Observational studies: going beyond the boundaries of randomized controlled trials

Wenying Yang, Alexey Zilov, Pradana Soewondo, Ole Molskov Bech, Fawzia Sekkal, Philip D. Home

ARTICLE INFO

The term observational study describes a wide range of study designs including prospective and retrospective cohort studies, case-control studies, and cross-sectional studies, a defining feature of which is that any intervention studied is determined by clinical practice and not the protocol. Data from large, prospective observational studies provide information about the safety and efficacy of medicines in daily clinical use. Such observational studies are generally carried out once a medicine has received approval from regulatory agencies. Observational trials have inherent limitations in terms of their susceptibility to bias and confounding, restricting their ability to define causality. However, their strengths include that they reflect daily clinical practice more closely than randomized controlled trials (RCTs), both in terms of the heterogeneous patient populations that are included, and the medical interventions that they receive. Therefore, observational trials can provide clinically relevant information that is not necessarily provided by RCTs. Given the limitations of an observational study approach, it is important to optimize their study design to maximize their validity, and thus, in particular, known causes of bias and confounding should be measured. Medical investigators, health authorities, and the pharmaceutical industry all have important roles to play in designing, approving, and performing observational studies.

© 2010 Elsevier Ireland Ltd. All rights reserved.

1. Introduction

While appropriately designed randomized controlled trials (RCTs) are rightly judged to provide the highest quality of efficacy and safety data for evidence-based medicine decision making, they do have some limitations; the highly selected patient populations included in many RCTs may not be representative of features of the typical patient. For example, patients may be younger, may have less severe disease, and usually lack the range of common comorbidities that clinicians see in their patients. It is unfortunately common for RCTs to compare new treatments with placebo rather than current best treatment options, making it difficult for clinicians to differentiate between treatments. In addition, because RCTs are expensive to run, there are limits in terms
of the number of participants that can be included and the length of follow-up in the study. As a result, RCTs may miss quite common adverse events, or those events that develop over longer timeframes, and usually provide poor estimates of their frequency. Accordingly, although RCTs are important in providing estimates of the effectiveness and common adverse events of therapies, other types of study are necessary to provide important additional information that is relevant to day-to-day prescribing decisions in the clinic.

The limitations of RCTs outlined above have led to a call for more clinically relevant approaches to supplement the important information that RCTs provide [1–5]. In the US, this approach has been termed comparative effectiveness research (CER) [1]. The US Institute of Medicine (IOM) defines CER as: “the generation and synthesis of evidence that compares the benefits and harms of alternative methods to prevent, diagnose, treat and monitor a clinical condition, or improve the delivery of care. The purpose of CER is to assist the consumer, clinicians, purchasers, and policy makers to make informed decisions that will improve health care at both the individual and population levels” [1]. Inherent in this approach is a focus on how treatments perform in day-to-day clinical practice in typical patient populations and in comparison with other effective treatments rather than placebo [1]. In addition to examining how RCTs can be made more relevant for CER, an integral part of this approach is to supplement evidence from RCTs with information from systematic reviews and observational research. Accordingly, this paper will examine studies that can be classified as observational, and discuss how these types of studies can be used to provide information on how drugs perform in real-life clinical practice, how observational trials compare with RCTs, and what aspects of the observational study design are important to maximize validity. In addition, the roles, responsibilities, and expectations of investigators, health authorities, and industry in designing, approving, and performing observational studies will be examined.

2. Methods

To inform this review, a literature search of the MEDLINE/PubMed databases was performed using the terms observational study or studies, comparative effectiveness research alone or in conjunction with the terms insulin or diabetes. Relevant studies were identified and reviewed. In addition, non-peer-reviewed material was identified by internet searches.

3. What is an observational study?

The term observational study covers a wide range of study designs, a common feature of which is that they are non-interventional, in the sense that the study protocol does not determine the precise features of any therapy given to the participants in the study. There are three main types of observational studies: cohort (prospective and retrospective), case control (nested case control), and cross-sectional.

