



EUROPEAN MEDICINES AGENCY
SCIENCE MEDICINES HEALTH

9 December 2013
EMA/717922/2013 Rev. 1

Highlights from the “Workshop on methods for efficacy studies in the everyday practice”

European Medicines Agency, 24-25 October 2013

On 24 and 25 October 2013, the European Medicines Agency organised a “Workshop on methods for efficacy studies in the everyday practice” as a preparatory step to the development of scientific guidance on post-authorisation efficacy studies (PAES), which the European Medicines Agency will develop according to a mandate set out in the new pharmacovigilance legislation and subsequent to any outputs of the European Commission on the situations in which such studies may be required. The objectives of the workshop were to understand strengths and weaknesses of different design options to study efficacy in the conditions of the everyday medical practice, to issue recommendations on best use of methods to account for bias and confounding and to identify needs for the improvement of methods in the field of efficacy studies. Five main topics were addressed by invited experts in working groups: pragmatic trials, observational studies, registries, the use of electronic health records for pragmatic trials, and methods to control for confounding.

Pragmatic trials

Pragmatic trials are not as well codified as Phase III trials but every step of a controlled trial can be relaxed to make it pragmatic. Pragmatic trials are randomised trials where patient follow-up is akin to an observational study. The expert group discussed the range of interventions where pragmatic trials are particularly appropriate (as opposed to Phase III trials which answer questions on a specific molecule) and where randomisation is useful (as opposed to observation). It is particularly important to ensure external validity and generalizability, in particular because the need for consent can have a major impact on the population enrolled and results should be translated into a real-world setting. For ethical reasons, pragmatic trials are less appropriate where efficacy is uncertain. Additional key issues are the impact of the Clinical Trials Directive on the feasibility of pragmatic trials in practice and the need to ensure that the initial diagnosis by the GP can be relied upon as pragmatic trials tend not to include confirmatory tests.

There are many relevant design options for the design of pragmatic trials and the expert group discussed several ones. It concluded that the current paradigm of clinical trials holds but more Baskerville type designs could be used where patients determine how long they stay in any arm of the trial before switching or withdrawing. In addition further exploration of ‘Latin-square’ designs and



minimization of design elements by remote electronic follow-up for events of interest were considered useful.

Cluster randomised trials were considered a useful tool for the measurement of effectiveness in particular when investigating rare outcomes. The expert group considered that this method needs to be further developed and could be more frequently utilised. Stepped-wedge designs introducing a drug in one area first and then randomising it sequentially in other areas could also be useful.

The expert group highlighted that some of the control mechanisms in place in phase III trials are relaxed in pragmatic trials. Investigators should therefore report quality metrics, i.e. measures quantifying to what extent and which control mechanisms were relaxed. Adherence to the CONSORT statement was considered crucial for reliable reporting of results.

Observational studies

The expert group agreed that, except in specific circumstances, the aim of observational post-authorisation efficacy studies is not to demonstrate the efficacy of a drug: this is the role of randomised clinical trials (RCTs). Once efficacy has been demonstrated, observational studies are useful to study effect modifiers, namely variables that may influence the level of efficacy of the drug and have been controlled for in the RCTs. Examples of effect modifiers are drug doses that have been prescribed by physicians or taken by patients and differ from those used in the RCT, treatment schedules, patient sub-groups defined by factors such as age, co-morbidities and use of concomitant drugs, and factors related to a defined country or health care system. With the possibility to use historical data, observational studies are also useful when a rapid answer to an efficacy question is needed or when the comparator drug used as reference changes over time.

The expert group discussed several aspects of the design of observational efficacy studies that would increase the confidence in the reliability of results. In such studies, investigators do not interfere with the allocation of treatments and physicians and patients know which treatment has been prescribed. Bias may occur if the assessment of outcomes cannot be blinded, for example in studies with secondary use of data. Observational efficacy studies therefore require exposures and outcomes with a high specificity which can be measured with objective criteria. Ideally, the same outcomes as those used to prove efficacy in RCTs should be used. Ability of the study to correctly measure the relevant confounding factors and effect modifiers is another element to be considered to be confident in its results.

The focus of observational efficacy studies on the assessment of effect modifiers influences the data analysis and the choice of statistical parameters. An intention-to-treat analysis should always be performed but, depending on the aim of the study, alternative analyses may be needed, for ex. if the objective is to determine the level of efficacy depending on patients' compliance to treatment.

More generally, the credibility of results of observational studies could be increased with documented use of strict standards of quality control.

