

'Real-world' clinical trials in diabetes care: meaningful or meaningless?

PHILIP HOME

Key words: clinical trials, real-world, diabetes

Abstract

So-called 'real-world' studies seem increasingly popular in diabetes care, as are the economic evaluations in secondary literature based upon them. The term is usually used for pharmacoepidemiological uncontrolled observational studies of different designs. Interpretation of the study findings is, however, badly undermined by the very reasons that the randomised controlled blinded study was invented – namely, non-medication study effects and biases in investigator selection and behaviour. In diabetes studies, glucose control seems particularly susceptible to such effects, perhaps through changes in patient motivation and education. Further, insulin studies are heavily influenced by baseline factors such as the site of starting insulin, the health circumstances of the patient at the time and the clinician involved. It is rare to see these issues adequately addressed or attempts made to understand their influence. In this article an attempt is made to discuss some of the issues further.

Background and need

Recent years have seen the development and introduction of a very welcome myriad of new therapies to aid the management of diabetes. These have included new classes of agents with unique properties (eg, sodium-glucose linked transporter type 2 blockers), derivatives within class with new properties (eg, insulin analogues) and 'me-too' additions within class. The early development pathway is often not well documented, but nearly always ends with published phase 2 and pivotal phase 3 randomised controlled trials (RCTs), sometimes placebo-controlled, sometimes active-controlled (with advantages and disadvantages to both). Where possible, these are double-blinded. The ideas behind the modern medication RCT can be traced back to many writings, amongst them Bernard who advocated blinding, Fisher who was strong on randomisation, and Bradford Hill, the last often credited with drawing together these ideas, and an author on the seminal MRC streptomycin tuberculosis study published in 1948.¹

RCTs are, however, not without problems, and the reader might like to reflect why there has never been such a study of stopping smoking for amelioration of heart disease risk (still an unknown) or, indeed, even for reducing lung cancer risk. Of these concerns, the most telling are narrow participant selection and, for safety outcomes, small study size. In many ways these problems come from the same source, namely cost, as providing blinded trial drugs across multiple centres with careful external monitoring and central assessment of outcome measures and central management is expensive. While the population willing to take part in studies may itself be a biased group, the requirements of statistical power with smaller numbers mean that population homogeneity is desirable at entry, a common example in insulin studies being that people with recurrent severe hypoglycaemia are excluded. While in recent years upper age restrictions have largely disappeared, numbers included tend to be relatively small and this approach has not resolved the difficulty in establishing efficacy and safety in, say, older people or some ethnic groups. These problems are often cited as the reason for doing 'real-world' studies, across broader populations unselected for willingness, or access, to join in RCTs.²

Terminology

The terminology used in publications of non-RCT studies of medications is often casual. The broad category of studies would be 'observational', but clearly observational studies extend to many areas of medicine beyond efficacy and safety of interventions. Smoking, air pollution, and COVID-19 severity studies are obvious examples. 'Real-world' studies in diabetes are for the most part pharmacoepidemiological (including pharmaco-economic) studies, although the term is little used. Reference can also be found to 'real-world clinical trials', but the use of 'trials' here is a misnomer, used casually or to mislead the reader, as any kind of trial ('to try') has the requirement of use for the purposes of study, and not use in clinical practice (a 'trial of therapy' is legitimate in clinical practice but is not relevant here).

The term 'real-world' as used here is also misleading. In 'reality' all studies including RCTs exist, and are done in living humans! The impression that seems to be conveyed is that the study represents the use of the medication as it is really used in clinical practice, rather than the limited circumstance of the RCT. However, a reading of nearly all pharmacoepidemiological studies suggests that this is not usually the case. Perhaps coming closest are studies based on electronic clinical records of large swathes of people – for example, based on a large number of UK general practices. But even here the practices contributing

Translational and Clinical Research Institute, Newcastle University, UK

Address for correspondence: Professor Philip D Home

Translational and Clinical Research Institute, The Medical School, Framlington Place, Newcastle upon Tyne NE2 4HH, UK