3.1. Cohort studies

These studies take a group of individuals who share a common characteristic (for example, a particular treatment, date of birth, or who have a particular disease), and who are then followed up at specified intervals to determine outcomes. Cohort studies are useful for identifying the incidence of a particular outcome over time. The Framingham study is an example of a prospective cohort study in that it followed up a large cohort of individuals over many years to identify potential risk factors for cardiovascular diseases [6]. A retrospective cohort study uses data that may have been collected for another purpose to answer a particular question, or simply to provide a resource for future studies. For example, a recent retrospective cohort study used a UK general practice research database to investigate the risk of incident myocardial infarction, congestive heart failure, and all-cause mortality associated with oral glucose-lowering drugs in 91,000 people with diabetes [7].

3.2. Case-control studies

These are usually retrospective and aim to identify predictors of a particular outcome [8]. To do this, a population with a particular outcome is compared with a matched population without the outcome. Researchers can then look at variables in each group to see if there are any factors that are associated with having or not having the outcome. The difficulty in case-control studies is in ensuring that cases and controls are a representative sample from the same source population [9]. This has resulted in the development of nested case-control study designs, where cases and controls are both selected from a well-defined source population [9]. In addition, because case-control studies are usually retrospective, they can be affected by recall bias, if, for example, cases and controls are asked to accurately answer questions about their exposure to risk factors over many years. Case-control studies are particularly valuable in investigating rare diseases, or chronic diseases such as cancer, cardiovascular disease, or diabetes, which may result from long-term exposure to particular risk factors. In these scenarios, RCTs would be difficult to perform, even if the risk factor was already known and an intervention existed to change it. Furthermore, RCTs of sufficient numbers are expensive, and carry the difficulty of recruiting sufficient numbers of individuals with the condition to secure adequate statistical power. In chronic conditions there is, furthermore, usually a need for long-term follow-up.

3.3. Cross-sectional studies

These provide information on the prevalence of a particular condition at a single time point [8]. Cross-sectional studies can also be used to indicate potential associations by examining the presence or absence of risk factors and their association with disease prevalence in the selected population [8]. However, the inherent limitation of a cross-sectional study is that sampling only takes place at one time point, so it can be difficult or impossible to infer cause and effect. Many cross-sectional studies take the form of questionnaires
and thus can be prone to responder bias, in that those individuals who give their time to respond to a questionnaire may not be representative of the whole population. This type of study may also be subject to recall bias [8].

Because observational studies include a wide range of study designs, this review will limit its focus to prospective cohort studies of glucose-lowering therapies.

4. Purpose of an observational study

There are four main reasons why interventional approaches may not be appropriate. First, if the size of effect is very large, observational studies may be adequate to demonstrate effectiveness [5]. However, this is probably most relevant in a historical context (e.g., the first use of penicillin for bacterial infections, or of insulin in people with type 1 diabetes), as, in the current environment, interventional approaches would be required for the initial demonstration of more marginal changes in efficacy and safety. Second, interventional approaches may be inappropriate where sufficient study participants can never be recruited, either for rare conditions or to identify rare adverse events [5]. Third, interventional approaches may be impossible for ethical or other reasons, for example where they test a strongly and widely held clinical opinion. Finally, interventional approaches may not accurately reflect real-life clinical practice, due to patient selection or the power of investigator intervention.

Findings from prospective cohort studies of medicines are important because they provide information about the safety, efficacy, and tolerability of medicines while in daily clinical use, and in particular in the environment of other medical conditions and treatments. These studies are generally carried out once a medicine has received approval from regulatory agencies. Therefore, the information they generate supplements data from the RCTs that were used to support marketing approval. An important aspect of the quality of observational trials in this post-marketing period is that they should systematically collect safety data in a large cohort of patients as representative as possible of the whole population.

