilkins. 88.; Grobbee, D. E., & Hoes, A. W. (2014). Clinical epidemiology. Jones & Bartlett Publishers.
Clinical endpoint/ outcome	An aspect of a subject's clinical or health status that is measured to assess the benefit or harm of an intervention. A clinical endpoint describes a valid measure of clinical benefit due to intervention: the impact of the intervention on how a subject feels, functions and survives. It is clinically relevant, sensitive (responsive to change) and is both accepted and used by physicians and patients. Clinical endpoints may be a clinical event (e.g. mortality,) a composite of several events, a measure of clinical status (e.g. blood pressure), or health related quality of life (HRQoL). (Adapted from EunetHTA, 2013)	EUnetHTA (2013) Methodology Guidelines. Endpoints used for relative effectiveness assessment of pharmaceuticals: Clinical Endpoints. http://www.eunethta.eu/sites/5026.fedimbo.belgium.be/files/Clinical%20endpoints.pdf
Clinical Guideline/ Medical Guideline	A document produced with the intention to guide decisions and inform criteria with regards to disease prevention, disease diagnosis, disease management, and treatment in the routine clinical setting. (Adapted from Council of Europe, 2001)	Council of Europe (2001). Developing a methodology for drawing up guidelines on best medical practices. http://www.leitlinien.de/mdb/edocs/pdf/literatur/coe-rec-2001-13.pdf
Clinical significance	The practical importance of and benefit from a treatment. It describes whether and to which extent the intervention has a real genuine, palpable, noticeable effect on daily life. Clinical significance is usually informed by effect size and statistical significance. (See also "statistical significance") (Adapted from Kazdin, 1999 and Redmond, 2001).	Kazdin, A.E., *The Meanings and Measurement of Clinical Significance*. Journal of Consulting and Clinical Consulting **67** (3): 332–9. doi:10.1037/0022-006x.67.3.332, 1999.

Clinical Trial	Any investigation in human subjects intended to discover or verify the effects of an intervention, such as clinical, pharmacological and/or other pharmacodynamic effects, and/or to identify any adverse reactions to an intervention. (Adapted from ICH, 2001)	Guideline, I. H. T. (2001). Guideline for good clinical practice. *J Postgrad Med*, *47*(1), 45-50.
Cohort study	A study in which a group of subjects, sampled from a certain source population, are classified according to their exposure or determinant status and followed over time to ascertain the occurrence of a certain outcome. (Adapted from Rothman, 2008)	Rothman, K.J. Greenland, S. Lash, T.L. (2008). Modern Epidemiology. Lippincott Williams & Wilkins. 88.
Comparative effectiveness research	The conduct and/or synthesis of research comparing different benefits and harms of alternative interventions and strategies to prevent, diagnose, treat, and monitor health conditions in routine clinical practice (i.e. the real world setting). Comparative effectiveness research includes both primary data collection and secondary analyses (such as systematic literature reviews, meta-analyses and economic evaluations). (IOM, 2009 & Sox, 2009)	The Institute of Medicine (2009) - Initial National Priorities for Comparative Effectiveness Research; http://www.iom.edu/reports/2009/comparativeeffectivenessresearchpriorities.aspx; Sox, H.C. and S. Greenfield, Comparative effectiveness research: a report from the Institute of Medicine. Ann Intern Med, 2009. 151(3): p. 203-5.
Comparator	Reference intervention to which safety, efficacy and/or effectiveness of a health intervention (e.g. pharmaceutical product) are compared. In the case of clinical trials for pharmaceutical products, comparators can comprise a placebo treatment (placebo-control trials), available standard of care, and/or a licensed medication (active-control trial). (Adapted from ICH, 2001)	Guideline, I. H. T. (2001). Guideline for good clinical practice. *J Postgrad Med*, *47*(1), 45-50.
Conditional marketing authorisation	A one-year marketing authorization within the European Union with annual review by the European Medicines Agency (EMA), and which applies in specific cases: (a) Seriously debilitating or life-threatening diseases; (b) Emergency threats determined by the WHO, or the EU Commission; (c) Orphan medicinal products. Accelerated drug approval is the near equivalent of conditional marketing authorisation in the USA. (Boon, 2011)	Boon, H., Conditional Marketing Authorisations in the European Union, in FDA ODAC meeting. 2011: Silver Spring, MD; Goozner, M., Accelerated drug approval: FDA may get tougher; companies cite hurdles. J Natl Cancer Inst, 2011. 103(6): p. 455-7
Confounder	An extraneous variable in a statistical model that correlates with both the dependent variable (e.g. outcome) and the independent variable (e.g. intervention). Lack of consideration of confounders in statistical analyses can lead to spurious statistical relationships between the dependent and independent variables. (Adapted from Greenland, 2001)	Greenland, S., & Morgenstern, H. (2001). Confounding in health research. Annual review of public health, 22(1), 189-212.
Confounding bias	Systematic error that occurs when the estimate of a measure of association between exposure (e.g. healthcare intervention) and outcome (e.g. health status) is distorted by the effect of one or several extraneous variables (confounding factor(s)) that are independently related to the exposure and outcome. (Adapted from Strom, 2006)	Strom B.L., Kimmel S.E. (2006). Textbook of pharmacoepidemiology. John Wiley & Sons, Ltd.

Covariate	A variable that may be predictive of the outcome under study. Alternative terms are *explanatory variable*, *independent variable*, or *predictor*. Depending on its causal impact on the outcome under study, a covariate may be of direct interest, or act as a confounder or effect modifier (Adapted from J. M. Last, 2001).	Last, J.M., (ed.). A Dictionary of Epidemiology [4th ed.]. Oxford UP. ISBN 0-19-514168-7, 2001.
Coverage decisions	Decisions taken by healthcare payers/ insurers to determine allocation of resources with regards to which health interventions to reimburse, and the extent of reimbursement associated with the interventions covered by the payment package. (Steiner, 1996)	Steiner C.A., Powe N.R., Anderson G.F., Das A (1996). The review process used by U.S. health care plans to evaluate new medical technology for coverage. Journal of General Internal Medicine; 11(5); pp 294-302.
Cross-design evidence synthesis	Evidence synthesis comprising the pooling and analysis of data from different studies. The studies from which data is combined can differ in relation to their design types, clinical setting, outcome measures, study interventions, study parameters, or patient population. (See also: "meta-analysis") (Athanasiou, 2011)	Athanasiou T., Darzi A. (2011). Evidence Synthesis in Healthcare: A practical Handbook for Clinicians. Springer, ISBN 978-0-85729-175-8, doi: 10,1007/978-0-85729-206-3
Cross-sectional study	A study in which one ascertains exposure status and outcome status for the observed population at one specific point in time. (Adapted from Rothman, 2008)	Rothman, K.J. Greenland, S. Lash, T.L. (2008). Modern Epidemiology. Lippincott Williams & Wilkins. 88.
Direct treatment comparisons	The comparison of the relative effect(s) of several interventions for a particular therapeutic indication in a trial setting. The basis for comparison can vary (e.g. clinical endpoints, adverse effect rates, or drug adherence rates) and the trial type may vary (e.g. randomised controlled clinical trial, pragmatic clinical trial or observational study). (Adapted from Hetland et al, 2009)	Hetland, M. L., Christensen, I. J., Tarp, U., Dreyer, L., Hansen, A., Hansen, I. T., ... & Østergaard, M. (2010). Direct comparison of treatment responses, remission rates, and drug adherence in patients with rheumatoid arthritis treated with adalimumab, etanercept, or infliximab: results from eight years of surveillance of clinical practice in the nationwide Danish DANBIO registry. *Arthritis & Rheumatism*, *62*(1), 22-32.
Discrete-event simulation model	The key components of a discrete-event simulation (DES) model are the units that are simulated, the type of events that can occur, and the rules that govern the progress from one to the next event. Units are typically individual patients, the events could for example be stroke or increase or decrease of blood pressure in cardiovascular disease. Individual level RCT data are used to specify the functional form of the attributes that characterize each unit. Such data also serve to set up the rules that govern the changes in the attributes. DES provides a framework for individual level stochastic simulation: first, a virtual patient is generated, then the time of occurrence and type of the first event are simulated, the event rates and probabilities are updated in the light of the event, the second event is simulated using the new rates and probabilities, and so on. This scheme is repeated until a pre-defined stopping rule is met (for example the virtual patient reaches the end of the follow-up period). Events thus affect the variables that characterize the patient and his or	Karnon J, Stahl J, Brennan A, Caro JJ, Mar J, Möller J; ISPOR-SMDM Modeling Good Research Practices Task Force. Modeling using discrete event simulation: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force--4. Value Health. 2012 Sep-Oct;15(6):821-7. doi: 10.1016/j.jval.2012.04.013. J. J. Caro, “Pharmacoeconomic analyses using discrete event simulation,” PharmacoEconomics, vol. 23, no. 4, pp. 323–332, 2005. J. Banks, J. S. Carson, and B. L. Nelson, Discrete-event system simulation. Prentice Hall, 1996.

	her disease state. These variables, in turn, influence the rates of future events. Through this process, the patient's full trajectory over the time span of the simulation is generated. This is repeated for a large number of patients. At the end, the outcome variables are summarized.	
Drivers of effectiveness	Non-drug variables (i.e. related to the healthcare system, the disease, the patient, or the actual use of drug) for which a variation in distribution between real-world settings, as opposed to the setting in RCTs, would impact the outcome of a treatment. In turn, the effectiveness of the drug can be understood as the result of the modification of the pharmacological efficacy by the distributions of these drivers of effectiveness. (See also "effectiveness" and "efficacy") (GetReal, 2016)	IMI-GetReal Glossary Workgroup, 2016
Drug utilisation	Defined by the World Health Organisation as the the marketing, distribution, prescription and use of drugs in a society, with special emphasis on the resulting medical, social, and economic consequences. This definition contains 2 aspects: the process of drug utilisation, that is the movement of drugs along the drug chain in society, and how drug utilisation relates to the effects of drug use (Baksaas, 1981; Lee & Bergman 1989) (See also "drug utilisation studies") (Sacristén, 1994)	Sacristén, J. A., & Soto, J. (1994). Drug utilisation studies as tools in health economics. *Pharmacoeconomics*, *5*(4), 299-312.
Drug utilisation studies	Research designed to investigate drug utilisation. (See also: "drug utilisation") (Adapted from Sacristén, 1994)	Sacristén, J. A., & Soto, J. (1994). Drug utilisation studies as tools in health economics. *Pharmacoeconomics*, *5*(4), 299-312.
Effect modification	Occurs when the magnitude of the effect of the primary exposure on an outcome (i.e., the association) differs depending on the level of a third variable. (Adapted from VanderWeele, 2009)	VanderWeele, T. J. (2009). On the distinction between interaction and effect modification. *Epidemiology*, *20*(6), 863-871.
Effectiveness	The extent to which an intervention does more good than harm when provided under the usual circumstances of health care practice. (See also "ideal vs.usual conditions") (High Level Pharmaceutical Forum, 2008)	High Level Pharmaceutical Forum 2005-2008. Final Conclusions and Recommendations of the High Level Pharmaceutical Forum. URL: http://ec.europa.eu/pharmaforum/docs/final_conclusions_en.pdf
Effectiveness studies	Clinical studies investigating health interventions whose design does not follow the design of a randomised controlled clinical trial and aims to reflect health intervention effectiveness in routine clinical practice. Effectiveness studies do not typically include randomisation of trial subjects, but there are exceptions (e.g. pragmatic clinical trials). For the purposes of GetReal, effectiveness studies include, but are not limited to, the following: pragmatic clinical trials, non-interventional/ observational studies, drug utilisation studies, post-authorisation efficacy/safety studies. (See also: "real-world studies", "drug utilisation study",	IMI-GetReal Glossary Workgroup, 2016

	"pragmatic clinical trial" and "non-interventional/ observational study") (IMI-GetReal, 2014)	
Efficacy	The extent to which a healthcare intervention does more good than harm as compared to another healthcare intervention under ideal conditions. (See also "ideal vs.usual conditions") (High Level Pharmaceutical Forum, 2008)	High Level Pharmaceutical Forum 2005-2008. Final Conclusions and Recommendations of the High Level Pharmaceutical Forum. URL: http://ec.europa.eu/pharmaforum/docs/final_conclusions_en.pdf Accessed 7 March 2014
Efficacy study	A study aiming to measure the effect of a drug under highly controlled conditions (i.e. in the setting of randomised controlled clinical trials). The study serves to prove the causal relationship between an intervention and an outcome, thus answering the question "can it work?" in an ideal world.	Luce, B.R., et al., *EBM, HTA, and CER: clearing the confusion.* Milbank Q, 2010. **88**(2): p. 256-76.
Efficacy-effectiveness gap	The observed discrepancy between effects of a health intervention in routine clinical practice as compared with the effects demonstrated in randomised controlled clinical trials. (Adapted from Eichler et al., 2011)	Eichler, H. G., Abadie, E., Breckenridge, A., Flamion, B., Gustafsson, L. L., Leufkens, H., ... & Bloechl-Daum, B. (2011). Bridging the efficacyâ€“effectiveness gap: a regulator's perspective on addressing variability of drug response. Nature Reviews Drug Discovery, 10(7), 495-506.
Electronic health/medical record (EHR/EMR)	An electronic record of health-related information on an individual that can be created, gathered, managed, and consulted by authorized clinicians and staff within one healthcare organization. Patient health-related information may include all of the key administrative clinical data relevant to that person's care under a particular provider, including demographics, progress notes, problems, medications, vital signs, past medical history, immunizations, laboratory data and radiology reports. (Adapted from FDA 2013 and IOM, 2013)	Centers for Medicare en Medicaid Services http://www.cms.gov/Medicare/E-Health/EHealthRecords/index.html?redirect=/ehealthrecords/; FDA (2013). Best Practices for Conducting and Reporting Pharmacoepidemiologic Safety Studies Using Electronic Healthcare Data.
Electronic healthcare data	An organized set of healthcare data or collection of files available by computer through electronic format. It is derived from a raw electronic healthcare database. Electronic healthcare data include administrative claims data and electronic medical record (EMR) data. (Adapted from Hartzema, 2008 and FDA, 2013)	Hartzema, Tilson, and Chan. Pharmacoepidemiology and Therapeutic Risk Management. Cincinnati: Harvey Whitney Books, 2008.; FDA (2013). Best Practices for Conducting and Reporting Pharmacoepidemiologic Safety Studies Using Electronic Healthcare Data.
(Clinical) Equipoise	An ethical criterion for recommending subject participation in clinical trials that states that a subject can only be referred to study participation if there is genuine collective professional uncertainty from the medical community regarding the best health intervention for that subject. (Adapted from Miller, 2003)	Miller, F. G., & Brody, H. (2003). A critique of clinical equipoise: therapeutic misconception in the ethics of clinical trials. *Hastings Center Report*, *33*(3), 19-28.
External validity/	Whether the results of a study can be reasonably applied to a definable group of patients in a particular clinical setting in routine practice.	Rothwell, P. M. (2005). External validity of randomised controlled trials:"to whom do the results of this trial apply?". *The Lancet*, *365*(9453), 82-93.