Tel: +44 (0)19 1208 7880

E-mail: philip.home@newcastle.ac.uk

Br J Diabetes 2021;21:169-172

<https://doi.org/10.15277/bjd.2021.317>

to such databases are not likely to be representative of prescriber or user populations. More commonly, in commercially supported studies there will be selection bias towards prescribers known to the sponsor, and of course then an unknown bias from the investigator/sponsor relationship. Sometimes retrospective criteria are used to try to ameliorate sponsor bias – for example, by stipulating that the participants have already been prescribed the medication in advance of recruitment³ – but in practice that interval is often so short that the medication can be begun and then study entry delayed for that interval. Bias can then affect the way the medication is used over that time, or for any further duration of prospective data collection – for example, by increased monitoring and contact.

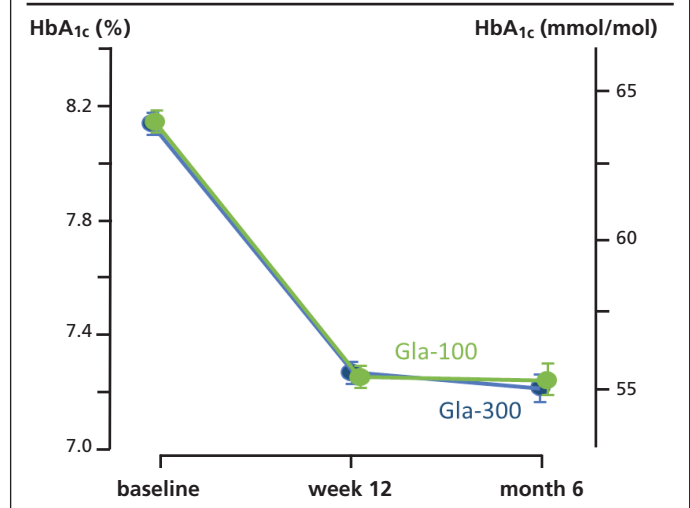
These problems may be partially ameliorated by introducing a random selection element into study centre recruitment, or using a database constructed for unrelated reasons. In one of my own insulin starter studies the sponsor's insulin (basal insulin alone) approach was used by 52% of the total study population, others using competing regimens,³ but often the population is selected for the specific medication under study.

How do problems of interpretation arise?

Many of the difficulties are simply those which arise from the very reasons blinded, controlled, randomised studies were devised – namely, study effects contaminating the intervention effects. Study effects can be seen in RCTs, but here the effect by virtue of blinding and randomisation are equal (except by play of chance) in the control and active intervention study arms. Thus, statistical comparisons at endpoint, or change from randomisation to endpoint, should reflect the difference between active and control populations alone. It is useful to see here that such non-medication study effects are quite large, larger than any difference arising from the medication per se. Thus, in Figure 1 we see an improvement in HbA_{1c} to a very useful extent with a new insulin, but this is also seen with the control arm insulin.⁴ What is telling here is that, prior to randomisation, both populations were being managed with the control arm insulin already (or with a therapy known to give similar HbA_{1c} levels), so the improvement in the control arm, and thus the new intervention arm, must be purely a study participation effect. In the Discussion section of similar studies it is often incorrectly said that glucose control improved with the intervention by n % – all that can really be said is that glucose control improved by n % in the intervention arm, with the emphasis on 'arm' rather than the intervention.

It is not possible to say precisely from these studies what drives the improvement in glucose control. It could, for example, be enhanced patient education, enhanced glucose monitoring or enhanced motivation from taking part in a clinical trial. In an early clinical trial of multiple injection therapy versus pump therapy,⁵ the biggest improvement in glucose control came between people agreeing to take part in the study and a screening visit (ie, before any kind of intervention), indicating that even the idea of increased focus on someone's insulin therapy can influence the behaviours improving outcome.

