The Role of Surrogate Endpoints in the Evaluation of Efficacy and Safety of Therapeutic Interventions in Diabetes Mellitus

Aleksandra Wieczorek1, Przemyslaw Rys1, Iwona Skrzekowska-Baran2 and Maciej Malecki3

1 HTA Consulting, Krakow, Poland. 2 NovoNordisk Polska, Poland. 3 Department of Metabolic Diseases, Jagiellonian University College of Medicine, Krakow, Poland. Address correspondence to: Maciej Malecki, e-mail: mmalecki@cm-uj.krakow.pl

Manuscript submitted October 26, 2008; resubmitted November 21, 2008; accepted November 23, 2008

Abstract

In this paper, we examine the concept of surrogate endpoints (i.e. substitute outcome measures) and review their use in clinical trials involving therapies for diabetes mellitus using the example of metformin. Trials such as DCCT and UKPDS, in which patient-important endpoints were evaluated, are relatively rare in diabetology. Clinical decisions, therefore, are often based on evidence obtained using surrogate outcomes, usually fasting or postprandial glycemia or glycated hemoglobin level. In contrast to patient-important endpoints, surrogates do not describe direct clinical benefit to the patient. However, a proven association between a surrogate and patient-important endpoint is essential to draw appropriate therapeutic conclusions. In the process of new drug development, the duration of follow-up, sample size and methodology of the studies initially available are often inadequate to demonstrate the effect of the intervention on patient-important endpoints. Evidence concerning the effect of an intervention on surrogate outcomes usually comes first, followed only later by reports describing its influence on patient-important endpoints. Metformin may serve as an example in several ways. The first publications reported beneficial effects on glycemic control and body weight. Outcomes from the subsequent UKPDS study suggested the patient-important efficacy of metformin measured as a reduction in mortality and a decrease in the incidence of diabetic complications, including myocardial infarction. This reasoning process worked for some but not all strategies. It is particularly questionable whether a change in surrogate endpoint was associated with a potential deterioration in patient-important outcomes. Defining the general relationship between surrogates widely used as measures of metabolic control and patient-important endpoints remains an important challenge in contemporary diabetology.

Keywords: diabetes · patient-important endpoint · clinical study · metformin · HbA1c · metabolic control

Vocabulary of endpoints description in clinical trials

Understanding and interpreting the methodology and results of clinical studies requires familiarity with a wide range of terms and definitions which, if not used carefully, may be difficult or confusing for the reader of clinical literature. The different concepts associated with describing an endpoint may serve as an example. An endpoint (or an outcome) is any measurable effect (usually, but not necessarily, related to health) observed in individuals participating in a clinical trial. The effect of an external factor (for example, tobacco smoking as predictive factor or a prophylactic cholesterol lowering drug as therapeutic factor) on this kind of selected endpoint is investigated in
临床研究文献 [1]。阅读临床研究文献可能会遇到不同的描述，包括结局/终点：主要与次级、硬与软、临床重要性、患者重要性以及替代。主要与次级的结局是其中一个最复杂的；它涉及研究者的选择测量结果，为研究者提供选择，影响样本大小计算、分析和报告结果。在一些实验中，主要终点可能为死亡或中风。在其他资源使用和在其他研究中，平均胆固醇水平，肌酐水平，血压，或每平方厘米的毛发数量。这些结果可能或可能不会是客观的（硬，可再现性和不会影响测量值的测量过程，如死亡，截肢的臂，胆固醇水平）或主观的（软，主观的，可能由测量过程影响，如描述病人的情绪在某个时刻）。

在比较结论与主要结论之间（硬与软，主要与次级），数值上并非临床重要或患者重要的终点可能会在主要与次级条件下发生变化。使用临床改善的几个方式，其中之一是问自己，如果结果与正常结果相同，这会是仅利用治疗改变时会得到的。如果某种硬的与软的或非临床重要与次级结果（或替代）终点是不重要的，这一方式可能会更容易。一个方式的定义，说患者的重要结果是问我们自己和我们的患者一个问题，如果这是唯一结果改变导致的。在某种情况下，这种结果可能在使用治疗方面会因其非毒性而得到使用。有死亡或截肢的臂考虑到治疗效果，这种结果可能不重要或至少对于决策最为重要。