Generalisability / Applicability	(Adapted from Rothwell, 2005)	
Health economic model	A logical mathematical framework demonstrating the quantitative relationship between a defined set of variables (e.g. cost, effectiveness, net benefit) based upon an explicit set of parameters and assumptions. The purpose of modeling is to structure evidence on clinical and economic outcomes in a form that can help to inform decisions about clinical practices and health-care resource allocations. (Adapted from Weinstein, 2003).	Weinstein, M. C., O'Brien, B., Hornberger, J., Jackson, J., Johannesson, M., McCabe, C., & Luce, B. R. (2003). Principles of Good Practice for Decision Analytic Modeling in Health - Care Evaluation: Report of the ISPOR Task Force on Good Research Practices—Modeling Studies. *Value in health*, *6*(1), 9-17.
Health survey	Questionnaires designed to collect descriptions of health status and well-being, healthcare utilization, treatment patterns, and health-care expenditures from patients, providers, or individuals in the general population. (Garrisson, 2007)	Garrison et al. (2007). Using Real-World Data for Coverage and Payment Decisions: The ISPOR Real-World Data Task Force Report. Value in Health 10:5, 2007.
Health Technology Assessment	The systematic evaluation of the properties and effects of a health technology, addressing the direct and intended effects of this technology, as well as its indirect and unintended consequences, and aimed mainly at informing decision-making regarding health technologies. (HTAi, 2014)	HTAi - HTA glossary (http://htaglossary.net/health+technology+assessment+%28HTA%29)
Heterogeneity (meta-analysis)	The variability between studies included in an evidence synthesis, due to differences in population, outcomes, interventions, or study design. Such variability may lead to differences in the observed treatment effects that exceed what is expected by chance. In that case, there is statistically detectable heterogeneity. (See also "meta-analysis" and "cross-design evidence synthesis") (Adapted from Higgings, 2008 and Dias, 2013)	Higgins, J.P.T, Green,S. (eds.), Cochrane handbook for systematic reviews of interventions [Part 2, Chapter 9.5]. Vol. 5. Chichester: Wiley-Blackwell, 2008. Dias, S., Sutton, A.J., Welton, N.J., Ades, A.E., *Evidence synthesis for decision making 3 heterogeneity—subgroups, meta-regression, bias, and bias-adjustment*, Medical Decision Making 33.5: 618-640, 2013.
Heterogeneity (network meta-analysis)	An assumption sometimes made in network meta-analysis that the heterogeneity parameters are equal across all possible treatment comparisons , i.e. the variance of the random effects is the same for all treatment comparisons in the network. Such an assumption can be useful when the data are sparse (e.g. few studies per comparison), and simplifies the estimation in the presence of multi-arm studies. (Adapted from Higgins,1996, Salanti, 2008 and Dias, 2013)	Higgins, J.P.T., Whitehead, A., 1996. Borrowing Strength from External Trials in a Meta-Analysis. Stat. Med. 15, 2733–2749. doi:10.1002/(SICI)1097-0258(19961230)15:24<2733::AID-SIM562>3.0.CO;2-0. ; Salanti, G., Higgins, J.P., Ades, A.E., Ioannidis, J.P., 2008a. Evaluation of networks of randomized trials. StatMethods MedRes 17, 279–301. doi:10.1177/0962280207080643; Dias, S., Sutton, A.J., Welton, N.J., Ades, A.E., Evidence synthesis for decision making 3 heterogeneity—subgroups, meta-regression, bias, and bias-adjustment, Medical Decision Making 33.5: 618-640, 2013.
Hierarchical model	A generalization of linear and generalized linear modeling in which regression coefficients are themselves given a model, whose parameters are also estimated from an underlying distribution. For example, in a random-effects meta-analysis, the relative treatment effect parameters in the individual studies are drawn from a random effects distribution, which allows for statistical heterogeneity between studies. (See also "meta-	Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. (2014). Bayesian data analysis [3rd ed., Part 1, Chapter 5]. Boca Raton, FL, USA: Chapman & Hall/CRC.

	analysis" and "heterogeneity (meta-analysis)") (Adapted from Gelman, 2014)	
Homogeneity (meta-analysis):	The absence of heterogeneity between studies included in an evidence synthesis. (See also "heterogeneity (meta-analysis)") (IMI-GetReal, 2016)	IMI-GetReal Glosssary Workgroup, 2016
Ideal vs. Usual conditions	The context in which a treatment is being studied consists of the mode of administration, side-effects and their treatment, diet, auxiliary care, associated treatments, etc. All the characteristics of this context can be called contextual factors or extraneous factors, i.e. characteristics of the population (e.g. genetics, behaviour towards the drug, co-morbidities), the healthcare delivery system (e.g. physician behaviour of prescription, guidelines etc.), the actual use of the drug (compliance, health beliefs, co-medication, etc.). The ideal condition refers to the situation where the setting is experimental and controlled, and where the contextual factors are fixed and equalized in the two (or more) therapeutic groups through randomization, blinding and/or standardisation. For instance, the intervention status will be set at the beginning of the trial, and emphasis is put on not to change intervention status overtime, etc. Often, the design is optimised to show the most benefit of the investigated interventions. The usual condition (or routine clinical practice) refers to what really happens when an intervention is prescribed by a physician, to a patient. In routine clinical practice, the contextual factors are not fixed: they vary according to the physician's usual habits, the disease severity, patients' preferences, etc.	Schwartz, D. and J. Lellouch, Explanatory and pragmatic attitudes in therapeutical trials. J Chronic Dis, 1967. 20(8): p. 637-48.
Indirect treatment comparisons	The comparison of several (i.e. 2 or more) interventions (e.g. A and B) for a particular therapeutic indication in the absence of a trial that directly compares the interventions in question. This implies that comparison is done based upon a third intervention (e.g. C) against which A and B have been directly compared in a trial setting. (Adapted from ISPOR, 2011)	Hoaglin, D. C., Hawkins, N., Jansen, J. P., Scott, D. A., Itzler, R., Cappelleri, J. C., ... & Barrett, A. (2011). Conducting indirect-treatment-comparison and network-meta-analysis studies: report of the ISPOR Task Force on Indirect Treatment Comparisons Good Research Practices: part 2. *Value in Health*, *14*(4), 429-437.
Individual patient data (IPD) / Patient-Level Data	The raw data for subjects included in a study, as opposed to aggregate data (summary data for the comparison groups in a study). (Adapted from Bandolier, 2014)	Bandolier Glossary - IPD; http://www.medicine.ox.ac.uk/bandolier/booth/glossary/individual.html

Information bias	A flaw in measuring exposure, covariate, or outcome variables that results in different quality (accuracy) of information between comparison groups. The occurrence of information biases may not be independent of the occurrence of selection biases. (See also: "bias", "selection bias") (Porta, 2008)	Porta, M., ed. (2008). *A Dictionary of Epidemiology* (Fifth ed.). New York: Oxford University Press. p. 128. ISBN 978-0-19-531449-6.
Informative prior	A prior distribution used to incorporate prior knowledge into a Bayesian analysis. This is useful when some parameters might otherwise not be identifiable, such as the random effects variance when only a small number of studies is available. Informative priors can be based on external sources of evidence or expert judgment. A strongly informative prior may have a noticeable influence on analysis results, and therefore could invalidate results if it is biased. (Adapted from Gelman, 2014)	Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. (2014). Bayesian data analysis [3rd ed., Part 1, Chapter 2.4]. Boca Raton, FL, USA: Chapman & Hall/CRC.
Internal validity	The extent to which study attributes (e.g. study design) keep the possibility of systematic errors (i.e. biases) to a minimum. (See also "Bias") (Adapted from Rothwell, 2005)	Rothwell, P. M. (2005). External validity of randomised controlled trials:"to whom do the results of this trial apply?". *The Lancet*, *365*(9453), 82-93.
Large simple trials	Large simple trials are pragmatic, randomised clinical trials with minimal data collection protocols that are narrowly focused on clearly defined outcomes that are important to patients as well as clinicians. Their large sample size provides the adequate statistical power to detect a small difference in effects between treatments in a situation where a moderate difference in an important outcome may be important. Additionally, LST's include follow-up that mimics normal clinical practice. LST's are by definition pragmatic clinical trials. (See also "pragmatic clinical trials") (Adapted from Stroup 2011, Peto 1995)	Stroup, T.S., What can large simple trials do for psychiatry? Am J Psychiatry, 2011. 168(2): p. 117-9.; Peto, R., R. Collins, and R. Gray, Large-scale randomized evidence: large, simple trials and overviews of trials. J Clin Epidemiol, 1995. 48(1): p. 23-40; Clinical Trials Transformation Initiative. Large Simple Trials: Facilitating the Use of Large
Longitudinal study	A study in which subjects are followed over time with continuous or repeated monitoring of risk factors or health outcomes, or both. (Adapted from BMJ, 2014)	BMJ, Longitudinal studies http://www.bmj.com/about-bmj/resources-readers/publications/epidemiology-uninitiated/7-longitudinal-studies
Marketing authorisation	A licence given by a regulatory authority to a pharmaceutical manufacturer allowing for the marketing of a specific product within the jurisdiction of the regulatory agency. The decision for granting a marketing authorisation is based primarily on the quality, safety and efficacy of the new medicinal product. (Adapted from MHRA, 2014)	MHRA - Marketing Authorisations (http://www.mhra.gov.uk/Howweregulate/Medicines/Licensingofmedicines/Marketingauthorisations/index.htm#l3)
Markov model	Stochastic multi-state transition models. Transition between different states in the model are based upon pre-defined conditional probabilities that depend only on the current states (i.e. future evolution depends only on the current state, not the past states). (See also: "multi-state transition model") (Adapted from Briggs, 1993)	A. Briggs and M. Sculpher, "An introduction to Markov modelling for economic evaluation," *PharmacoEconomics*, vol. 13, no. 4, pp. 397–409, Apr. 1998. Sonnenberg FA, Beck JR. Markov models in medical decision making: a practical guide. Med Decis Making. 1993 Oct-Dec;13(4):322-38.

Medicine Adaptive Pathways to Patients (MAPP's)	A prospectively planned, flexible approach to regulation of drugs and biologics. It starts with the early authorisation of a medicine in a restricted patient population, followed by iterative phases of evidence-gathering and the adaptation of the marketing authorisation to allow broader patient populations to access the medicine. Adaptive pathways are sometimes known as 'staggered approval' or 'progressive licensing'. (Adapted from Eichler, 2012 and EMA, 2014).	Eichler, H. G., Oye, K., Baird, L. G., Abadie, E., Brown, J., Drum, C. L., ... & Hirsch, G. (2012). Adaptive licensing: taking the next step in the evolution of drug approval. *Clinical Pharmacology & Therapeutics*, *91*(3), 426-437.; http://www.ema.europa.eu/ema/index.jsp?curl=pages/regulation/general/general_content_000601.jsp&mid=WC0b01ac05807d58ce
Meta-analysis	Statistical methods allowing for the combination of study outcomes from several studies, while also accounting for uncertainties resulting from the combination of data from varying sources. (See also "evidence synthesis") (Adapted from Greenland S., 2008)	Greenland S, O' Rourke K: Meta-Analysis. Page 652 in Modern Epidemiology, 3rd ed. Edited by Rothman KJ, Greenland S, Lash T. Lippincott Williams and Wilkins; 2008.
Meta-regression	Assessment of the extent of the effect of moderator variables on outcomes presented by a model in the context of meta-analysis. This is achieved through the use of regression techniques, and can be based on both aggregate and individual patient level data. Data for the regression analysis may originate, and be synthesised, from several studies. (See also: "meta-analysis") (Adapted from Stanley, 1989)	T.D. Stanley and Stephen B. Jarrell, (1989). Meta-regression analysis: A quantitative method of literature surveys. *Journal of Economic Surveys*, 19(3) 299-308.
Misclassification bias	Possibility of error associated with the classification of study participants with regards to their exposure status, outcome status, or disease status. Random occurrence of information bias is known as 'non-differential misclassification' while non-random occurence is known as 'differential misclassification'. Missclasification bias falls within the category of information bias. (See also: "Bias" and "Information bias".) (Adapted from Strom, 2006 and Delgado-Rodriguez, 2004)	Strom B.L., Kimmel S.E. (2006). Textbook of pharmacoepidemiology. John Wiley & Sons, Ltd.; Delgado-Rodríguez, M., & Llorca, J. (2004). Bias. Journal of epidemiology and community health, 58(8), 635-641.
Multi-state transition model	Multi-state transition models are used to model prognosis for clinical problems with ongoing risks. The model assumes that the patient is always in one of a finite number of clinical conditions, referred to as states. All events of interest are modelled as transitions from one state to another. Multi-state transition models can be either discrete or continuous time models. For the former the transitions are only possible at certain time points. The time interval from one time point to the next one is referred to as a cycle. The net probability of making a transition from one state to another during a single cycle is called transition probability. This transition probability can be constant over time or not. Multi-state transition models can model a cohort of patients or individuals. In the latter case they are called microsimulation models. A prominent example is the illness-death model with three states representing health, illness and death. (Adapted from Sieber et al, 2012)	Siebert U, Alagoz O, Bayoumi AM, Jahn B, Owens DK, Cohen DJ, Kuntz KM; ISPOR-SMDM Modeling Good Research Practices Task Force. State-transition modeling: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force--3. Value Health. 2012 Sep-Oct;15(6):812-20. doi:10.1016/j.jval.2012.06.014.

Network meta-analysis (NMA)	An extension of meta-analysis, allowing for the comparison of the relative effects of multiple treatments, either with or without the presence of a common comparator against which all interventions are studied. NMA methods take into account Direct and Indirect Treatment Comparisons and pairwise meta-analysis. (see also: "transitivity assumption", "direct treatment comparisons" and "indirect treatment comparison") (Adapted from Athanasiou, 2011 & Hawkins, 2011)	Athanasiou T., Darzi A. (2011). Evidence Synthesis in Healthcare: A practical Handbook for Clinicians. Springer, ISBN 978-0-85729-175-8, doi: 10,1007/978-0-85729-206-3; Hoaglin, D. C., Hawkins, N., Jansen, J. P., Scott, D. A., Itzler, R., Cappelleri, J. C., ... & Barrett, A. (2011). Conducting indirect-treatment-comparison and network-meta-analysis studies: report of the ISPOR Task Force on Indirect Treatment Comparisons Good Research Practices: part 2. Value in Health, 14(4), 429-437.
Non-informative prior	A prior distribution that does not contribute at all to the posterior estimation of a Bayesian analysis. Such priors are often "improper", in that they are not true probability distributions. Although they let the data "speak for themselves", they have the disadvantage that models using non-informative priors may not be identifiable or difficult to estimate. It is, however, increasingly acknowledged that non-informative may not exist, since all priors contain some information. (Adapted from Gelman, 2014 and Lambert, 2005)	Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B., Bayesian data analysis [3rd ed., Part 1, Chapter 2.8]. Boca Raton, FL, USA: Chapman & Hall/CRC, 2014; Lambert, P.C., Sutton, A.J., Burton, P.R., Abrams, K.R., Jones, D.R., 2005. How vague is vague? A simulation study of the impact of the use of vague prior distributions in MCMC using WinBUGS. Stat Med 24, 2401–2428. doi:10.1002/sim.2112
Non-interventional study / Observational study	A study where the investigator does not interfere with choice of the prescribed health intervention i.e. interventions are prescribed in the usual manner in accordance with clinical practice. The assignment of the patient to a particular therapeutic strategy is not decided in advance by a trial protocol but falls within current practice and the administration of the intervention is clearly separated from the decision to include the patient in the study. No additional diagnostic or monitoring procedures are applied to the patients and epidemiological methods are used for the analysis of collected data. (See also "real world study") (Adapted from EC, 2001) Footnote: Depending on local regulations, something considered to be interventional in one country (blood tests and patient questionnaires) may not be in others.	European Commission (2001). DIRECTIVE 2001/20/EC OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL. *Official Journal of the European Cmmunities.*
Observational data	Data collected from populations as present in the routine setting of healthcare (i.e. outside the setting of a randomised controlled trial). Sources of observational may data include: routine clinical practice, patient registeries, hospital claims databases/administrative data, health surveys, electronic health records, medical chart reviews and post-marketing safety studies. (Adapted from Machin, 2007)	Machin D., Campbell M.J., Walters S.J. (2007). Medical Statistics: A Textbook for the Health Sciences (4th ed.). Wiley, Chapter 12.
Patient-Reported Outcome (PRO)	A measurement based on a report that comes directly from the patient (i.e., study subject) about the status of a patient's health condition without amendment or interpretation of the patient's response by a clinician or anyone else. A PRO can be measured by self-report or by interview provided that the interviewer records only the patient's response. (FDA, 2009)	FDA (2009) - Guidance for Industry Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims.