Observational trials can also provide important data sources for health economic (HE) models. The Center for Outcomes Research Diabetes Model provides information on long-term complications, life expectancy, quality-adjusted life expectancy and total costs based on baseline cohort characteristics, past history of complications, current and future diabetes management and concomitant medications, screening strategies, and changes in physiological parameters over time [10]. Due to their size, observational studies can provide the necessary data for populating the model, including baseline disease status and demographic information, disease natural history and outcomes, and drug usage. These models can be used to predict long-term HE outcomes. In addition, as observational trials are non-interventional, individuals are only enrolled in the study once their local clinician makes the decision to initiate them on the treatment under study. Given that prescribing and dosing decisions are made in accordance with local prescribing practices, observational study databases include useful country-specific information that can provide targeted HE information [10–13].

5. Differences between observational studies and RCTs

While RCTs are a core part of the drug evaluation process there is a growing appreciation that other methods of investigation, including observational trials, can add valuable information. A major criticism of RCTs reflects their applicability to the real world. For example, the majority of RCTs have rigorous inclusion and exclusion criteria, which in some cases can mean that the type of patient enrolled into a particular trial bears little resemblance to the typical patient seen in the clinic [14]. RCTs are often performed in specialist centers rather than in the primary care centers where the majority of patients would receive their treatment [2]. Some of the differences between RCTs and clinical practice are summarized in Table 1 [2]. The table also includes suggestions for study designs that would more closely reflect real-life clinical practice.

6. Designing the ideal observational study

6.1. Countering bias

A criticism of observational studies is that they are open to the effects of bias, and that the bias may be hidden. The main causes of bias in a cohort study are patient selection and loss to follow-up [8]. Selection bias implies that the participants entered into the study are not a representative selection of the background population. For example, if a non-blinded, non-randomized cohort study were set up to test oral glucose-lowering monotherapy versus combination therapy, it would be possible that investigators would assign individuals with more severe diabetes to the combination therapy group due to the perception that these individuals would benefit from more intensive therapy. However, the effect of this selection bias would worsen the overall outcome in the combination therapy group. Interventional trials counter selection bias by randomizing patients to particular treatments. For observational trials, an important mechanism to reduce selection bias is to limit enrolment to individuals who have been already begun study treatment according to routine clinical practice, and to admit all such individuals within a geographical or time constraint. Loss to follow-up can also bias results because there is the potential that those patients who stop participating in the study may have done so due to a treatment-related effect (for example lack of effectiveness or poor tolerability). Therefore, efforts should be made to minimize loss to follow-up, particularly in studies that follow populations over extended periods [8]. All studies should be on an intention-to-treat basis, ie, follow-up is independent of whether the initial study treatment is continuing.
6.2. **Addressing confounding**

Confounding occurs when an independent variable is associated both with the variable of interest and with the outcome [8]. Common confounding variables include age, gender, race/ethnicity, and smoking status, but these can often be measured and statistical adjustments made for them. More problematic are hidden confounders, such as when the presence of a comorbidity affects the treatment people are started on. For example, metformin has been traditionally avoided in people with heart, liver, and kidney disease, while insulin is used more extensively in people with other illnesses. Accordingly, people on metformin will appear to survive better than people on insulin, although the effect is not related to the treatment. The effect of confounding in RCTs can be limited by randomization such that, in a large patient population, confounding factors will be equally distributed between the treatment groups. For observational studies it is important that any relevant confounding factors are identified and measured and adjusted for in the analysis; however, often they cannot be identified from the datasets used [15]. Thus, despite attempts to address confounding in the study design, it is possible that results may be affected by residual confounding by unknown or unmeasured factors that affect the outcome but have not been accounted for in the analysis [15].