Figure 1. Change in HbA_{1c} in control (Gla-100) and new intervention (Gla-300) arms in people with type 2 diabetes entering a randomised controlled trial, having previously been using a basal plus meal-time insulin regimen. The prior basal insulin was dominantly the same as the basal insulin used in the control arm. After Riddle *et al*, 2014,⁴ with permission



Whatever the causes of the effect of being included in any study, they can be large and, indeed, dwarf effects of the therapy itself. In the A₁chieve study, of three different insulin analogue regimens, large improvements in glucose control from poor baseline levels in people not previously on insulin were not unexpected – the absolute improvement in control with a new medication is known to correlate with baseline levels.^{6,7} But similarly large improvements were seen in people transferred from other insulins, despite controlled trial evidence that HbA_{1c} is difficult to improve with a new insulin regimen. A clue to what is going on was the evidence that body weight did not rise, something that, in the context of big changes in HbA_{1c} and thus marked improvement in glycosuria, strongly suggests positive lifestyle change. Indeed, other surrogate outcomes also improved, notably lipids and blood pressure, and these improvements cannot be attributed to the insulin. Clearly then it was the events surrounding starting the insulin analogue that improved glucose control, and not the analogue itself.

For insulin it is worth thinking about the circumstances under which insulin gets started in anyone with diabetes. Our default thinking is that an insulin starter is someone in our clinics with control above target, often for a little longer than is desirable. But in reality, and the proportions do vary globally, a large percentage of insulin starts occur in other circumstances, notably during hospital admissions when improved control is mandated by a concomitant illness or an imminent procedure, or in ambulatory care after a referral from a non-specialist because of control difficulties (Table 1). In both these scenarios the patient will often not be previously known to the insulin team, and then enhanced diabetes education, improved self-monitoring, and

Table 1 Some problems underlying interpretation and usefulness of pharmacoepidemiological studies in diabetes; these are rarely recorded, adjusted for, or discussed in diabetes real-world studies

Problem area	Examples
Study population biases	
Investigator-related	Prior relationship with sponsor, study funding relationship with sponsor
Study-related	Study halo effect
Patient-related	Activated interest in glucose control, patient education and lifestyle change, enhance self-monitoring
Confounding issues	
Circumstance of starting new therapy	Inpatient emergency, inpatient procedure, concomitant illness, continuing ambulatory care, referral for poor control, referral for injection therapy
Site of starting new therapy	Hospital inpatient, specialist diabetes service, office-based specialist, primary care
Practitioner advising start new therapy	Insulin specialist, diabetes specialist, endocrinologist, diabetes care team, primary health care team.
Outcome and monitoring issues	
Use of diverse laboratories and diverse assays	Biochemical analytes including HbA _{1c}
No adjudication of true health outcomes	Uncertainty over stroke, MI
Poorly recorded health outcomes	Hypoglycaemia in many clinical records and all coding databases
Missing data	True health outcomes (eg, MI) occurring in remote healthcare sites
Therapy-related biases	
Position in glucose-lowering algorithm	Metformin early, insulin late
Guideline use for specific populations	GLP-1RA use if prior CVD, SGLT2 blocker use if HF or CKD progression
Contraindications (actual and historic) and positive indications	Metformin with CKD, advanced HF, liver disease; insulin use in the presence of other complex conditions affecting glucose control

CKD, chronic kidney disease; CVD, cardiovascular disease; GLP-1RA, glucagon-like protein-1 receptor agonist; HF, heart failure; MI, myocardial infarction; SGLT2, sodium-glucose transporter-2.

reasons to self-motivate will be of significance. Accordingly, it would be surprising if there was not improvement in surrogate outcomes, independent of any therapy change. Curiously, in pharmacoepidemiological studies as a whole, and certainly in 'real-world' studies in diabetes, the circumstances under which a new medication is started are very rarely recorded.

Confounding

Confounding is a huge problem in epidemiological studies. Living in proximity to roads is associated with a number of adverse health risks, but of course associates with urban pollution as a whole, poor health education and the problems associated with health deprivation. Too often authors are allowed to get away with some brief statement about association and causation, with no attempt to identify or quantify potential issues. In diabetes pharmacoepidemiological studies other confounders are rife and too often ignored. Particular problems concern metformin, and separately insulin (Table 1).