Phase 2 trials	Trials in which the primary objective is to explore therapeutic efficacy and safety in patients. (NIH, 2001)	National Institute of health (2001) - Glossary of terms for human subjects protection and inclusion issues. http://grants.nih.gov/grants/peer/tree_glossary.pdf
Phase 3 trials	Trials implemented to study the efficacy of the biomedical or behavioral intervention in large groups of human subjects (from several hundred to several thousand) by comparing the intervention to other standard or experimental interventions as well as to monitor adverse effects, and to collect information that will allow the intervention to be used safely. (NIH, 2001)	National Institute of health (2001) - Glossary of terms for human subjects protection and inclusion issues. http://grants.nih.gov/grants/peer/tree_glossary.pdf
Phase 4 trials	Studies implemented in a post-marketing setting that are designed to monitor effectiveness of the approved intervention in the general population and to collect information about any adverse effects associated with widespread use. (NIH, 2001)	National Institute of health (2001) - Glossary of terms for human subjects protection and inclusion issues. http://grants.nih.gov/grants/peer/tree_glossary.pdf
PICOT(S)	PICOT is an acronym for a method used in evidence based medicine to frame and answer a clinical quesion. "P" stands for patient, problem or population, "I" for intervention, "C" for comparison, control or comparator, "O" for outcomes and "T" for time. Alternative versions also include "S" for study design. (Riva, 2012)	Riva, J. J., Malik, K. M. P., Burnie, S. J., Endicott, A. R., & Busse, J. W. (2012). What is your research question? An introduction to the PICOT format for clinicians. The Journal of the Canadian Chiropractic Association, 56(3), 167–171.
Post-authorisation	The period after market authorization of a specific pharmaceutical/ medical device product. (Adapted from Rang, 2006)	Rang H.P. (2006). Drug Discovery and Development: technology in transition. Elsevier press (2006).
Post-authorisation efficacy studies	Studies conducted within the authorised therapeutic indication to complement available efficacy data in the light of well-reasoned scientific uncertainties on aspects of the evidence of benefits that should be, or can only be, addressed post-authorisation(EU Pharmacovigilance Directive). Footnote: Although the term refers to "efficacy", PAES studies collect data in a setting that reflects general clinical practice rather than a randomised controlled trial. Therefore it may be more appropriate to think of these studies as providing "effectiveness" data rather than "efficacy" data.	EMA (2015) - Scientific guidance on post-authorisation efficacy studies: http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2015/11/WC500196379.pdf.
Post-authorisation safety studies	Any study relating to an authorised medicinal product conducted with the aim of identifying, characterising or quantifying a safety hazard, confirming the safety profile of the medicinal product, or of measuring the effectiveness of risk management measures. Such studies may be aimed at collecting data to enable the assessment of safety of medicinal products in everyday medical practice. (EU, 2010)	EU Pharmacovigilance Directive 2010/84/EU
Post-marketing	The period after launch of a specific pharmaceutical/ medical device product in the market. This follows the post-authorisation phase. (Wyeth et al., 2013)	Wyeth et al. (2013). Postmarket Safety Surveillance of Drugs and Therapeutic Biologics. http://www.fda.gov/downloads/forpatients/about/ucm410175.pdf

Pragmatic clinical trial	A study comparing health interventions among a randomised, diverse population representing clinical practice, and measuring a broad range of health outcomes. To ensure generalizability, pragmatic trials should represent the intended patients to whom the treatment will be applied as best as possible. For instance, inclusion criteria would be broad (e.g. allowing co-morbidity, co-medication, wider age range, etc.), the follow-up would not be (or not much) interventional and allowing for treatment switching etc. (See also "large simple trials" and "real-world studies") (Adapted from Schwartz, 1967, Tunis, 2003 & Roland, 1998)	Schwartz, D., & Lellouch, J. (1967). Explanatory and pragmatic attitudes in therapeutical trials. *Journal of chronic diseases*, *20*(8), 637-648.; Tunis, S. R., Stryer, D. B., & Clancy, C. M. (2003). Practical clinical trials: increasing the value of clinical research for decision making in clinical and health policy. Jama, 290(12), 1624-1632.; Roland, M. and D.J. Torgerson, What are pragmatic trials? BMJ, 1998. 316(7127): p. 285.
Pre-authorisation	The period prior to the granting of the marketing authorisation ("approval") / license for a specific pharmaceutical/ medical device product. (see also: "pre-authorisation drug development") (Rang H.P., 2006)	Rang H.P. (2006). Drug Discovery and Development: technology in transition. Elsevier press (2006).
Pre-authorisation drug development	Research and development performed in order to discover new active substances and subsequently demonstrate safety and efficacy of a new drug in order to gain marketing authorisation. This includes the discovery phase, pre-clinical studies (Phase I) and clinical studies (Phase II-III). (Rang H.P., 2006)	Rang H.P. (2006). Drug Discovery and Development: technology in transition. Elsevier press (2006).
Predictive modelling	The activity of developing, validating or adapting models which relate a dependent variable with a set of independent variables in a manner similar to multiple regression analysis. For the purposes of the IMI-GetReal project, this predictive modelling is desinged to predict the (relative) effectiveness of a medical intervention from available clinical and trial-based efficacy/effectiveness data. (Wilson, 2004)	Wilson, A. M., Thabane, L., & Holbrook, A. (2004). Application of data mining techniques in pharmacovigilance. British journal of clinical pharmacology, 57(2), 127-134.
Propensity score	The conditional probability of assignment to a particular intervention given a vector of observed covariates. As such, the propensity score summarizes information of multiple confounders and can be used to adjust for confounding by means of matching, weighting or regression adjustment. (Adapted from Rosenbaum, 1981)	Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, *70*(1), 41-55.
Protopathic bias	Protopathic bias arises when the initiation of an intervention (exposure) occurs in response to a symptom of the (at this point, undiagnosed) disease under study (outcome). This sort of bias falls within the category of information bias. (Adapted from Strom, 2006 and Delgado-Rodriguez, 2004)	Strom B.L., Kimmel S.E. (2006). Textbook of pharmacoepidemiology. John Wiley & Sons, Ltd.; Delgado-Rodríguez, M., & Llorca, J. (2004). Bias. Journal of epidemiology and community health, 58(8), 635-641.
Randomisation	The process of assigning trials participants to treatment or control groups, using an element of chance to determine the assignments. Randomization is a tool for providing comparable participant groups regarding all measurable and unmeasurable characteristics, apart from	Strom, B. and S.E. Kimmel, Textbook of Pharmacoepidemiology. 2006: John Wiley & Sons, Ltd.

	the intervention ("active" or control). It ensures the initial comparability of patients between the intervention groups. By randomization, one hopes to equalize the distributions of confounding factors, whether they are known or unknown. (Adapted from Strom, 2006)	
Randomised Controlled clinical Trial (RCT)	Clinical trial designed to test the efficacy of a medical intervention within a population of selected subjects. Subjects are subjected to rigorous inclusion and exclusion criteria, and upon inclusion, are randomised into separate treatment and control groups. Considered the gold standard for clinical trials. (See also "clinical trial") (Adapted from Solomon, 2009)	Solomon P., Cavanaugh M.M., Draine J. (2009). Randomoized Controlled Trials: design and implementation for community-based psychosocial interventions. Oxford Press.
Real World Data (RWD)	An umbrella term for data regarding the effects of health interventions (e.g. safety, effectiveness, resource use, etc) that are not collected in the context of highly-controlled RCT's. Instead, RWD can either be primary research data collected in a manner which reflects how interventions would be used in routine clinical practice or secondary research data derived from routinely collected data. Data collected include, but are not limited to, clinical and economic outcomes, patient-reported outcomes (PRO) and health-related quality of life (HRQoL). RWD can be obtained from many sources including patient registries, electronic medical records, and claims databases. (See also "randomised controlled clinical trial", "real-world evidence" and "real-world study")(Adapted from Garrisson, 2007 and GetReal, 2016)	Garrison et al. (2007). Using Real-World Data for Coverage and Payment Decisions: The ISPOR Real-World Data Task Force Report. Value in Health 10:5, 2007.; IMI-GetReal Glossary Workgroup 2016.
Real World Evidence (RWE)	Real World Evidence (RWE) is the evidence derived from the analysis and/or synthesis of real-world data (RWD). (See also "real-world data", "real-world study")(GetReal, 2016)	IMI-GetReal Glossary Workgroup, 2016
Real World Study (RWS)	Studies investigating health interventions whose design does not follow the design of a highly-controlled RCT and aims to reflect health intervention effectiveness in routine clinical practice. Real world studies do not typically include randomisation of trial subjects, but there are exceptions (e.g. pragmatic clinical trials). For the purposes of GetReal, real-world studies include, but are not limited to, the following: pragmatic clinical trials, non-interventional/ observational studies, drug utilisation studies, post-authorisation efficacy/safety studies. RWS, by definition, generate RWD, which can subsequently be analysed and/or synthesised to produce RWE. (See also: "real-world data", "real-world evidence", "effectiveness study", "drug utilisation study", "pragmatic clinical trial" and "non-interventional/ observational study") (GetReal, 2016)	IMI-GetReal Glossary Workgroup, 2016
Referral bias	Occurs if reasons for referring a patient to a health intervention/ health care institute are related to the drug exposure status e.g. when the use	Strom B.L., Kimmel S.E. (2006). Textbook of pharmacoepidemiology. John Wiley & Sons, Ltd.; Delgado-Rodríguez, M., & Llorca, J. (2004). Bias.

	of the drug contributes to the diagnostic process. This sort of bias falls within the category of selection bias. (See also "Selection bias") (Adapted from Strom, 2006 and Delgrado-Rodriguez, 2004)	Journal of epidemiology and community health, 58(8), 635-641.
Registry	Database resulting from prospective, observational cohort studies of patients who have a particular disease and/or are receiving a particular treatment or intervention. (Garrisson, 2007)	Garrison et al. (2007). Using Real-World Data for Coverage and Payment Decisions: The ISPOR Real-World Data Task Force Report. Value in Health 10:5, 2007.
Relative effectiveness	The extent to which an intervention does more good than harm, when compared to one or more alternative interventions for achieving the desired results and when provided under the routine setting of health care practice. (See also "ideal vs.usual conditions") (High Level Pharmaceutical Forum, 2008)	High Level Pharmaceutical Forum. High Level Pharmaceutical Forum 2005-2008. Final Conclusions and Recommendations of the High Level Pharmaceutical Forum. URL: http://ec.europa.eu/pharmaforum/docs/final_conclusions_en.pdf Accessed 7 March 2014
Relative efficacy	The extent to which an intervention does more good than harm when compared to one or more alternative interventions under ideal conditions. (See also "ideal vs.usual conditions") (High Level Pharmaceutical Forum, 2008)	High Level Pharmaceutical Forum. High Level Pharmaceutical Forum 2005-2008. Final Conclusions and Recommendations of the High Level Pharmaceutical Forum. URL: http://ec.europa.eu/pharmaforum/docs/final_conclusions_en.pdf Accessed 7 March 2014
Risk-sharing agreements	Agreements concluded by payers and pharmaceutical companies to diminish the impact on the payer's budget of new and existing medicines brought about by either the uncertainty of the value of the medicine and/or the need to work within finite budgets. (Adamski, 2010)	Adamski, J., Godman, B., Ofierska-Sujkowska, G., Osińska, B., Herholz, H., Wendykowska, K., ... & Gustafsson, L. (2010). Risk sharing arrangements for pharmaceuticals: potential considerations and recommendations for European payers. BMC health services research, 10(1), 153.
Safety	The presence or absence of an adverse effect as a direct or indirect result of treatment intervention. (See also "adverse event") (ICH, 1994)	ICH (1994). Clinical Safety Data Management: defnitions and standards for expedited reporting. http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E2A/Step4/E2A_Guideline.pdf
Scientific advice	Advice given by a regulatory/ reimbursement authority to a manufacturer on appropriate tests and studies to be performed during product development/ application for product reimbursement, in order to avoid major objections being raised during evaluation of the marketing authorisation application/ reimbursement application. (Adapted from EMA, 2009)	EMA - Scientific advice and protocol advice. http://www.ema.europa.eu/ema/index.jsp?curl=pages/regulation/general/general_content_000049.jsp&mid=WC0b01ac05800229b9
Selection bias	A systematic error introduced due to the selective inclusion of subjects who differ in characteristics from the target population, or selective drop-out of subjects in a study. (Adapted from Strom, 2006 and Delgrado-Rodriguez, 2004)	Strom B.L., Kimmel S.E. (2006). Textbook of pharmacoepidemiology. John Wiley & Sons, Ltd.; Delgado-Rodríguez, M., & Llorca, J. (2004). Bias. Journal of epidemiology and community health, 58(8), 635-641.
Self-selection bias	An error occuring when study participants decide themselves to particpate in or to leave a study based on both drug exposure and change in health status. This sort of bias falls within the category of selection bias. (See also "Selection bias") (Adapted from Strom, 2006	Strom B.L., Kimmel S.E. (2006). Textbook of pharmacoepidemiology. John Wiley & Sons, Ltd.; Delgado-Rodríguez, M., & Llorca, J. (2004). Bias. Journal of epidemiology and community health, 58(8), 635-641.