6.3. **Importance of size**

An advantage of observational trials is that they are far less costly per participant to run than RCTs, and it is thus possible to afford to recruit more participants. This gives them high statistical power and an ability to detect rare adverse events that would not be detected by RCTs. By the time they reach marketing approval many pharmaceutical agents will have been tested in only 1000-5000 individuals [16]. While testing in this number of individuals can demonstrate efficacy and provide information on safety, rare adverse events will often be missed. For example, 3000 individuals would need to be treated to detect a single adverse event with 95% confidence if it affected 1 in 1000 patients and 6500 patients would need to be treated to detect 3 cases of a rare event with 95% confidence if it affected 1 in 1000 patients. Therefore, observational trials often include the types of people under-represented in RCTs, such as those with comorbidities, those receiving multiple pharmacological interventions, people in full-time employment, or those investigators think would be unlikely to adhere to treatment protocols [2]. In designing an observational trial, eligibility

<table>
<thead>
<tr>
<th>Clinical practice</th>
<th>Typical RCT</th>
<th>Observational studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Most management in primary care with a proportion in secondary care</td>
<td>Clinical specialists. Secondary care populations, and participants recruited via advertisement</td>
<td>Representative of all patient populations</td>
</tr>
<tr>
<td>All diagnosed patients</td>
<td>Patients carefully selected to create homogeneous study population</td>
<td>Need well-characterized study population but should minimize inclusion/exclusion criteria</td>
</tr>
<tr>
<td>Diagnosis by clinical picture, sometimes without robust confirmation</td>
<td>Primary diagnosis confirmed by defined criteria and/or laboratory tests/imaging</td>
<td>Require a minimum of confirmatory testing</td>
</tr>
<tr>
<td>Patients treated with and without comorbidities</td>
<td>Patients with significant comorbidities usually excluded</td>
<td>Patients selected on the basis of use of some product or carrying some diagnosis</td>
</tr>
<tr>
<td>Physician and patient interaction determines treatment choice</td>
<td>Randomization determines treatment allocation</td>
<td>Variable. May reflect clinical practice, or cluster randomization or other mechanism to allow exercise of clinical judgement</td>
</tr>
<tr>
<td>Active treatment given to all appropriate patients</td>
<td>Placebo or active comparator controlled</td>
<td>Comparator of ‘usual care’, which may include treatments in routine use</td>
</tr>
<tr>
<td>Follow-up by clinical need</td>
<td>Structured and frequent follow-up</td>
<td>Flexible follow-up more reflective of clinical practice</td>
</tr>
<tr>
<td>Data collection by clinical need</td>
<td>Extensive data collection</td>
<td></td>
</tr>
<tr>
<td>Patients aware of their treatment</td>
<td>Treatment blinded where feasible</td>
<td>Open label</td>
</tr>
<tr>
<td>Partial non-adherence usual (affected by treatment convenience, efficacy, adverse events, economic factors)</td>
<td>Protocol compliance encouraged and monitored</td>
<td>Non-adherence an endpoint to be investigated for cause</td>
</tr>
</tbody>
</table>

Table adapted from Freemantle and Strack, 2009 [2].
criteria should be defined but should be broad enough so that the study population is representative of the type of people under management in the general population.

6.5. Clear rationale and study methodology

It is nevertheless important that observational studies comply with the core principles of medical research as outlined in ethical guidelines such as the Declaration of Helsinki [17], procedural guidelines such as those of the International Committee on Harmonization [18], and other guidelines such as those produced by the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) group [15]. These guidelines stress that the study design should reflect a thorough knowledge of the background scientific literature (condition and treatments), including gaps in the knowledge base, and that there should be a clear scientific rationale for undertaking the study [15,17]. The protocol should state specific objectives, including study population, exposures, clearly defined outcomes, and statistical methods [15]. The protocol should also include information about how the study size was calculated [15]. Statistical methodology should be described including the intended statistical tests, the primary outcome and secondary outcomes to be tested, an indication of how potential confounders and effect modifiers would be dealt with, and how missing data would be imputed.

It is important that the study should have adequate timelines to maximize follow-up by physicians and patients. As observational trials are non-interventional, the protocol should include only general guidance for the physician on follow-up, leaving decisions about follow-up intervals to the discretion of the physician. In this way, the trial can most closely approximate local clinical practice.

6.6. Global reach

The external validity of observational studies that are only performed in a few countries can be questioned because clinical practice can vary considerably from country to country, as will ethnic mix and use of resources. By including a wide range of countries in the study a more accurate picture can be formed of how a study product will perform globally. Furthermore, potentially interesting local variations in practice and performance can be investigated.