Metformin has conventionally been used (since 1998) as first-line therapy in a stepped algorithm, and hence usually in people with the shortest diabetes duration. Furthermore, its contraindication to use with renal impairment or with more advanced heart failure and, in some minds, liver disease has meant that it will inevitably be associated with better health in terms of long-term outcomes than other glucose-lowering medications, as indeed is found for comparisons with sulfonylureas and insulin.⁸ Insulin, by contrast, is enriched in use in people whose health is compromised by concomitant medical conditions, from myocardial infarction to chemotherapy. The bias effect here is very large; as severe adverse outcomes occur in only a small percent of our

populations per year and are predominantly in those with prior health impairment. Accordingly, restricted or enhanced use markedly biases outcome rates even sometimes several-fold. Prescribing bias is obvious in other studies, none more so than a study of the incidence of pancreatitis with and without exposure to glucagon-like peptide-1 (GLP-1) based therapies.⁹ In that study the GLP-1 therapy arm had a massively different baseline risk for pancreatitis across a wide range of known risk factors. The findings are uninterpretable.

Again, studies very rarely adequately assess these issues – often only going as far as adjusting by the Charlson Comorbidity Index at best. We know that the risk of severe hypoglycaemia (SH) is markedly increased in people with gastrointestinal, pulmonary and even skin disease.¹⁰ If a study of SH fails to account for these, it is difficult to be confident in any conclusion.

Data quality and health economic analysis

Collection of outcome data, whether surrogate measures or health events, is usually rigorously standardised in RCTs. Thus, HbA_{1c} is measured in a central laboratory, hypoglycaemic events are recorded by electronic diary and confirmed by a standard meter, and true outcomes are adjudicated. In many 'real-world' studies the data are taken from the routine clinical record, with lack of standardisation and uncertainty over such things as to whether a myocardial infarction was robustly diagnosed. Indeed, because the site of diabetes care is often different geographically from the site of, say, a vascular event or an eye procedure, it may be unreliably captured.

Particular problems surround hypoglycaemia. General practice databases are generally of coded events and much outpatient

hypoglycaemia simply goes uncoded. Certainly there is no mapping of grades of hypoglycaemia or its confirmation status to most GP clinical records, and indeed this is poor even in specialist units. At best, all hypoglycaemia information in pharmaco-epidemiological studies is to be viewed with caution.

Since health economic analysis depends on ascribing costs to health events (usually as health events saved being the offset for increased medication costs), it will be evident that cost-effectiveness calculations will also be unreliable. In short-term studies, changes in metabolic control (HbA_{1c} and sometimes other measures) have to be modelled to calculate the likely influence on true adverse outcomes over the years going forward – evidently, if the improvement in glucose control is overestimated, so will be the cost benefit. This was seen in a real-world-based analysis of Swedish data, where an HbA_{1c} gain was found in the context of starting insulin degludec and, although modest, would contribute to cost savings from a reduction in modelled long-term complications.¹¹ To a limited extent, such calculations will be able to estimate the cost-effectiveness of the intervention as a whole (the medication, the education, the self-monitoring, and the motivation) but will not be able to say what the benefit of the medication itself is, if indeed any.

Conclusion and way forward

Presently, it is difficult to see that 'real-world' studies of medications in diabetes care contribute anything usefully generalisable to our practice. Indeed, they are ignored by guideline developers and the health economists that advise them. In time there may be a way forward using true electronic health records given three conditions. Firstly, all the records for each person must be linked from all the different sites and practitioners delivering care; secondly, common definitions should be used for outcomes and measurements should be uniformly standardised; and, thirdly, the circumstances of care interventions (eg, outpatient or admission, professional affiliation of prescriber, referral or continuing care) need be recorded. But this is far from being realised currently. A problem here is that this approach will create a data monster, and it is unclear that statistical techniques yet exist to make all the appropriate adjustments, or whether – even if that is done – the findings will be applicable to the individual people with diabetes we serve.