	and Delgrado-Rodriguez, 2004)	
Stakeholder	For the purposes of the IMI-GetReal project, a stakeholder is an individual, organisation or initiative that participates in, is involved with, influences the outcomes of, or is influenced by the outcomes of, or the implications of, the IMI-GetReal project. (Adapted from Varvasovszky, 2000 and Freeman, 2010)	Varvasovszky, Z., & Brugha, R. (2000). A stakeholder analysis. Health policy and planning, 15(3), 338-345.; Freeman, R. E. (2010). Strategic management: A stakeholder approach. Cambridge University Press.
Standard of care	Care delivered by a healthcare provider for a specific patient which should correspond to the care that an averagely competent physician in the same field would provide under similar circumstances. (Adapted from Strauss, 2009)	Strauss D.C., Thomas J.M. (2009). What Does the Medical Profession Mean By "Standard of Care?" JCO Nov 10, 2009: e192-193; published online on September 21, 2009.
Statistical significance	A measure that describes the extent to which conclusions from a statistical analysis are supported by the data. Alternatively, it can be defined as the conclusion that findings from statistical analysis are true (i.e. not true due to random chance) for the sample studied, thus expected to be highly reliable. Footnote: High statistical significance does not necessarily mean that the finding is of great importance or decision-making utility. (See also "clinical significance") (Adapted from Redmond, 2001)	Redmond, C., Colton, T., *Clinical significance versus statistical significance, Biostatistics in Clinical Trials*. Wiley Reference Series in Biostatistics [3rd ed.]. West Sussex, United Kingdom: John Wiley & Sons Ltd. pp. 35–36. ISBN 0-471-82211-6, 2001.
Sub-group analysis	Anlaysis conducted to assess whether, and how, observed endpoints in a study are sensitive to/ affected by characteristics attributable to a specific sub-group of the study population. Such characteristics can relate to, for example, age, gender, genotype, or patient history. (Adapted from ICH, 1998)	ICH (1998) - E9 Guideline: Statistical Princinples for Clinical Trials. http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E9/Step4/E9_Guideline.pdf. Biomarkers Definitions Working Group.Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. Clin Pharmacol Ther. 2001 Mar;69 (3):89-95.
Surrogate endpoint/ outcome	A variable that provides an indirect measurement of effect in situations where direct measurement of clinical effect is not feasible or practical. It is intended to replace a clinical endpoint of interest that cannot be observed in a trial. A surrogate endpoint is expected to predict clinical benefit (or harm or lack of benefit or harm) based on epidemiologic, therapeutic, pathophysiologic, or other scientific evidence. (Adapted from ICH, 1998 and Biomarkers Definitions Working Group, 2001)	ICH (1998) - E9 Guideline: Statistical Princinples for Clinical Trials. http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E9/Step4/E9_Guideline.pdf; Biomarkers Definitions Working Group.Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. Clin Pharmacol Ther. 2001 Mar;69 (3):89-95.
Time horizon	Period of time within a pharmacoeconomic model over which costs and health outcomes are gathered during a pharmacoeconoimc evaluation. (Ademi, 2013)	Ademi, Z., Kim, H., Zomer, E., Reid, C. M., Hollingsworth, B., & Liew, D. (2013). Overview of pharmacoeconomic modelling methods. *British journal of clinical pharmacology*, *75*(4), 944-950.
Transitivity assumption	If the safety, efficacy, effectiveness or HRQoL of health intervention A is quantitatively related to that of B, and that of B to C, then A is related to C for all A, B and C in the domain of the relation. This is a fundamental assumption of indirect treatment comparisons and network meta-	Athanasiou T., Darzi A. (2011). Evidence Synthesis in Healthcare: A practical Handbook for Clinicians. Springer, ISBN 978-0-85729-175-8, doi: 10,1007/978-0-85729-206-3; Efthimiou O., Debray, T.P.A., van Valkenhoef G., et al. (2016). GetReal in network meta-analysis: a review of

	analysis. (see also:"mixed treatment comparisons") (Athanasiou, 2011 and Efthimiou, 2016)	methodology. Wiley, Research Synthesis Methods 1759-2887, http://dx.doi.org/10.1002/jrsm.1195
Vague (weakly informative) prior	A prior distribution that provides a very small amount of information that allows a model to be estimated, while not having a large influence on the posterior estimation of a Beysian analysis. For example, a vague prior might specify that heterogeneity will not cause more than five orders of magnitude differences in the odds ratio, or that a change in blood pressure lies in the -500 to +500 mmHg range with 95% probability. In practice, vague priors may offer a better compromise between objectivity and practicability. (Adapted from Gelman, 2014)	Gelman, A., Carlin, J.B., Stern, H.S., & Rubin, D.B., Bayesian data analysis [3rd ed., Part 1, Chapter 2.9]. Boca Raton, FL, USA: Chapman & Hall/CRC, 2014.
Work package	A specific subset of a project assigned to the execution of a specific aim. Work packages are defined by brief statements of: Activity Description, Activity Resources of Skill and Expertise, Activity Estimates of Effort and Duration, Activity Schedule, Activity Risks, and Activity Budget. (Weiss, 1992)	Weiss J.W. (1992). 5-Phase Project Management: a practical planning & implication guide. Addison-Wesley, 1992. ISBN 0201563169, 9780201563160

5. Comments Received during Consultation Rounds in 2015 & 2016

A. Comments received during public consultation 2015

Organisation	Glossary Term	Comments Received	Authors' Response
EMA	**Adaptive Clinical Trial**	Given this is an EU funded project, a reference to the CHMP Reflection paper on methodological issues in confirmatory clinical trials planned with an adaptive design would be appropriate. There is another good reason to do so I will explain later on. http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500003616.pdf Both the FDA and CHMP guidance is clearly focussed at Adaptive RCTs.	Incorporated: The source mentioned was consulted and the definition was adapted accordingly. The source has been listed as a reference.

Pfizer	**Bias**	Systematic (non-random) errors in values of **parameters** that are the object of measurement… Another book useful here is: TL Lash, MP Fox. 2009. Applying quantitative bias analyses to epidemiologic data	Incorporated: spelling of 'parameters' corrected as indicated by reviewer.
NICE	**Case-control study**	"A retrospective study (applied to registry data) in which past exposure of a group with an outcome (cases) is compared to the exposure of a group without the outcome (controls). To determine the proportion of people with or without a certain outcome, has/had exposure or characteristic of interest." (Polygenisis et al., 'ISPOR Taxonomy of Patient Registries: Classification, Characteristics and Terms', 2013)	Partially incorporated: Although proposed definition highlighted some missing elements (e.g. retrospoctive study), the authors find the limitation of case-control studies to those applied to registry data too restrictive. It is suitable for the purposes of the ISPOR report cited in the reference, but is too narrow in scope for GetReal.
NICE	**Cohort study**	"A prospective, descriptive study model, applied in registries and for studies evaluating comparative effectiveness, safety, and/or quality of care. Cohort studies follow a group of people with certain shared characteristics to observe whether or not they develop a particular endpoint or outcome over time." (Polygenisis et al., 'ISPOR Taxonomy of Patient Registries: Classification, Characteristics and Terms', 2013)	Not incorporated: A cohort study can refer to either prospective or retrospective studies. In addition, restricting the definition to registries alone makes the scope too narrow. It is suitable for the purposes of the ISPOR report cited, but is too narrow in scope for GetReal.
NICE	**Comparative effectiveness research**	Removed the term "systematic" from the first sentence and qualify that this is in routine clinical practice. Purpose is also removed as we do not define the purpose of other terms). Added clarification that comparative effectiveness research is conducted in a real world setting. "The conduct and synthesis of research comparing different benefits and harms of alternative interventions and strategies to prevent, diagnose, treat, and monitor health conditions in routine clinical practice (the real world setting). Comparative effectiveness research includes both primary data collection and secondary analyses (such as systematic literature reviews, meta-analyses and economic evaluations)"	Incorporated: The proposed changes to the definition were implemented.

NICE	**Effectiveness**	Definition expanded. The distinciton between efficacy and effectiveness isn't always made; there is confusion in the literature because of the colloquial use of the term "effectiveness", based on its definition in other (non-scientific) contexts, when the intention is to refer to efficacy. The health benefits associated with an intervention in routine clinical practice (i.e. 'real world' studies) where multiple variables that might influence patient outcomes (such as concomitant medication and co-morbidities) are introduced.	Not incorporated: Although the suggestions made by the reviewer are useful, authors agreed to adopt the definitions of the High Level Pharmaceutical Forum for the terms Efficacy, Effectiveness, Relative Efficacy and Relative Effectiveness for consistency and clarity.
ZIN	**Effectiveness**	I would suggest to use the HPLF definition instead	Incorporated: The HLPF definition for the terms Efficacy, Effectiveness, Relative Efficacy and Relative Effectiveness were adopted.
NICE	**Effectiveness study**	Expanded using elements from the PRECIS tool and AHRQ tool. Effectiveness studies (also referred to as real-world studies) measure the outcomes (both positive and negative) associated with a health intervention when prescribed in routine clinical practice and usual prescription conditions. Observational/non-interventional studies, pragmatic studies and large simple trials are all types of effectiveness studies. Key defining criteria of effectiveness studies include: multiple active comparators; no eligibility criteria; clinically meaningful health outcomes such as quality of life, functional capacity and mortality (as opposed to objective outcomes); long duration; and large sample size. Randomisation does not preclude a study from being able to measure effectiveness. Randomisation does not preclude a study from being able to measure effectiveness, however these trials would need comparator arms other than placebo to be representative of real-world populations.	Partially incorporated: The authors believe that the listing of "no eligibility criteria" is too restrictive, as it would exclude Pragmatic Clinical Trials (PCT's) from what would be defined as effectiveness studies. Therefore, this has been changed to broad eligibility criteria. The suggestions made regarding the key defining criteria to effectiveness was added to the definition. Additionally, the note on randomisation not precluding a study from meauring effectiveness has been incorporated into the definition.

NICE	**Efficacy**	Agreeing with the definition but adding further supporting information: "The extent to which a healthcare intervention produces a therapeutic benefit as compared to a placebo intervention under ideal conditions (i.e. the highly-controlled conditions of a randomised controlled trial)." This way, RCTs/efficacy trials are distinguished from realworld/effectivness trials. 'RCTs' that assess effectiveness (pragmatic trials) fall into the realworld bucket. Replaced therapeutic effect with therapeutic benefit to avoid any confusion with effectiveness.	Partially incorporated: The authors agree to specifying that highly-controlled conditions ion RCT's are meant by ideal conditions, therefore this has been added to the definition. As for the re-wording of 'therapeutic effect' to 'therapeutic benefit', this has been avoided. It is best to consistently use the term therapeutic effect to match the definitions of effectiveness, relative efficacy and relative effectiveness.
NICE	**Efficacy study**	This definition could be perceived to be misleading: It provides RCTs as only one 'type' of efficacy study. What examples would it give for the other types of efficacy study? If this is referring to the well-known hierarchy of evidence (pyramid) concept, RCTs are the only 'efficacy' trials in the pyramid. The others (with the exception of systematic reviews and metanalysis) are RW studies: cohort studies, case-control studies etc. The wording has been modified to address this: "Efficacy studies, often referred to as explanatory studies, measure the therapeutic benefits of a drug under highly controlled conditions (i.e. randomised controlled trials) and prove a causal relationship between an intervention and an outcome. They serve to answer the question of "can it work?", in an ideal world. They will also try to explain why it works. By contrast, effectiveness studies answer the question "does it work?""	Incorporated: The comment made by the reviewer is a very good one. Additionally, the supposed changes are too. Therefore, the recommended changes to the definition have been incorporated.

NICE	**Large simple trials**	Most bodies (ISPOR, PCORI, Quintiles) use the terms 'large simple trials' and 'pragmatic clinical trials' interchangably (ISPOR). Others refer only to pragmatic trials and don't use the term LST. The only body to explicitly distinguish between the two was the National Information Center on Health Services Research and Health Care Technology (NICHSR): - Large simple trials: retain the methodological strengths of prospective, randomised design, but use large numbers of patients, more flexible patient entry criteria and multiple study sites to generate effectiveness data and improve external validity. Fewer types of data may be collected for each patient, easing participation by patients and clinicians. Prominent examples of include the GISSI trials of thrombolytic treatment of acute myocardial infarction (AMI) (Maggioni 1990), the ISIS trials of alternative therapies for suspected AMI (Fourth International Study of Infarct Survival 1991), and the CATIE trial of therapies for schizophrenia (Stroup 2003). - Pragmatic trials are a related group of trial designs whose main attributes include: comparison of clinically relevant alternative interventions, a diverse population of study participants, participants recruited from heterogeneous practice settings, and data collection on a broad range of health outcomes. They explain that some large simple trials are also pragmatic trials.	Incorporated: The point raised by the reviewer here is a good one: although the terms 'large simple trials' and 'pragmatic clinical trials' are used interchangeably, the authors have agreed to adopt the ideas demonstrated by the NICHSR which identify PCT's as a sub-set of large simple trials. The definition has thus been updated accordingly. A cross-reference to PCT's has also been added.
Pfizer	**Line 165**	If the intervention is dictated or prescribed by the investigator, rather than being self-selected by the patient, then the study design is referred to as a quasi-experimental design. See Shadish, Cook, and Campbell.2001.Experimental and Quasi-Experimental Designs for Generalized Causal Inference.	Not incorporated: Although the point raised by the reviewer is factually correct, it does not fit in the flow of ideas in this narrative. This paragraph specifically focuses the consensus on features of observational studies. Therefore the paragraph was left unchanged.

Pfizer	**Longtitudinal Study**	Three books also useful here are Fairclough, D.L. 2010. Design and analysis of quality of life studies in clinical trials. 2nd edn. Boca Raton, Florida: Chapman & Hall/CRC. Fitzmaurice, G.M., Laird, N.M. and J.H. Ware. 2011 . Applied longitudinal analysis. 2nd edn. Hoboken, New Jersey: John Wiley & Sons. Singer, J.D. and J.B. Willett. 2003. Applied longitudinal data analysis: Modeling change and event occurrence. New York, NY: Oxford University Press.	Not incorporated: Given that there are no objections to the current definition which would require referall to alternative sources, it is not relevant to cite the references mentioned here. However, the authors also agree that the books mentioned provide good references for definitions of the term.
NICE	**Non-interventional/ Observational study**	Little, if any controversy over this, except with local variation, and a footnote has been added to this effect. Additionally, all the potential study types have been listed, so there is clarity when people cross reference this definition from others. "A study where the investigator does not interfere with choice of the prescribed health intervention i.e. interventions are prescribed in the usual manner in accordance with the terms of the marketing authorisation. The assignment of the patient to a particular therapeutic strategy is not decided in advance by a trial protocol but falls within current practice and the administration of the intervention is clearly separated from the decision to include the patient in the study. No additional diagnostic or monitoring procedures are applied to the patients and epidemiological methods are used for the analysis of collected data. Types of observational studies include: routine clinical practice, patient registeries, hospital claims databases/administrative data, health surveys, electronic health records, medical chart reviews and post-marketing safety studies. " Footnote: Depending on local regulations, something considered to be interventional in one country (blood tests and patient questionnaires) may not be in others.	Partially incorporated: Although the authors agree to the suggestions made regarding local variation, there are some controversies regarding the listing of studies made here by the reviewer: the list conflates RWD sources and RWD-generating study designs. The authors therefore suggested adding a cross-reference to real-world studies (RWS) in stead. A foot note has been added regarding the local variation in what constitutes an intervention.
NICE	**Observational data**	"Data collected from patients populations as present in the routine setting of healthcare (i.e. outside the setting of a randomised controlled trial). Sources of observational data include: routine clinical practice, patient registeries, hospital claims databases/administrative data, health surveys, electronic health records, medical chart reviews and post-marketing safety studies."	Incorporated: The extra examples of observational data listed by the reviewer have been added to the definition.