6.7. Ethical approval

There is some debate about the requirement for ethical approval of observational trials. For example, it can be argued that there is little that would require ethical consideration in an observational trial in which people with type 2 diabetes are treated according to standard clinical practice and then followed up at regular intervals [19]. However, an investigator participating in any study may have an interest in including people in that study (such as funding to their institution, or the prospect of travel to meetings), and perhaps an interest in publishing or presenting study results. Ethical committees sometimes also play an important role in promoting good quality research by identifying poorly designed studies, in which inappropriate designs, unrepresentative or small samples, and incorrect methods of analysis might result in useless results [20]. Accordingly, ethics approval is required in most countries for any study where patient data are used for any purpose beyond clinical care. Appropriately then, participants are expected to provide informed consent if their medical records are to be accessed or additional measurements performed even if treatment is not changed, and the format of that consent and the information provided for it should be subject to ethics review.

7. Roles, responsibilities, and expectations in observational studies

7.1. Medical investigators

The role of medical investigators in observational trials is no less important than it is in RCTs. However, as observational trials reflect by definition real-life clinical experience, and are carried out predominantly in primary care facilities, decisions about initiating patients on a particular treatment, dose/dosing frequency, and timing of follow-up should be at the discretion of the investigator and should reflect local practice. Physicians should ensure that participation in an observational trial does not affect their prescribing decisions, so that prescribing decisions should be made based on usual clinical judgement and not distorted by participation in the study. In prospective observational studies there will be a protocol defining the data items to be extracted from the participant’s clinical record, and, in some cases, this may drive some specific investigations, for example for the purposes of safety monitoring. In general, however, participant management, after informed consent, will be according to usual practice.

A medical investigator’s primary responsibility is to the people under their care, in any study of any type. Such responsibility will fall on a qualified physician, even if other multidisciplinary team members are involved. This duty of care to participants means that investigators need to ensure that any observational trials in which they participate are approved by local regulatory bodies.

Medical investigators also have a direct or indirect contractual responsibility to the funder of any study to ensure that they accurately report safety and efficacy outcomes according to agreed procedures. Any publications or presentations of the results should be within guidelines for appropriate reporting of observational trials, such as those published by STROBE.

Observational studies give the medical community an opportunity to share latest clinical data/practice and to report insights from their own clinical practice. Investigator participation in observational studies may be simpler than participation in RCTs (no special data collected, no defined study visits), but motivation to participate may be lower. Industry funding can assist medical investigators to generate clinical data that is relevant for their country/region. This may be particularly important for investigators in countries that do not have the healthcare infrastructure to participate in multinational RCTs. In addition, such trials allow investi-
be balanced and should take care not to influence current regulations. Participation in observational studies must also involve committees to ensure that their trial complies with local work of local regulatory agencies/health authorities/ethical committees to develop useful links with local, regional, and global investigators in their field. Post-marketing surveillance for adverse events is mandatory, but through support of observational studies industry can generate data to confirm/refute wider hypotheses and generate epidemiological data for the benefit of the medical community.

7.2. Health authorities

In general, local and/or national ethics committees’ and health authorities’ approval should be obtained before conducting an observational study. Different requirements will apply regarding the approval of an observational study according to country-specific regulations. Accordingly, structures must be in place to ensure the competent and timely review needed for such functions.

The role of local ethical committees is to ensure that the wellbeing of the individual takes precedence over investigator, funder, and sponsor interests in accordance with ethical guidelines such as the World Medical Association’s Declaration of Helsinki [17] or the International Conference on Harmonisation (ICH) Guideline for Good Clinical Practice [18]. Common themes of these guidelines include that research must be based on a thorough knowledge of the scientific literature, according to a clearly described protocol, that research should be carried out by appropriately qualified individuals, and that participants should be informed of the aims, methods, sources of funding, anticipated benefits, and potential risks or discomforts associated with the trial, as well as any other relevant information. Having received this information, individuals should be given the opportunity to freely give their consent to participate, having been informed of their right to refuse participation or withdraw consent at any time [17,18].