根据当前定义，美国食品和药物管理局（FDA），替代终点是一个用于替代患者重要终点的替代定义，即用其描述一个治疗效果。在次级定义中，替代定义可能会被认为太严格 [7]。例如，ACE抑制剂和钙通道阻滞剂可能具有相似的血压降低效果，但它们对心血管事件的影响可能不同，这意味着血压控制可能不完全是唯一因素。对于治疗效果的治疗结果 [8-10] 。许多作者，包括Busye, Molenbroeks和Friedman [7, 11, 12]，对做重要贡献。例如，ACE抑制剂和钙通道阻滞剂

要总结，替代终点并不直接描述治疗效果。通过声明和验证（在背景下示例统计相关性）或在生物学考虑因素（在这种情况下造成因果关系）。研究组的终点包括生物化学（例如，血流胆固醇水平）以及病理生理学（例如，血流血压）以及形态学变量（例如，冠状动脉或左心室肥大）。一些作者认为替代终点可能不是直接参与病理生理学的疾病，但反映了活动过程导致不利事件。例如，患者可以作为替代终点在HbA1c。血红蛋白A1c的使用在替代终点（替代定义）的定义中可能不会改变。这种定义可能会直接导致不利事件 [13]。例如，一个替代的定义会是水平的HbA1c。对血糖的影响取决于该水平，其敏感性反映了血糖暴露。在几周内，

To summarize, a surrogate does not describe direct benefit to the patient and its usefulness in evaluating such benefit depends on how much it allows for prediction of treatment effect on patient-important out-
comes (Figure 1). A parameter that is potentially useful as a surrogate endpoint may be:

- directly involved in the pathophysiological pathway of the disease (i.e. it is identical with one of the mechanisms, on which a medication or intervention used in treatment of the disease acts), for example coronary artery stenosis as a surrogate for myocardial infarction or cardiac death;
- related to, but not directly involved in, the pathophysiology of the disease (i.e. directly associated with one of the stages of the disease occurrence and thus correlating with its progression or response to treatment), for example serum cholesterol level as a biomarker of myocardial infarction risk;
- not directly associated with any disease mechanism or its response to treatment (for example HbA1c level as a surrogate for the incidence of late complications of diabetes mellitus)[14].

It is sometimes neither possible nor feasible to perform a valid evaluation of the efficacy and safety of the treatment based on patient-important endpoints in a reasonably short time. Trials, in which those clinically relevant endpoints are evaluated, require a relatively long time and large investigated groups. These conditions complicate the organization and increase the costs of such trials, but the most important problem is the time usually required before the results are known, a period usually measured in years. Meanwhile, patients and physicians expect new medications to be introduced onto the market as rapidly as possible.

Therapeutic decisions based exclusively on surrogate endpoints have their advantages and disadvantages. On the one hand, such decisions may result in the quicker introduction of a specific intervention into clinical practice, as was the case with certain antiviral drugs used in HIV infection therapy, where the decision was based on their ability to increase CD4+ cell count [15]. Such a policy also makes it possible to reduce the costs of assessing a health technology or medical product. On the other hand, in certain situations this may lead to false conclusions resulting in unfavorable clinical consequences. For example, sodium fluoride was proven to increase bone mineral density (BMD), but, contrary to expectations, clinical trials demonstrated that it had no effect on the fracture rate in postmenopausal women [16]. Worse still, certain antiarrhythmic drugs were introduced on the basis that they suppressed premature ventricular contractions with the hope of saving lives. The results of clinical trials looking at patient-important outcomes proved this supposition to be tragically wrong [17]. Use of surrogates is therefore hampered by the fact that the evaluation of their association with relevant clinical endpoints may be incomplete, inadequate or downright erroneous.