Conflict of interest The author has received funding to himself or to the institutions with which he is associated for his research, speaking and advisory activities in relation to members of all classes of glucose-lowering

agents including those named in Figure 1 (Sanofi), and the sponsors of the A₁chieve (Novo Nordisk) and CREDIT (Sanofi) studies identified as references 6 and 3.

Funding No funding or other support was sought or provided for the writing of this invited article.

References

1. Marshall G, Blacklock JWS, Cameron C, *et al.* Streptomycin treatment of pulmonary tuberculosis. *Br Med J* 1948;**2**(4582):769–82. <https://doi.org/10.1136/bmj.2.4582.769>
2. Yang W, Zilov A, Soewondo P, Bech OM, Sekkal F, Home PD. Observational studies: going beyond the boundaries of randomized controlled trials. *Diabetes Res Clin Pract* 2010;**88**(Suppl 1):s3–s9. [https://doi.org/10.1016/S0168-8227\(10\)70002-4](https://doi.org/10.1016/S0168-8227(10)70002-4)
3. Home PD, Dain M-P, Freemantle N, *et al.* Four-year evolution of insulin regimens, glycaemic control, hypoglycaemia and body weight after starting insulin therapy in type 2 diabetes across three continents. *Diabetes Res Clin Pract* 2015;**108**:350–9. <https://doi.org/10.1016/j.diabres.2015.01.030>
4. Riddle MC, Bolli GB, Ziemann M, Muehlen-Bartmer I, Bizet F, Home PD. New insulin glargine 300 units/ml versus glargine 100 units/ml in people with type 2 diabetes using basal and mealtime insulin: glucose control and hypoglycemia in a 6-month randomized controlled trial (EDITION 1). *Diabetes Care* 2014;**37**:2755–62. <https://doi.org/10.2337/dc14-09912>
5. Reeves ML, Seigler DE, Ryan EA, Skyler JS. Glycemic control in insulin-dependent diabetes mellitus: comparison of outpatient intensified conventional therapy with continuous subcutaneous insulin infusion. *Am J Med* 1982;**72**:673–80. [https://doi.org/10.1016/0002-9343\(82\)90479-X](https://doi.org/10.1016/0002-9343(82)90479-X)
6. Home P, Naggar NE, Khamseh M, *et al.* An observational non-interventional study of people with diabetes beginning or changed to insulin analogue therapy in non-Western countries: the A₁chieve study. *Diabetes Res Clin Pract* 2011;**94**:352–63. <https://doi.org/10.1016/j.diabres.2011.10.021>
7. DeFronzo RA, Stonehouse AH, Han J, Wintle ME. Relationship of baseline HbA_{1c} and efficacy of current glucose-lowering therapies: a meta-analysis of randomized clinical trials. *Diabet Med* 2010;**27**:309–17. <https://doi.org/10.1111/j.1464-5491.2010.02941.x>
8. Schramm TK, Gislason GH, Vaag A, *et al.* Mortality and cardiovascular risk associated with different insulin secretagogues compared with metformin in type 2 diabetes, with or without a previous myocardial infarction: a nationwide study. *Eur Heart J* 2011;**32**:1900–08. <https://doi.org/10.1093/eurheartj/ehr077>
9. Singh S, Chang HY, Richards TM, Weiner JP, Clark JM, Segal JB. Glucagonlike peptide 1-based therapies and risk of hospitalization for acute pancreatitis in type 2 diabetes mellitus: a population-based matched case-control study. *JAMA Intern Med* 2013;**173**:534–9. <https://doi.org/10.1001/jamainternmed.2013.2720>
10. Zoungas S, Patel A, Chalmers J, *et al.* Severe hypoglycemia and risks of vascular events and death. *N Engl J Med* 2010;**363**:1410–18. <https://doi.org/10.1056/NEJMoa1003795>
11. Jendle J, Ericssonb Å, Ekman B, *et al.* Real-world cost-effectiveness of insulin degludec in type 1 and type 2 diabetes mellitus from a Swedish 1-year and long-term perspective. *J Med Econ* 2020;**23**:1311–20. <https://doi.org/10.1080/13696998.2020.1805454>