Health authorities also need to ensure the protocol is of an appropriate quality and with an adequate and stringent approval process to provide useful safety and efficacy information. They need to ensure that the implementation of an observational trial will not unduly influence physicians’ prescribing decisions.

National health authorities expect that any observational trial will be a source of safety information for the medications under study; this can be particularly helpful in those countries that lack centralized reporting of adverse events. In addition, data from observational trials can be used to provide HE information tailored to a particular country, to assist in making decisions regarding the inclusion or exclusion of drugs in formulary lists.

7.3. Industry

Funding by industry enables large-scale observational studies and analyses to be carried out. Industry may be involved at all stages of the trial from conception and study design through to data analysis and reporting.

Pharmaceutical companies must work within the framework of local regulatory agencies/health authorities/ethical committees to ensure that their trial complies with local regulations. Participation in observational studies must also be balanced and should take care not to influence current clinical practice. Companies should ensure establishment of internal guidelines (SOPs) on how to plan and implement studies and register study results on publically accessible sites. Industry should ensure that any publications coming from such trials comply with guidelines (eg, STROBE) [15,21].

Observational trials can provide industry with important data that complements and extends information from RCTs. Primarily, they are used to provide important safety information on adverse events that may not be identified by RCTs. In addition, observational trials provide pharmaceutical companies with information about how their product performs in daily clinical practice, in relatively unselected patient populations.

8. Summary

Observational trials provide important information about the safety and efficacy of pharmaceutical therapies according to local standards of clinical care. Observational studies tend to have minimal inclusion and exclusion data, so the heterogeneous populations enrolled tend to reflect the types of cases physicians see in their clinics. In contrast, RCTs aim to enrol very homogeneous populations and so, for example, they may exclude patients with common comorbidities or severe disease. Observational trials are non-interventional, so patients are prescribed a treatment according to their physician’s judgement or local clinical practice and receive only standard clinical care. Again, this contrasts with typical RCTs, in which enrolled individuals receive intensive clinical support and follow-up. Therefore, while RCTs are essential to prove the safety and efficacy of drug therapies, observational trials provide valuable additional information about how therapies perform in real life.

Conflict of interest

All authors were members of the AChieve study advisory board and received support from Novo Nordisk to attend advisory board meetings. Prof. Wenying Yang reported no other conflicts of interest. Dr. Alexey Zilov has lectured in seminars and meetings arranged by different companies, and has received honorarium from Novo Nordisk, Sanofi Aventis, Eli Lilly, Novartis and Berlin Chemie. Dr. Pradana Soewondo has received reimbursement for attending symposia from Novo Nordisk, Sanofi-Aventis, Merck MSD Inc, and Novartis and has received honoraria for speaking from Novo Nordisk, Sanofi-Aventis, Merck MSD Inc, and Eli Lilly. He has received funding for carrying out post-marketing surveillance from Novo Nordisk and Sanofi-Aventis. Ole Molskov Bech is an employee of Novo Nordisk International Operations A/S. Prof. Fawzia Sekkal reported no conflicts of interest. Prof. Philip D. Home or the institutions with which he is associated receive funding for his research, education or advisory activities from AstraZeneca, Boeringher Ingelheim, BMS, Eli Lilly, GlaxoSmithKline, MannKind Corp, Merck MSD Inc, Novo Nordisk, Novartis, Roche Pharma, Sanofi-Aventis, and Xoma Inc.
Authorship

All authors made substantial contributions to the conception and design of the paper or the acquisition of data. All authors had the opportunity to critically review the manuscript during development and the final draft of the manuscript was approved by all authors prior to submission.

Acknowledgements

The authors take full responsibility for this paper but are grateful to Mike Lappin, PhD, of Watermeadow Medical (supported by Novo Nordisk A/S) for writing assistance.

REFERENCES