NICE	**Post-authorisation efficacy study**	Although the term refers to "efficacy", these studies collect data in a setting that reflects general clinical practice rather than a randomised controlled trial. Therefore it may be more appropriate to think of these studies as providing "effectiveness" data rather than "efficacy" data.	Incorporated: The comment made by the reviewer here has been added as a footnote to the definition.
MHRA	**Pragmatic Clinical Trial**	There is some internal inconsistency in the definition of pragmatic clinical trials between the two documents that were circulated, one indicating that a randomisation ratio changing over time being part of the definition of what constitutes PCT (Section 6, paragraph 3), and this aspect being absent / replaced in the definitions in the second document.	Incorporated: Even though the definitions in the documents referred to both indicate that randomisation is part of PCT's definition, the wording in the glossary was not as clear as in the report. Therefore, the definition was re-worded to highlight the fact that the designing of PCT's involves randomisation of the trial population.
NICE	**Pragmatic clinical trial**	See comment on 'large simple trials'	Incorporated: The point raised by the reviewer here is a good one: although the terms 'large simple trials' and 'pragmatic clinical trials' are used interchangeably, the authors have agreed to adopt the ideas demonstrated by the NICHSR which identify PCT's as a sub-set of large simple trials. The definition has thus been updated accordingly. A cross-reference to PCT's has also been added.

EMA	**Propensity-Based Trial**	This is a bit of a novel concept, mainly because I am not clear what they are proposing. The reference in the glossary does not talk of trials, and is more of a Propensity Score (PS) overview. PS are often calculated a posteriori to adjust the data set. The use a priori to define which treatment patient is given seems curious, because their propensity to receive a particular treatment will be 1. A situation where treatment assignment is guided by underlying covariates, which are used to define treatment can easily be seen, but there is no need to formally invoke propensity scores to assign patients, although the cut-off value to dichotomise between treatment A or B may be driven by propensity scores. Taking off my methodology hat, I would be very nervous to do this. PS are based on exposure alone, and never on outcome, which makes them very useful in many pharmacoepidemiological situations. If you are going to have an assignment based on observed covariates, I would have thought that it should be based on outcomes. As an example, if we know men do better on treatment A and women on treatment B, then this makes sense to assign men preferentially to treatment A. – but if we know that in the past more men have proportionally been treated with treatment A and more women with treatment B, all this tells us is previous prescribing habits. Propensity Scores are just doing this for many, many covariates so don't solve a fundamental scientific problem. In summary, I would suggest to GetReal that before this is included in the glossary, a clear understanding of what this means, based on relevant references from the literature, how and why it may be applied, and the scientific question it is trying to answer, should be properly worked up. There is of course the potential I have misunderstood what they mean here, but as this is a glossary, it is perhaps the 1 document where clear understanding is key.	Incoporated: The points raised by the reviewer on the novelty of this term is correct. To avoid confusion, authors have decided to remove this term from the GetReal glossary until it could be discussed in further detail by the Glossary Working Group.
ZIN	**Relative efficacy**	Maybe you should also include the relative efficacy definition of the HPLF?	Incorporated: This suggestion has been implemented.

EMA	**RWD**	The issue here is that the definition of Efficacy and Effectiveness which is never going to be uniform between groups and I believe can never be solved. I appreciate the consortium has to make a definition somehow, for the purpose of the project, but you can end up tying yourself in knots if you are not careful. Efficacy: The extent to which a healthcare intervention produces a therapeutic effect as compared to a placebo intervention in a clinical trial setting. (Adapted from High Level Pharmaceutical Forum, 2008) The problem here is that we wish to define efficacy as something that is collected in a clinical trial , and then say that efficacy data is something collected outside of a clinical trial. With globally loose definitions of efficacy and effectiveness, if you take 2 different authors for your definitions you get tangled. This can of course be solved by talking about benefits and risks, or favourable and unfavourable effects, or, and perhaps easier for GetReal, simply being especially strict with your definitions, and every time you see the word anywhere in any document (including glossaries) that you take the time to consider it. As an aside the FDA Adaptive design Guidance states that FDA assess effectiveness from Phase 3 RCTs (lines 180 – 196 of the document linked in the spreadsheet), another reason to link to EMA documents! The US Code of federal regulations (in particular 314.126) also talks of effectiveness http://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfcfr/cfrsearch.cfm?fr=314.126 From a US FDA perspective, effectiveness is measured in RCTs, which may explain why there is sometimes a lack of linguistic harmony between stakeholders.	Incorporated: The reviewr makes a good point; controversy regarding efficacy and effectiveness should be resolved before a clear distinction between RWD and non-RWD can be made. The definitions offered by this glossary for efficacy and effectiveness have been checked to ensure that not confusion can occur as a result of their direct misinterpretation.

NICE	**RWD**	The original definition referred to observational studies as if they were different from registry studies and EHR. This has been clarified. Also, although not explicitly stated anywhere in the review above (apart from consultation comments on the ISPOR taskforce), we infer that RWD are primary (ie excluding systematic reviews and models etc). Systematic reviews and metanalyses are however included in many definitions of "comparative effectiveness research". NB: ABPI refer to other outcomes (treatment pathways, service models and patient preference/experience/compliance), which I have not included but implied that there are other outcomes collected. "An umbrella term for primary data on the effects of health interventions that are not collected in the context of conventional randomised controlled trials. Real world data (RWD) are collected both prospectively and retrospectively from observations of routine clinical practice. Data collected include, but are not limited to, clinical and economic outcomes, patient-reported outcomes (PRO) and health-related quality of life (HRQoL). RWD can be obtained from many sources including pragmatic clinical trials and observational studies (of which there are many types - see also "non-interventional/observational study")"	Partially incorporated: It would seem restrictive if we explicitly limit RWD to primary data collection; the synthesis and linking of different RWD sources should also be possible when generating RWE. Therefore the authors decided to not incoporate this. However, the authors did agree on the reviewers' extra examples of data collected, which have subsequently been added to the definition.
NICE	**RWE**	The ISPOR task force offer the following explanation which the IMI-GetReal Glossary Workgroup deemed too vague the project's purposes: “data conjures the idea of simple factual information, whereas evidence connotes the organization of the information to inform a conclusion or judgment. Evidence is generated according to a research plan and interpreted accordingly, whereas data is but one component of the research plan. Evidence is shaped, while data simply are raw materials and alone are noninformative.”	Not incorporated: Despite this clarification comment being relevant for general discussion, it does not lead to changes to the definition proposed. Therefore, it has not been incorporated in the glossary.

EMA	**RWS**	I have no idea why adaptive clinical trial is in here, and would strongly push for its deletion. Adaptive clinical trials are (almost always) randomised controlled clinical trials and so therefore fall under the umbrella category. I'd also suggest removing bridging study for similar reasons. These are often far from being 'real-world' the canonical example being RCTs in Japanese patients.	Incorporated: This argument has been repeated by several reviewers: adaptive clinical trials and bridging studies should not be part of RWS, since they primarily fall under the umberalla of RCT's. This can also be clearly seen in a publication by the CHMP with the title "Reflection paper on methodological issues in confirmatory clinical trials planned with an adaptive design". Therefore, adaptive clinical trials and bridging studies have been removed from the list of RWS's.
ZIN	**RWS**	RWS definition: revise term scientific studies; perhaps clinical studies	Incorporated: The authors agree that changing the term 'scientific studies' to 'clinical studies' provides a more accurate definition. Therefore, the proposed definition has been changed accordingly.
MHRA	**RWS**	The papers highlight that the various definitions are controversial and this is endorsed. I think that the definitions would benefit from some further reflection. In particular, the definitions mix the data source with the experimental design. The paper includes a pragmatic design as RWD, and more controversially (incorrectly) all adaptive clinical trials, many of which are very clearly RCTs. The border between an RCT (the definition of which doesn't necessitate tight inclusion and exclusion criteria), a large simple trial and a pragmatic trial can be blurred. A large simple trial collecting data with an unlicensed treatment, with informed consent, collecting data through CRFs that may or may not be collected in routine clinical practice would appear to be a CT, as defined by the relevant legislation. A large simple trial using EHRs may be RWD. Without considering both the data collection and the trial design it is not clear that definitive definitions can be reached. Furthermore, it is not clear how single arm clinical trials are classified in the definitions.	Incorporated: The points raised by the reviewer are quite relevant. To begin with, the definition of RWS has been updated to exclude adaptive clinical trials and bridging trials. Moreover, a clear distinction has been made between the definitions of RCT's (which have strict inclusion criteria), and large simple trials and PCT's that have broader inclusion and exclusion criteria. We agree that definitive definitions can only be achieved by considering aspects of data collection and trial design, to which effect we have included elements of these in the definitions for the following terms: RWD, RWS, non-interventional/observational studies, PCT's and RCT's.
NICE	**RWS**	Should consider listing more examples of study types (routine clinical practice, patient registries, hospital claims databases/administrative data, health surveys, electronic health records, medical chart reviews and post-marketing safety studies), but these are captured within "observational studies", which is cross referenced. Should also consider adding the following statement: "Real world studies do not typically include randomisation, but there are exceptions (for example, pragmatic trials)"	Partially Incorporated: The authors believe that the list of RWS suggested by the review conflate RWS types with RWD types. Therefore, alternative examples to RWS have been added to the list. On the other hand, the authors agree with the suggested statement regarding exceptional RWS's that include randomisation of trial subjects by design. The definition has thus been updated accordingly.

B. Comments received during internal review 2016

Organisation	Glossary Term/ Proposal New Term	Comments Received	Authors' Response
Merck	Adaptive Clinical Trial	The word "trials" is used when it should be "trial". The sentence "By means of an adaptive design, researchers therefore have the opportunity modify the trial procedures at different stages on the basis of analyzing interim data from study subjects." is redundant as well as gramatically incorrect and could be eliminated.	The authors believe that this sentence is needed for a clear explanation of a complicated concept. We therefore propose to keep the statement. The grammatical errors will be corrected accordingly.
HAS	Adaptive clinical trial	… By means of an adaptive design, researchers therefore have the opportunity to modify	The authors agree with the grammatic correction proposed and have incorporated it in the modified definition.
UCMU	Adaptive clinical trials	This is only one specific example of an adaptive trial. Also trials (very common) in which there is an interim analysis are adaptive trials as they may be terminated early or be prolonged. Those are also adaptations. Any trial for which the protocol allows (intended) changes to protocol elements such as type and number of patients, nature or dosage of the intervention or duration and size of the study are adaptive trials. The trial where the response to treatment leads to modification is just one example.	The authors agree with the statement that patient response to treatment is only one basis upon which adaptive trials can be modified. The emphasis in the current definition is on the fact that results from interim analyses are what guide researchers' choices to intentionally modify different aspects of trial protocol(s). This is reflected in the statement: "Such modifications may include, but are not limited to: selected drug dosages, sample size, and patient selection criteria. "
Bayer	Adaptive clinical trials	is there any workingpackage delivering more information regarding this method? Otherwise I would put more detail e.g. when to use it best etc. in here	According to our knowledge, work packages 2 and 3 are specifically addressing the use of adaptive clinical trials. Although more information is always welcome, we believe that it is beyond the scope of this definition to give information on when best to use such trials.
F.Hoffmann-La Roche	Administrative claims data	Why is it called "administrative"?	This was done to ensure that the name of this type of data coincides with the name used in the reference for this definition.
UMCG	Aggregate (Study) Data	Could add something like “derived from individual patient data using a (statistical) method of aggregation (e.g. calculating the average).”	The authors agree with this comment and have incorporated it in the modified definition.
F.Hoffmann-La Roche	Aggregate (Study) Data	Aggregated data should not only limit to clinical trials. Any summary data of a study should be called aggregated data.	The authors agree with this comment and have incorporated it in the modified definition.
UMCG	Aggregate Data Drug Information	Please add “ (ADDIS)” or “ADDIS: “ to the Term because the acronym is much more widely known	The authors agree with this comment and have incorporated it in the modified definition.

	System		
Amgen	Aggregate Data Drug Information System	Should "multiple-criteria decision making" be changed to "multiple-criteria decision analysis"?	The authors agree with this comment and have incorporated it in the modified definition.
UMCU	Alternative study design	I agree that the RCT is in general the preferred design, but not sure whether everyone will agree that all non-randomized trials should be considered alternative study designs. Shouldn't we label it 'non-randomized studies, and then say that 'This refers to all alternative study designs investigating ….'	The authors believe we should avoid dividing along the lines of randomisation as an aspect of trial design, since pragmatic clinical trials are alternative study designs which can include randomisation. Therefore, we propose to keep the current definition.
HAS	Alternative study design	Is it not too restrictive?	The authors believe that the nature of definition by exclusion (all but convnentional RCT) is inclusive, rather than exclusive. Therefore, we propose to keep the current definition.
Roche	Bayesian methods	I think the current proposal is not satisfactory and a bit misleading: - "the posterior may depend heavily on a prior as well as a likelihood ratio [...]": the posterior is fully determined by the prior and what is called "likelihood ratio" here; the sentence is not wrong but I do not think it puts the emphasize on the essence of Bayes theorem - "likelihood ratio": this is a bit confusing as it brings to mind the likelihood-ratio statistics, which is not what is ment here My proposal: Statistical methods based upon Bayes' Theorem, which shows how to update prior knowledge in the light of new data (i.e. posterior probability ∝ likelihood x prior probability). The prior probability distribution is based on subjective opinion, or on objective evidence, such as the results of previous research, or both and is explicitly included in subsequent calculations. Statistical inference is then based on suitable summaries from the posterior probability distribution.	The authors agree to the proposal made here. It is clearer and more concise. The proposed definition has been used to modify the previous one.
UMCU	Bias	Systematic (non-random) errors in values of parameters that are the object of measurement. --> this shoudl read study instead of measurement	The authors agree with this comment and have incorporated it in the modified definition.
UMCU	Case control study	Can also be prospective (see e.g. Grobbee & Hoes Clinical Epidemiology)	The authors agree with this comment and have incorporated it in the modified definition.
UMCU	Case control study	group without the outcome (controls) --> the controls have not experienced the outcome yet, but are not non-diseased by	The authors agree with this comment and have incorporated it in the modified definition.