Registries

Registries allow collecting data on patients diagnosed with a certain disease or treated with a certain drug in a defined setting. Established registries provide an opportunity to assess patient outcomes, including effectiveness. The expert group considered that registries can be used when data on exposure, outcomes and confounders are available in the registry, or supplementary data collection or

linkage are feasible, as these situations allow systematic data collection on hard clinical outcomes. Registries are particularly important for understanding real-world treatment use and off-label use.

Registries allow for a wide variety of observational study design options including prospective cohort studies with nested case-control analysis, inception cohorts, retrospective cohorts for events with short induction times, natural history studies, cohort studies with internal comparators, linkage and/or supplementary data collection, and case-control studies. Disease registries can also be used as a source of subjects for RCTs, in particular when the medicine is already marketed, when the follow-up is adequate for study purposes and if there is interest in a non-randomised comparator group. It is important to point out that registry infrastructure does not support randomisation at the point of care.

Registries allow for large numbers of subjects to be followed, which is an asset for rare safety events and for studying treatment heterogeneity. They are most appropriate when information is not available in other settings, when long-term follow-up is available, for events that would not come to the attention of traditional care providers or health care systems, when data on patient-reported information is needed and for rare conditions.

Amongst their limitations, the expert group highlighted situations when the disease or exposure classification is not specific enough, when follow-up is not possible or available, for comparative effectiveness studies where reasonably unbiased comparators cannot be identified, when the treatment of interest is not used in patients or subgroups of interest, and where there is little awareness about the disease in a given country.

As for any other epidemiological source of data, data quality is key for the success of the research. Measures to improve the quality of data, the validity of studies and the usefulness of results from registries include using common terminologies and data dictionaries/definitions, quality control of laboratory and measurements data and standards for collection of patient-reported information. Moreover, in terms of data interpretability, it is important to describe the representativeness and generalizability of a registry, and whether it covers the relevant patients and periods of interest.

Use of electronic health records for pragmatic trials

The expert group noted that using clinical practice databases to facilitate the conduct of randomised clinical trials is a new area and that significant challenges need to be resolved.

For some RCTs, the potential benefits of electronic health records (EHR) databases are significant, in particular those in which the outcomes are clinically important acute events (e.g. death and onset of new disease) that are likely to be well recorded. In addition, long-term low-cost follow-up is possible and patients with rare diseases can be identified by automated screening of a large population.

Particular concerns are that the quality and completeness of data must be sufficient to ensure the trial findings are robust and that adequate patient consent procedures must be introduced. It was also noted that important variations exist between databases and consequently that the implementation of RCT processes might also vary. Administrative requirements, coding conventions, quality of data, ability to link to additional data sources and the ability to provide further clinical details on request are all likely to be specific to the EHR system. Moreover, interpretation of the data often requires significant expertise.

Some advances are required if the use of EHR databases in RCTs is to become a routine procedure. Harmonisation of legal requirements and administrative procedures across databases would be desirable, possibly with groups of databases forming networks giving access to very large and diverse populations of patients. Development of software to allow and record randomisation and, in some

cases, collect specific tests and variables whilst ensuring the trials caused minimal additional work for database users would also be required. In addition, moves toward higher data quality and better coding procedures should be continued.

The consensus was that for certain kinds of large pragmatic trials, EHR databases are an important and promising resource.

Methods to control for confounding

In observational studies of drug effects, confounding by indication and channelling of treatments are amongst the main challenges when evaluating efficacy endpoints. From the outset, the expert group acknowledged that for well-measured confounders there is little difference in results between different methods used to address confounding. The expert group focused therefore on methods to handle unmeasured and mismeasured confounding. Strengths, weaknesses and suitability of the following options in different situations were discussed, including the need for additional data collection, use of self-controlled designs, use of instrumental variable analysis, proxy adjustment via score methods such as propensity scores and disease risk scores and use of active comparators.

When it is possible to identify a subset of the observational study population that mimics the RCT population, confidence in the overall study results may be increased if the same results as the RCT results are found in that subset of the population. Therefore, some phase III studies should ideally be designed in anticipation of such post-authorisation study designs and analyses. Similarly, if negative results are obtained for outcomes where negative results are expected then other positive results from the same study can also be considered more reliable. Finally, the importance of sensitivity analyses to test the robustness of study results was emphasised.

The following issues that make analyses particularly challenging were identified: strong adherence to treatment guidelines makes allocation of treatment less random, prescribing tends to be highly selective immediately after marketing, in which case development of disease risk scores pre-marketing may be useful, and many chronic conditions have time-varying exposures, in which case marginal structural modelling may be appropriate.

In summary, whilst the expert group recognised and discussed a variety of challenges related to confounding and channelling, it also identified a range of potential ways to address these as well as scenarios that would enhance confidence in post authorisation efficacy study results.