		definition	
HAS	Case-control study	Not necessarily an exposure. Suggestion : characteristics of a group	The authors agree that "characteristics of the group" provides a broader interpretation. However, we propose to use "exposure" to remain consistent with the terminology used in the references.
HAS	Clinical endpoint/ outcome	A clinical endpoint is an aspect of a subject's clinical or health status that is measured to assess the benefit or harm of an intervention. A clinical endpoint describes a valid measure of clinical benefit due to intervention: the impact of the intervention on how a subject feels, functions and survives. It should be clinically relevant, sensitive (responsive to change) and is both accepted and used by physicians and patients. Clinical endpoints may be a clinical event (e.g. death, stroke, mortality,) a composite of several events, a measure of clinical status, a clinical parameter (e.g. pain, blood pressure) or health related quality of life (HRQoL). (Adapted from EunetHTA, 2013)	The authors agree to use the example of clinical status (blood pressure) and have modified the definition accordingly. On the other hand, we recommend stating that the clinical endpoint is clinically relevant, rather than should be.
HAS	Clinical Guideline/ Medical Guideline	A document produced with the intention to guide decisions and criteria (difficult to understand 'guide criteria') with regards to disease diagnosis, disease management, and treatment (could include detection and prevention) in the routine clinical setting. (Adapted from Council of Europe, 2001)	The authors agree with the comments and have incorporated them in the modified definition.
Merck	Clinical Trial	Instead of "investigational product(s)", it would be more accurate to say "intervention".	The authors agree with this comment and have incorporated it in the modified definition.
HAS	Clinical Trial	product sounds too restricted to pharmaceuticals. It can be a procedure, a combination of intervention, a strategy, etc.	The authors agree with this comment and have incorporated it in the modified definition.
HAS	Comparative effectiveness research	The conduct and/or synthesis of research comparing different benefits and harms of alternative interventions…	The authors agree with this comment and have incorporated it in the modified definition.
HAS	Comparator	I would suggest a very short version as : "intervention (including placebo) to which another one to be tested is compared". Reason : making it longer would require longer developpements to make it exhaustive.	Although we agree that a shorter version is more concise, we believe the current version leaves less room for confusion. We propose to keep the current definition.
HAS	Confounding bias	It's a bias. Do we want to articulate this one specifically?	Since the use of RWD inherently involves addressing varying forms of biases, we believe it best to identify and define different sources of bias.
Merck	Cross-sectional study	Instead of "particular" it would be more correct to say "specified".	The authors agree with this comment and have incorporated it in the modified definition.
F.Hoffmann	Direct treatment	What is observational trial? Does a phase II non-randomized but	We have changed "observational trial" to "observational study"

-La Roche	comparisons	controlled trial provide direct treatment comparison? In another word, does direct treatment comparison need randomization?	to avoid confusion. According to our knowledge, randomisation is not a requriement for direct treatment comparisons. This is reflected in the current definition.
HAS	Drug utilisation	I would not keep the second sentence.	The authors agree with this comment and have incorporated it in the modified definition.
UMCG	Drug utilisation studies	Should come after “Drug utilisation” (both alphabetically and logically)	The authors agree with this comment and will ensure the order is changed in the final version of the glossary
UMCU	Effectiveness studies	Term pragmatic studies is used, this should be changed to pragmatic trials	The authors agree with this comment and have incorporated it in the modified definition.
HAS	Effectiveness study	Consistensy issue with the entry of "real-world study" of this glossary?	The authors agree with this comment thus have used the same definition for RWS to define effectiveness studies.
HAS	Efficacy	just to make sure: do we want the comparator here to be always a placebo?	Indeed not, we do not want to limit the comparator in the definition to placebo. This has been incorporated into the modified defintion.
HAS	Efficacy study	See entry "efficacy" = the comparator here in efficacy study is not referred to as always a placebo.	The authors agree with this comment and have incorporated it in the modified definition of the term "efficacy".
HAS	Efficacy-effectiveness gap	just for confirmation: do we want to keep the word "randomised"?	According to the authors' knowledge, we do want to keep the word "randomised", in combination with "controlled" as well.
HAS	Electronic health/medical record (EHR/EMR)	Is the second sentence necessary?	In the authors' opinion, the second sentence should be kept to shed light on the fact that the items within an EHR are not universally consistent.
F.Hoffmann -La Roche	Electronic healthcare data	EHD can be used for effectiveness as well, can't it?	In the authors' opinion, EHR's can be used for providing information on effectiveness.
HAS	Electronic healthcare data	I would remove "which can be used... studies". It is only one possible use among other. Or do we want to stress this use in particular?	The authors agree with this comment and have incorporated it in the modified definition.
HAS	EMA Conditional marketing authorisation	(c) Orphan medicinal products)bracket is a typo.	The authors agree with this comment and have incorporated it in the modified definition.
HAS	(Clinical) Equipoise	An ethical criterion for recommending subject participation in clinical trials that states that a subject can only be referred to study participation if there is genuine collective (remove professional?) professional uncertainty from the medical community between the standard health intervention and the new tested intervention. (Adapted from Miller, 2003)	We propose to keep the term "professional" in the definition in order to specifiy the sort of uncertainty meant. We also propose to keep the current wording "regarding the best health intervention" as it is more concise.
Merck	General comment	The main thing is that someone should go through the entire glossary and give them a unified voice. Right now, some entries are definitions, some are descriptions, and a couple are quotes.	The authors agree with this comment and have incorporated it in the modified glossary.

		Many incorporate the term being defined into the definition. It looks amateurish. This is an excellent argument for employing a language editor.	
Merck	General comment	There is a typo on the third page of the section defining real world data. They give the acronym for the National Pharmaceutical Council as NOC	The authors agree with this comment and have incorporated it in the modified glossary.
Merck	General comment	Insert page numbers.	The authors agree with this comment and have incorporated it in the modified glossary.
Amgen	General comment	Typo, change 'heath' to 'health' (page 5)	The authors agree with this comment and have incorporated it in the modified glossary.
UCMU	Ideal vs. usual conditions	The “trialist” term for contextual factors is more commonly “extraneous factors:, i.e., factors that my impact treatment effects but are not resulting from the inherent pharmacological properties of the compound under investigation. This includes placebo effects. Extraneous factors are not necessarily (but can be) fixed or equalized, they may be made the same across groups by blinding. Randomization in the absence of blinding will not make extraneous effects the same (= comparable). Randomization does not serve this purpose but rather makes groups comparable for prognosis: this is essential for valid comparisons in both trials under ideal and usual conditions. A distinction should be made by extraneous factors that are excluded (ideal) or included (usual) in the estimate of the intervention effects (benefits and risks) and those aspects of the trial that speak on the intervention itself, such as the requirement to keep the intervention fixed over the duration of the study and changes are not allowed.	Though the authors agree with pionts mentioned here, further seperating contextual factors as suggested would prolong the definition. We therefore propose to keep the definition as it appears in the modified version.
Amgen	Ideal vs. usual conditions	“usual conditions” is not very clear whereas “routine clinical practice” sounds better	The authors agree but also see the value of including the term usual conditions.
UCMU	Ideal vs. usual conditions	‘The “ideal” condition refers to the situation where the setting is experimental and controlled, and where the “contextual factors” are fixed and equalized in the two (or more) therapeutic groups through randomization and/or standardization.’ Would add text in red as in randomized studies (i.e. pragmatic trials) one might also describe conditions to be ‘as usual’ (apart from randomization and IC).	The authors agree with this comment and have incorporated it in the modified definition.
F.Hoffmann -La Roche	Ideal vs. Usual conditions	Do we really need to include these terms in the GetReal glossary? They do not really sound like scientific terms	The authors agree but also see the value of including the term ideal vs. usual conditions, which clearly summarises a central theme to the use of RWD.

HAS	Ideal vs. Usual conditions	It might also be so, that the design is optimised to show the most benefit of the investigated intervention (drug) under investigation. Suggest to remove (drugs) ; I don't think the last sentence is necessary. If we keep it I trust we should not only focus on "drug", but broaden to 'intervention'	The authors agree with this comment and have incorporated it in the modified definition.
HAS	Indirect treatment comparisons	I'd remove 'treatment', even though most of the time treatments are compared.	In the authors' opinion, the term treatment should be kept in line with terminology provided by varying literature sources.
UMCG	Individual patient data (IPD) / Patient-Level Data	The definition is confusing since it assumes a meta-analysis setting, although that is not necessary to explain IPD. Could be improved by removing “in each included trial” and “in each study”, and replacing “This term refers to the availability of” by “Individual patient data are the”.	The authors agree with this comment and have incorporated it in the modified definition.
HAS	Individual patient data (IPD) / Patient-Level Data	.... refers to raw data.... (the availability is another matter)	The authors agree with this comment and have incorporated it in the modified definition.
HAS	Information bias	It's a bias. Do we want to define specifically some biases (and not other ones?)	Since the use of RWD inherently involves addressing varying forms of biases, we believe it best to identify and define different sources of bias.
Merck	Internal validity	Instead of "Whether)", it would be more accurate to say "The extent to which".	The authors agree with this comment and have incorporated it in the modified definition.
HAS	Longitudinal study	A study in which subjects are followed over time. I would stop here and not specify further.	Though this would be more concise, we believe a little extra information is necessary. Therefore, we propose to keep the defintion as currently stated.
HAS	Marketing authorisation	I'd remove 'manufacturing'; I'd remove the last sentence	The authors agree to removing the word "manufacturing" and have incorporated this in the modified defnition. On the other hand, we propose to keep the second sentence on basis for marketing authorisation.
UMCG	Markov model	First sentence refers to the “Markov property” without a definition. The second sentence contains the definition, but that is not clear from the text.	The authors agree with this comment. To avoid confusion, we have removed the statement related to the Markov property in the modified definition.
F.Hoffmann -La Roche	Markov model	A general comment: how are terms selected to be included in this glossary? Markov model is more related to cost-effectiveness model.	Terms are included in the glossary based upon a call for relevant terms from GetReal members. The authors agree that Markov modelling relates to cost-effectiveness analysis yet RWD is often use to define transition probabilities within such models.
F.Hoffmann -La Roche	Meta-analysis	Why not use the Cochrane definition? A proper meta-analysis is based on a systematic review. The Greenland defintion omits the SR.	The authors only intended to define the concept of NMA, rather than stipulate that a systematic review is a compulsory component of it. To our knowledge, not all institutions conduct a

			systematic review for every NMA performed. Therefore, we propose to keep the current definition
UMCG	Meta-regression	The current definition does not make clear it is related to meta-analysis - it does not even refer to evidence from multiple studies	The authors agree with this comment and have incorporated it in the modified definition.
Merck	Meta-regression	The term "IPD data" is redundant. Please spell out the full term: "individual patient data".	The authors agree with this comment and have incorporated it in the modified definition.
HAS	Misclassification bias	Same comment about specific biases.	Since the use of RWD inherently involves addressing varying forms of biases, we believe it best to identify and define different sources of bias.
F.Hoffmann -La Roche	Mixed treatment comparisons	Suggest to refer to the publication of ISPOR task force on indirect treatment comparisons good research practices: part I. https://www.ispor.org/workpaper/interpreting-indirect-treatment-comparison-and-network-meta-analysis-studies-for-decision-making.pdf	Although we agree that this reference is relevant, it is beyond the scope of the glossary to do so. The references cited relate to the origin of the definition. However, we recommend forwarding this comment to the Framework team, for whom this is highly relevant.
University of Ioannina	Mixed treatmentcomparisons	This term has been used in the literature as an alternative term for network meta-analysis. Also note that another term is also "multiple treatments meta-analysis". The term "network meta-analysis", however, is the term currently used more often . I suggest removing this definition altogether, and just refer to network meta-analysis. Also, add another entry, "multiple treatments meta-analysis", and also refer to network meta-analysis.	The authors agree with this comment and have removed the definition accordingly to avoid redundancy.
University of Ioannina	Multi-state transition model	I suggest deleting the phrase "This transition probability can be constant over time or not." This is not true for continuous time models, and anyway I think it is confusing. Rephrase "can be encoded in an (n×n) transition matrix" to "can be encoded in an (n×n) transition probability matrix" (this is needed because there is also a transition rates matrix)	The authors agree with the comment regarding confusion of the phrase but cannot find this phrase in the current definition. We have checked that the phrase is not in the modified definition.
UMCG	Multi-state transition model	“In a model comprising n states, for example, all possible transition probabilities can be encoded in an (n×n) transition matrix.” seems unnecessary and may not be strictly true in time-varying models - I suggest to remove it	The authors agree with this comment and have incorporated it in the modified definition.
University of Ioannina	Network meta-analysis	See previous comment; rephrase to "NMA methods take into account both Direct and Indirect Treatment Comparisons". Also I suggest rephrasing "both with or without the presence " to "either with or without the presence	The authors agree with this comment and have incorporated it in the modified definition.
Amgen	New term	Could a new term be added defining 'Causal relationship' as this is referred to in a few other definitions?	The authors have selected several terms (in red) to include in collaboration with WP 2 & 4.

Amgen	New term	Could a new term be added defining "Confounder" or 'Confounding factor' before the term "Confouding bias" as confounders are discussed in some definitions?	The authors have selected several terms (in red) to include in collaboration with WP 2 & 4.
Amgen	New term	Should a definition for 'Design-adjusted' be added as this is referred to in some of the evidence synthesis approaches available	The authors have selected several terms (in red) to include in collaboration with WP 2 & 4.
Amgen	New term	Should a definition for 'effect modifiers' be added as these play a key role in NMA?	The authors have selected several terms (in red) to include in collaboration with WP 2 & 4.
Amgen	New term	Should a defintion for 'interaction' be added as investigating possible interactions is key in understanding potential effect modifers for NMA?	The authors have selected several terms (in red) to include in collaboration with WP 2 & 4.
Amgen	New term	I wonder if some of the terms we refer to in WP4 need to be added to the glossary, for example: (multiple) competing interventions, primary endpoint(s), multiple endpoints, repeated measures, nodes, clinical significance, statistical significance, heterogeneity, homogeneity, statistical power, missing data, R (software package), systematic literature review, prediction models, mathematical models, multi-state models, generalized linear models, sensitivity analyses, model validation, adherence/non-adherence, precision, validity (internal vs external), network (of evidence), types of evidence – direct randomized, indirect randomised, direct observational, indirect observational, hierarchical models, acyclic graph, covariate, variable classification, time-varying confounders, down-weighting, prior, informative prior, non-informative prior,	The authors have selected several terms (in red) to include in collaboration with WP 2 & 4.
Merck	"Stochastic"	This term appears in the definition of Markov model. It should be defined independently.	The authors believe this term may be too specific/technical to warrant a separate entry in the glossary,
UCMU	Non-interventional study / Observational study	Maybe terms should be split. I agree with description of non-interventional study. However an observational study would also be a study where 'the investigator does not interfere with choice of the prescribed health intervention' but where (some) 'additional diagnostic or monitoring procedures' may be applied I would think.	The authors recognise the point made but refer to he following one on national legislation defining how diagnostic procedures are considered as interventions. Therefore, to avoid confusion, we propose to regard both terms as synonyms and define them together.
Lilly	Non-interventional study / Observational study	Regarding PW's comment - what extra monitoring/diagnostic procedures are allowed in an observation study very much depends on the country/legislation. In my opinion the current definition of both terms is pretty similar.	The authors agree with the comment made.
HAS	Non-interventional study / Observational	...interventions are conducted in the usual manner.(again not to focus totally on pharma)	The authors agree with this comment and have incorporated it in the modified definition.

	study		
UMCU	Observational data	Not necessarily patient data; could also be data on subjects who have no disease at all (e.g. primary prevention)	The authors agree with this comment and have incorporated it in the modified definition.
HAS	Observational data	"Data collected from patients populations" I'd remove the word "patients"; Sources of observational data may include:	The authors agree with this comment and have incorporated it in the modified definition.
HAS	Phase 2 (2a, 2b) trials	a and b not needed unless we explain the difference in the text?	The authors agree with this comment and have incorporated it in the modified definition.
HAS	Phase 3 (3a, 3b) trials	idem a and b	The authors agree with this comment and have incorporated it in the modified definition.
Merck	Phase II & Phase III trials	I'm not sure of the definitions for Phase II and Phase III trials. My understanding is that the purposed of Phase II is to find the best dose, and Phase III is to produce data to win regulatory approval. These goals aren't mentioned in the definitions. Am I misunderstanding the concepts?	The goals cited for the different phases are partially true; other main differences between both phases is the study population size. To avoid discussions, we did not specify the aims of the different phases.
F.Hoffmann -La Roche	Phase x trials	Why phase I trial is not included?	GetReal stakeholders did not deem it relevant to include phase I studies, namely because the use of real-world data from human subjects in phase 1 is limited.
Merck	PICO	The full term, "PICOTS" is now widely used. It should either replace PICO or be included separately.	The authors agree with this comment and have incorporated it in the modified definition.
Amgen	PICO	Has ths further evolved into PICOS or even PICOS-T?	The authors agree with this comment and have incorporated it in the modified definition.
F.Hoffmann -La Roche	PICO	Why not PICOS? S stands for study type.	The authors agree with this comment and have incorporated it in the modified definition.
HAS	PICO	We might wish to add T = time, temporality and make it a PICOT	The authors agree with this comment and have incorporated it in the modified definition.
F.Hoffmann -La Roche	Post-authorisation	I am not sure whether it has to be before product launch.	The authors agree with this comment and have incorporated it in the modified definition.
HAS	Post-authorisation	"The period after market authorization of a specific pharmaceutical/ medical device product" I would stop here.	The authors agree with this comment and have incorporated it in the modified definition.
HAS	Post-authorisation efficacy studies	I'd suggest to simplify the first sentence "Studies conducted after the intervention has been authorized (marketed)"	The authors agree with the comments presented. At the time of first drafting, the EMA had not produced a definition on this concept. To avoid mistakes, we have proposed to take the definition used in their guidance on PAES.
HAS	Post-authorisation evaluation	With regard to the previous entry, the focus here might be more to explain the word "evaluation". Do we want to refer to whatever is not a study ? (not sure what we want to mean)	In light of definitions present on PASS and PAES, we believe this term is redundant. Therefore, to avoid the confusion, we propose to remove it from the glossary.
HAS	Post-marketing	"This supercedes the post-authorisation phase." follows ?	The authors agree with this comment and have incorporated it in the modified definition.

HAS	Pragmatic clinical trial	I have a problem with what we want to mean by an "interventional follow-up"	Please refer to the footnote in the definition for the term Non-interventional/Observational studies for more information on whether follow-up procedures are interventional or not.
UCMU	Pragmatic clinical trials	....A study comparing ~~several~~ health interventions.....Note: not neccessarliy several! ...~~diverse~~ population representing patients intended to use the intervention in clinical practice.... Note: there is reduncancy in the description of the patients! (in the next paragraph) I don't think pragmatic trials are necessarily a subcategory of large and simple trials. They may well be small. It is rather the reverse: large and simple trials could be a subcategory of pragmatic trials	In relation to comment 1: pragmatic trials theoretically compare at least 2 interventions (new + standard of care). Therefore, we propose to keep the term "severa"l. In relation to comment 2: this is clarified in the following sentence. Therefore, to avoid redundancy, we propose to not add this statement in the first sentence. In relation to comment 3: we agree with this and have incorporated this by deleting the statement from the current definition and adding a corresponding statement in the definition for Large Simple Trials.
UCMU	Pragmatic clinical trials	# "A study comparing…measuring a broad range of health outcomes. " I am not sure if this is a defining characteristic of Pragmatic trials (compared to explanatory trials, which may have many outcomes measured). Maybe better the "measuring the full range of health gains" or something? # Related to the above point: "Pragmatic clinical trials are a sub-category of large simple trials." I am not sure whether 'large simple trials' (LST) are really conceptually different? Only they need to be large in size? In the definition of Large Simple Trials it seems that follow-up mimicking usual practice is differentiates LST from Pragmatic Trial "Additionally, LST's include follow-up that mimics normal clinical practice." I think LSTs are actually a sub-category of pragmatic trials (namely those that are large). I would just say: "Large simple trials are (by definition) Pragmatic trials." # "To ensure generalizability, pragmatic trials should represent the patients to whom the treatment will be applied and treatment strategies should follow usual practice as best possible." I would add that also the treatment strategy should be in line with usual practice (see addition in red).	In relation to comment 1: we believe this may a defining feature of PCT's. Please bear in mind that here we are not coparing the number of outcomes to those measured in RCT's. Therefore, we propose to keep the statement unchanged. In relation to comment 2: we agree with this and have incorporated this by deleting this statement from the current definition and adding a corresponding statement in the definition for Large Simple Trials. In relation to comment 3: although this sounds intuitive, it may be confusing; what if the new intervention is not conform standard treatment strategy or requries a change to the stanard treatment strategy? We propose to keep the definition as is to avoid such confusion.
UMCU	Propensity score	consider adding "the propensity score summarizes information of multiple confounders and can be used to adjsut for confounding by means of matching, wieghting or regression adjustment"	The authors agree with this comment and have incorporated it in the modified definition.
UMCU	Propensity-based trials	Is there randomization involved? If so, what population is used to derive PS from? And isn't this just a stratified randomized trial? If not, consider dropping the term trial, since it is not a RCT	The authors agree with this comment and propose to remove the term. It has previously caused confusion.

UMCU	Protopathic bias	This is a form of confounding, not information bias	The authors agree with this comment and have incorporated it in the modified definition.
HAS	Protopathic bias	Just learnt that word here :) good indeed, and same remark as previously : do we want to detail biaises?	Since the use of RWD inherently involves addressing varying forms of biases, we believe it best to identify and define different sources of bias.
HAS	Randomisation	Randomization is the process of assigning trials participants to intervention or control groups, using an element of chance to determine the assignments. Randomization is a tool for providing comparable participant groups regarding all measurable and unmeasurable characteristics, apart from the intervention ("active" or control). It ensures the initial comparability of patients between the groups. "By randomization, one hopes to equalize the distributions of confounding factors, whether they are known or unknown" [2].	The authors agree with this comment and have incorporated it in the modified definition.
HAS	Real World Data (RWD)	do we need the word "conventional"? asking because I'm not sure I know what an "unconventional RCT" would be (or if interesting for us it should be an entry here.	The authors agree with this comment and have incorporated it in the modified definition.
UCMU	Real-World Data	I have a problem with this definition as it does not include (randomized) pragmatic trials that may very well provide RWD superior to observational studies	We have decided to avoid mixing data sources and study designs in the defintion for RWD. Therefore, we will ensure that PCT's are clearly mentioned in the definition for RWS.
GSK	Real-World Data	An umbrella term for data regarding the effects of health interventions (e.g. benefit, risk, resource use, etc) that are collected in a way which reflects how the intervention would be used in everyday clinical practice and in a population of patients in whom the intervention would normally be used. It is different from that collected in the context of conventional randomised controlled trials. Instead, real world data (RWD) may be collected both prospectively and retrospectively from observations of routine clinical practice or from pragmatic clinical trials. Data collected include, but are not limited to, clinical and economic outcomes, patient reported outcomes (PRO) and health-related quality of life (HRQoL). RWD can be obtained from many sources including patient registries, electronic medical records, and observational studies. (Adapted from Garrisson, 2007)	The authors partially agree with the ammendments made here and have incorporated parts of them into the modified defintion. As stated above, we have decided to avoid mixing data sources and study designs in the defintion for RWD.

CTU Bern	Real-World Data	The description in Real-World Data and Real-World Studies need to be consistent: The definition on Real-World Data does not explicitly mention pragmatic clinical trials but instead gives the impression that only observational studies can generate Real-World Data whereas the definition of Real-World Studies does mention that pragmatic clinical trials are Real-World Studies. Also, the study types that are mentioned should be defined in the glossary for example, I do not know what a "drug utilization study" is.	We have adjusted the definition such that it leaves room for PCT's to qualify as RWD but have also decided to remove study designs from the definitions. Therefore, we will ensure that PCT's are clearly meantioned in the definition for RWS. In realtion to the second comment, drug utilisation studies are defined above in the glossary.
Merck	Real-World Data	An umbrella term for data regarding the effects of health interventions (e.g. benefit, risk, resource use, etc) that are not collected in the context of conventional randomised controlled trials.(Is there a clear definition of "conventional"? It may be opposed to pragmatic trials but this contradicts the notion of the continuum in trials (PRECIS tool). Every trial provides some real world data that complements observationaldata). Instead, real world data (RWD) is collected both prospectively and retrospectively from observations of routine clinical practice. Data collected include, but are not limited to, clinical and economic outcomes, patient reported outcomes (PRO) and health-related quality of life (HRQoL). RWD can be obtained from many sources including patient registries, electronic medical records, and observational studies. (Adapted from Garrisson, 2007)	To avoid confusion, we have decided to remove the term conventional. This has been incorporated into the new definition.
Bayer	Real-World Data	An umbrella term for data regarding the effects of health interventions (e.g. safety, effectiveness, benefit, risk, resource use, treatment patterns, patient pathways, costing information, etc). Real world data may be collected either in a way which reflects how the intervention would be used in everyday clinical practice and in a population of patients in whom the intervention would normally be used (e.g. pragmatic clinical trials) or may be directly derived from (in usual care) routinely collected data (e.g. by prospectively and retrospectively conducted observational trials). It is different from data collected in the context of conventional randomised controlled trials. Data collected include, but are not limited to, clinical and economic outcomes, patient reported outcomes (PRO) and health-related quality of life (HRQoL). RWD can be obtained from many sources including patient registries, electronic medical records, and observational studies. à mix of sources and methods, I would choose either one (Adapted from Garrisson, 2007)	The authors partially agree to all comments made here. We have adjusted the definition to only include sources of data, rather than sources and study designs.

UCMU	Real-World Studies	Here too, there is inconsistency as it is stated that RWS do not include randomization of subject while (randomized) pragmatic clinical trials are listed as an option	The authors believe that this issue is addressed by the statement: "Real world studies do not typically include randomisation of trial subjects, but there are exceptions (e.g. pragmatic clinical trials)."
Merck	Real-World Studies	This refers to all clinical studies investigating health interventions whose design does not follow the design of a conventional? randomised controlled clinical trial and aims to reflect health intervention effectiveness in routine clinical practice. Real world studies donot typically include randomisation of trial subjects, but there are exceptions (e.g. pragmatic clinical trials). For the purposes of GetReal, real-world studies include, but are not limited to, the following: pragmatic clinical trials, non-interventional/ observational studies, drug utilisation studies, post-authorisation efficacy/safety studies. RWS, by definition, generate RWD, which can subsequently be analysed and/or synthesised to produce RWE. (See also: "pragmatic clinical trial", "noninterventional/ observational study", and "drug utilisation study") (IMI-GetReal, 2014)	To avoid confusion, we have decided to remove the term conventional. This has been incorporated into the new definition.
Bayer	Real-World Studies	This refers to all clinical (remove word clinical) studies (à RW studies do not need to be clinical, or did I tremendously misunderstood anything??) investigating health interventions whose designaims to reflect health intervention effectiveness in routine clinical practice. Real world studies do not typically include randomisation of trial subjects, but there are exceptions (e.g. pragmatic clinical trials). For the purposes of GetReal, real-world studies include, but are not limited to, the following: pragmatic clinical trials, non-interventional/ observational studies, drug utilisation studies, post-authorisation efficacy/safety studies. RWS, by definition, generate RWD, which can subsequently be analysed and/or synthesised to produce RWE. (See also: "pragmatic clinical trial", "noninterventional/ observational study", and "drug utilisation study") (IMI-GetReal, 2014)	The authors agree with this comment and have incorporated it in the modified definition.
UMCU	Referral bias	Consider adding that becasue of referral outcomes may be more likley diagnosed and consequently the exposure appears to be related to the outcome (ie bias). This is a form of information bias (becasue there is selective misclassification of the outcome among those not referred)	Although this is a very valuable comment, it moves a little beyond the scope of defining referral bias. We thereore propose to keep the definition unchanged.
HAS	Referral bias	Idem general remark about specific biases	Since the use of RWD inherently involves addressing varying forms of biases, we believe it best to identify and define different

			sources of bias.
HAS	Registry	Maybe rephrase the sentence because a registry is not per se a study but the database occuring from the structured prospective recording of etc.	The authors agree with this comment and have incorporated it in the modified definition.
HAS	Relative effectiveness	wording swap words : alternative interventions	The authors agree with this comment and have incorporated it in the modified definition.
LASER	RELATIVE EFFECTIVENESS	"one or more intervention alternatives": we could specify that when the drug is compared to ONE other, it is comparative effectiveness and when it is "MORE", i.e. any other drug, it is relative effectiveness;	Althought this may be true, we believe that not all stakeholders make such a distinguishment. Therefore, we propose to keep the modified defnition in its current form.
LASER	Relative efficacy	same for "relative/comparative efficacy"	Althought this may be true, we believe that not all stakeholders make such a distinguishment. Therefore, we propose to keep the modified defnition in its current form.
F.Hoffmann -La Roche	Risk-sharing agreements	Is this in the remit of GetReal?	This is a good question. To the authors' knowledge, this term was recommended by several consoritum members, since these agreement involve the collection of RWD to evaluate eventual outcomes.
HAS	Safety	maybe remove the word "treatment"	We believe that removing the word treatment might lead to the definition lacking much context. We therefore propose to keep the definition in its current form
HAS	Scientific advice	do we need the word "test" ?	For the sake of accuracy in relation to the reference, we propose to keep the word "tests".
UMCU	selection bias	it's a systematic error, due to selective inclusion, or selective drop-out. Note that if selection depends on e.g. exposure only, or outcome only, there will not be any bias. It takes both in order to induce a bias (see e.g Hernan, Epidemiology 2004). Representativeness is not necessarily related to selection bias. Can have very unrepresentative sample, yet no selection bias (when using the definition by Hernan). Consider to explicitly distinguish this form generalizability, which are two different things.	The authors agree with this comment and have incorporated it in the modified definition.
HAS	Selection bias	Idem general comment on specific biases	Since the use of RWD inherently involves addressing varying forms of biases, we believe it best to identify and define different sources of bias.
HAS	Self-selection bias	Idem general comment on specific biases	Since the use of RWD inherently involves addressing varying forms of biases, we believe it best to identify and define different sources of bias.
Merck	Stakeholder	The word "thereof" is used when "of" is sufficient.	The authors agree with this comment and have incorporated it in the modified definition.

HAS	Stakeholder	Sounds a bit vague, but maybe I'm not enough aware of the project	To clarify, the term was included in order to guide the development of early deliverables related to stakeholder identification outside the GetReal consortium.
HAS	Standard of care	do we want to keep "at a minimal level" ?	The authors agree with this comment and have incorporated it in the modified definition.
F.Hoffmann -La Roche	Sub-group analysis	The definition is limited to clinical trials. For GetReal, it is probably more relevant for RWD.	The authors agree with this comment. In an attempt to make the definition less RCT-centric, we have replaced the term "trial" with "study".
HAS	Sub-group analysis	Do we want to stress how important it is to plan subgroups analysis in advance in protocols to get valid statistical data?	We believe that this would be beyond the scope of this definition, since that statement is more of a recommendation. Therefore we propose not to do so here.
F.Hoffmann -La Roche	Time horizon	Is this in the remit of GetReal?	This is a good question. Although cost-effectiveness is not within the GetReal remit, RWD inevitably features in PE aspects. Therefore, it is worthwhile highlighting certain aspects. To clarify, this term was recommended by consortium members.
University of Ioannina	Transitivity assumption	Please add the following "Transitivity assumption is a fundamental assumption of Indirect Treatment Comparisons and network meta-analysis. It implies that we can combine the direct evidence from A vs. C and B vs. C studies to learn (indirectly) about the comparison A vs. B". Add as a reference the recent GetReal paper on NMA: http://onlinelibrary.wiley.com/doi/10.1002/jrsm.1195/abstract	The authors agree with this comment and have incorporated it in the modified definition.
F.Hoffmann -La Roche	Transitivity assumption	HRQoL should be included as well. It is not only for safety, efficay or effectiveness.	The authors agree with this comment and have incorporated it in the modified definition.
HAS	Transitivity assumption	If the safety, efficacy or effectiveness of health intervention A is quantitatively related to that of B, and that of B to C, then A is related to C for all A, B and C in the domain (typo) of the relation. (Athanasiou, 2011) (see also:"mixed treatment comparisons")	The authors agree with this comment and have incorporated it in the modified definition.
HAS	Work package	Not wure why we would need this one in particular (because if we do, why not task, deliverable etc. etc.) ?	We believe this term is needed for external consortium members who are not familiar with the use of work packages in European projects. To clarify, this terms was recommended by consortium members.
2nd Round 10.06.2016			

UMC	Ideal vs. Usual conditions	The context in which a treatment is being studied consists of the mode of administration, side-effects and their treatment, diet, auxiliary care, associated treatments, etc. All the characteristics of this context can be called contextual factors or extraneous factors, i.e. characteristics of the population (e.g. genetics, behaviour towards the drug, co-morbidities), the healthcare delivery system (e.g. physician behaviour of prescription, guidelines etc.), the actual use of the drug (compliance, health beliefs, co-medication, etc.).	The authors agree with this comment and have incorporated it in the modified definition.
UMC		The ideal condition refers to the situation where the setting is experimental and controlled, and where the contextual factors are fixed and equalized in the two (or more) therapeutic groups through randomization, blinding and/or standardisation. For instance, the intervention status will be set at the beginning of the trial, and emphasis is put on not to change intervention status overtime, etc. Often, the design is optimised to show the most benefit of the investigated interventions.	The authors agree with this comment and have incorporated it in the modified definition.
UMC		The usual condition (or routine clinical practice) refers to what really happens when an intervention is prescribed by a physician, to a patient. In routine clinical practice, the contextual factors are not fixed: they vary according to the physician's usual habits, the disease severity, patients' preferences, etc.	The authors agree with this comment and have incorporated it in the modified definition.
UMC	Pragmatic clinical trial	A study comparing several health interventions among a randomised, diverse population representing clinical practice, and measuring a broad range of clinically relevant health outcomes. To ensure generalizability, pragmatic trials should represent the intended patients to whom the treatment will be applied and the treatment strategy should be in line with clinical practice as best as possible. For instance, inclusion criteria would be broad (e.g. allowing co-morbidity, co-medication, wider age range, etc.), the follow-up would not be (or not much) interventional and allowing for treatment switching etc. (See also "large simple trials" and "real-world studies") (Adapted from Schwartz, 1967, Tunis, 2003 & Roland, 1998)	Although the authors acknowledge the value of the additions made here, we propose to keep the definition as seen in the current version. Namely, we propose to keep the terms broad range of outcomes (to contrast with trials with highly selective outcomes). We propose to avoid adding the statement on generalisability, since it could be argued that generalisability is not dependent on clinical practice but rather mainly on the study population. The minor edits have been incorporated into the modified version.
UMC	Real World Data (RWD)	An umbrella term for data regarding the effects of health interventions (e.g. safety, effectiveness, benefit, risk, **[similar to effectiveness and safety?]** resource use, etc) that are collected in the context of routine clinical practice are not collected in the context of randomised controlled clinical trials.[**Not true if it is a pragmatic trial, contradicts the definition of RWS**] Instead,	The authors agree with suggestion in between brackets and have incorporated it into The modified version. On the other hand, we propose avoiding the modification in blue; data collected from PCT's is considered RWD, although it does not represent the routine clinical context. To avoid confusion, we draw a line between randomised controlled clinical trials (RCT's)

		RWD can either be primary research data collected in a manner which reflects how interventions would be used in routine clinical practice or secondary research data derived from routinely collected data. Data collected include, but are not limited to, clinical and economic outcomes, patient-reported outcomes (PRO) and health-related quality of life (HRQoL). RWD can be obtained from many sources including patient registries, electronic medical records, and claims databases. (Adapted from Garrisson, 2007 and GetReal, 2016)	and PCT's in the glossary.
UMC	Real World Study (RWS)	Studies investigating health interventions whose design does not follow the design of a randomised controlled clinical trial and aim **[Is/seems contradictory to next statement and is not needed]** aiming to reflect intervention effectiveness in routine clinical practice. Real world studies do not typically include randomisation of trial subjects, but there are exceptions (e.g. pragmatic clinical trials). For the purposes of GetReal, real-world studies include, but are not limited to, the following: pragmatic clinical trials, non-interventional/ observational studies, drug utilisation studies, post-authorisation efficacy/safety studies. RWS, by definition, generate RWD, which can subsequently be analysed and/or synthesised to produce RWE. (See also: "effectiveness study", "drug utilisation study", "pragmatic clinical trial" and "non-interventional/ observational study") (GetReal, 2016)	The authors believe we should avoid this deletion; it is important to draw the line between RCT's and non-RCT's one of the conditions for RWS's. Please also refer to the comment above on distinguishing between RCT's and PCT's.
Novo Nordisk A/S & University Medical Center Utrecht (UMCU)	Effectiveness study	I suggest to add the following in **bold**: "Effectiveness studies (also referred to as real-world studies) measure the effects associated with a health intervention when prescribed in routine clinical practice and usual prescription conditions **with the objective to estimate effectiveness**."	Although a valuable suggestion, we have avoided using elaborate formatting (e.g. bold font, italics, underlining). We propose to keep it this way for consistency.
Novo Nordisk A/S & University Medical Center Utrecht (UMCU)	Efficacy-effectiveness gap	Does it have to be *observed*? Could be changed to: "Discrepancy between efficacy and effectiveness and may be observed when comparing effects of a health intervention in routine clinical practice with effects demonstrated in randomised controlled clinical trials."	We believe it would be best to include the word observed, both for the sake of consistency with the definition source and to imply that the efficacy-effectiveness is not per se prsent for all interventions.

Novo Nordisk A/S & University Medical Center Utrecht (UMCU)	Efficacy study	similar as to effectiveness studies, the objective could be added. Also, we cannot be certain of a causal relation. Suggested changes marked with **bold**: "Efficacy studies, often referred to as explanatory studies, **have the objective to estimate efficacy, forexample by measuring** the effect of a drug in under highly controlled conditions (i.e. in the setting of randomised controlled clinical trials (RCT's). **Ultimately, t**hey serve to prove the causal relationship between an intervention and an outcome, thus answering the question "can it work?" in an ideal world."	Although a valuable suggestion, we have avoided using elaborate formatting (e.g. bold font, italics, underlining). We propose to keep it this way for consistency.
Novo Nordisk A/S & University Medical Center Utrecht (UMCU)	Randomisation	randomization only insures comparable treatment groups at the time of randomization. It may be so, that treatment groups start different behaviour after randomization. Suggested change marked with **bold:** "Randomization is the process of assigning trials participants to treatment or control groups, using an element of chance to determine the assignments. Randomization is a tool for providing comparable participant groups regarding all measurable and unmeasurable characteristics, apart from the treatment (“active” or control). It ensures the initial comparability of patients between the treatment groups **at the time of randomization**. “By randomization, one hopes to equalize the distributions of confounding factors, whether they are known or unknown” [2]. "	Although a valuable addition, we believe it should be clear to readers from the defintiion context that comparability is guaranteed at the time of randomisation. We proopse thus to avoid this addition.
UMCU	Pragmatic clinical trial	I don’t think these trials are by design larger. Only in specific circumstances when compared to an explanatory trial addressing the same intervention effect. Suggest to delete.	The authors agree with the comment made. We ensured that the definition for PCT does not metnion trial size. We also only refer to large simple trials for cross-referencing purposes within the glossary.
Janssen en Janssen	General comment	· In the pre-amble, I think it would be important to state that a stark effectiveness / efficacy definition isn’t realistic and that models like PRECIS suggest a continuum, where a study can be more/or less pragmatic along a number of different dimensions · Not sure if any comment is required on cluster randomisation, or is implicit that not necessarily at patient level · Glossary does not have a comprehensive list of important biases (e.g. no immortal time bias)	In response to the first comment, the authors agree with this and refer the reader to page 10 where the third paragraph discusses this. In response to the second comment, we indeed imply that randomisation neet not necessarily occur at patient level. In reponse to the third comment, it would be a daunting task to display a comprehensive list of biases; instead, we mentioned the three main forms of biases (information, selection, confounding) and recommended other GetReal members to name the remaining necessary specific forms.

Janssen en Janssen	Protopathic bias	· Definition of protopathic bias is confusing or plain wrong by my understanding. I've givent he ENCePP definition below, which reflects my understanding: Protopathic bias arises when the initiation of a drug (exposure) occurs in response to a symptom of the (at this point undiagnosed) disease under study (outcome). For example, use of analgesics in response to pain caused by an undiagnosed tumour might lead to the erroneous conclusion that the analgesic caused the tumour. Protopathic bias thus reflects a reversal of cause and effect	The authors agree with this comment and have incorporated it in the modified definition.
UMCU	Aggregate (Study) Data	Summary data of the results of a study (e.g. on subjects in a trial), as opposed to Individual Patient Data which represents the raw data of each study subject. Summary data is derived from individual patient data using a (statistical) method of aggregation. An example of aggregate data are the estimated coefficients from a regression model (Adapted from Lyman, 2005)	Although a valid addition, we propose to keep the definition as it appear for brevity.
UMCU	Bayesian methods	Statistical methods that are based on Bayes' Theorem, which shows how to update prior knowledge in the light of new data (i.e. posterior probability ∝ likelihood x prior probability). Prior knowledge is defined in terms of probability distributions, and can be based on subjective opinion, on objective evidence, such as the results of previous research, or both. Statistical inference is then based on suitable summaries from the posterior probability distribution. (Adapted from Rothman, 2008 and HTA glossary)	The authors agree with this comment and have incorporated it in the modified definition.
UMCU	Bias	This is on the primary study level. We have more types of bias in a meta-analysis.	The authors agree with this comment and hope the necessary biases in meta-analyses have been covered in the updated glossary.
UMCU	Clinical significance	The practical importance of and benefit from a treatment. It describes whether and to which extend the intervention has a real genuine, palpable, noticeable effect on daily life. Clinical significance is usually informed by effect size and statistical significance (Adapted from Kazdin, 1999 and Redmond, 2001).	The authors agree with this comment and have incorporated it in the modified definition.
UMCU	Cross-design evidence synthesis	Cross-design synthesis is a specific type of synthesis. The current definition applies to more general form of synthesis	The authors acknowledge this comment. However, for the purposes of this glossary, the intention was to define the general concept of cross-design synthesis, rather than its specific meaning in meta-analysis. We terefore propose to keep the definition as it currently is, with the possibility of adding the more specific one later.

UMCU	External validity/ Generalisability/ Applicability	External validity = whether results from a certain study reasonably remain valid in a new group of patients Generalisability = whether results from a certain study remain valid across different populations, settings, etc. Applicability = refers to whether the results (or models) can be applied in routine care	The authors agree with the comment made on seperating the concept applicability from this list. However, in the context of GetReal work, the authors believe that external validity and generalisability are synonymous: the new group of patients and different populations refer to the real-world setting (routine clinical practice). We therefore propose to avoid this distinguishment in this glossary.
UMCU	Hierarchichal model	I think this is not clear enough. In Bayesian statistics, parameters are also drawn from a certain distribution: Multilevel (hierarchical) modeling is a generalization of linear and generalized linear modeling in which regression coefficients are themselves given a model, whose parameters are also estimated from data. For example, in a random-effects meta-analysis, the relative treatment effect parameters in the individual studies are drawn from a random effects distribution, which allows for statistical heterogeneity between studies.	The authors agree with this comment and have incorporated it in the modified definition.
UMCU	Internal validity	The extent to which study results (e.g. treatment effects, model predictions) are valid in new subjects from the same population or setting from which the results were obtained.	Although this is a valid, thorough definition, we beleive it is synonymous with the current version. Therefore, we propose to keep the current definition as it is.
UMCU	Non-informative prior	The term “non-informative priors” is often used when researchers want to conduct a Bayesian analysis and obtain results that are dominated by the data (and not influenced by the prior distributions). It is, however, increasingly acknowledged that non-informative may not exist, since all priors contain some information. Although it is still possible to implement “vague” or “diffuse” priors, resulting models may not be identifiable or difficult to estimate. For this reason, it is often preferred to implement weakly informative prior distributions	The authors agree with this comment and have partially incorporated it in the modified definition.
UMCU	Predictive modelling	The activity of developing, validating or adapting risk prediction models. These models combine multiple subject characteristics to provide estimates of absolute (rather than relative) risk in individual subjects	Although this is a valid definition, it may be too limited to risk prediction in its current form. We therefore propose to incorporate this partially, rather than fully.
UMCU	Propensity scoring	The conditional probability of assignment to a particular intervention given the observed data. It is often used to summarize information of multiple confounders when estimating treatment effects from non-randomized data sources. (Adapted from Rosenbaum, 1981)	Although the additions made here are more intuitive, we propose to keep the current definitions as it is.
UMCU	Relative efficacy	Does this term [ideal conditions] matter? It refers to the conditions that were implemented in the trial. These conditions are somehow	This is a good question. Due to the fact that multiple stakeholders distinguish between the ideal conditions of a trial,

		artificial but it does not matter whether they represent some ideal. As long as they are the same for every subject, the conditions are met.	in contrast to the usual conditions of the real-world, we believe that this term is important.
UMCU	Statistical significance	Check definition P-value, as statistical significance refers to this. Note that interpretation of P-value is complex and certainly does not relate a probability of something being true. Instead, the meaning of a P-value is conditional, i.e. given that we assume that …., what is the chance that … This is completely different from "What is the probability that these findings are true?"]	The authors agree with the comments made. Therefore, we propose to ensure that the term "probably true" remain in the current form of the definition and the non-specific term "hypothesis testing" be used, rather than specific reference being made to use of the P-value.