### Surrogate endpoints in diabetes mellitus trials

The surrogate endpoints most often used in diabetes mellitus studies include HbA1c level, FPG and PPG. The patient-important and surrogate endpoints related to diabetes mellitus are summarized in Table 1. It is interesting to note that some stages of a particular diabetic complication, for example nephropathy, should be considered patient-important while others tend to be classified as a substitute (end-stage renal disease vs. microalbuminuria) [18]. The pathophysiological rationale for the use of HbA1c, FPG and PPG in assessment of metabolic control in diabetes mellitus is well documented. However, this does not automatically mean that any therapeutic intervention resulting in improved metabolic control as assessed by these parameters reduces the risk of developing late micro- and macrovascular complications or of patients’ mortality. According to recommendations from the American Diabetes Association (ADA), World Health Organization (WHO) and American College of Endocrinologists (ACE), HbA1c level is considered the “gold standard” in assessment of metabolic control and the specific level of HbA1c constitutes the target at which treatment of both type 1 and 2 diabetes mellitus should be aimed [19]. It should be noted that in two fundamental diabetology studies, the DCCT in type 1 and the UKPDS in type 2 diabetes mellitus, pharmacological interventions leading to the
reduction of this parameter were associated with improvement in clinical endpoints, i.e. microvascular complications [20, 21].

Table 1. Comparison of endpoints used in diabetes mellitus clinical trials

<table>
<thead>
<tr>
<th>Patient-important (clinically important) endpoints</th>
<th>Surrogate endpoints</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mortality</td>
<td>Glycated hemoglobin level</td>
</tr>
<tr>
<td>Quality of life</td>
<td>Postprandial glycemia</td>
</tr>
<tr>
<td>Macrovascular complications</td>
<td>Fasting plasma glucose</td>
</tr>
<tr>
<td>Myocardial infarction</td>
<td></td>
</tr>
<tr>
<td>Coronary heart disease</td>
<td></td>
</tr>
<tr>
<td>Cerebral stroke</td>
<td></td>
</tr>
<tr>
<td>Cardiovascular death</td>
<td></td>
</tr>
<tr>
<td>Peripheral vascular disease (incl. amputation, claudication)</td>
<td></td>
</tr>
<tr>
<td>Microvascular complications</td>
<td></td>
</tr>
<tr>
<td>Blindness</td>
<td></td>
</tr>
<tr>
<td>End-stage renal disease</td>
<td></td>
</tr>
<tr>
<td>Foot ulceration</td>
<td></td>
</tr>
</tbody>
</table>

All three organizations (ADA, WHO and ACE) also underline the importance of normalizing FPG and PPG. Nevertheless, it must be stressed that clinical trials convincingly demonstrating a direct relationship between reduction in PPG and improvement in patient-important endpoints are not yet available. Several trials in patients with type 1 or type 2 diabetes demonstrated the favorable effect of PPG reduction on the intima-media complex thickness (IMT) [22-24]. IMT cannot, however, be considered to be a clear patient-important endpoint, although its correlation with the incidence of cardiovascular events (stroke, myocardial infarction and others) has been described in several studies that included non-diabetic subjects [25-27].

The relationships between the different surrogates reflecting metabolic control in diabetes mellitus are themselves worth consideration. The level of HbA1c alone does not convey comprehensive information about the amplitude and frequency of glycemic fluctuations, especially in patients with type 1 diabetes mellitus, which makes it an insufficient sole marker of glycemic control [28]. A series of trials was undertaken to investigate the relationships between HbA1c level and FPG as well as PPG. Earlier studies demonstrated a weaker correlation of HbA1c with PPG than with FPG [28]. Other reports indicated that relationships between HbA1c level and PPG and FPG may vary, depending on level of metabolic control. A study in patients with type 1 diabetes mellitus demonstrated that PPG in patients whose HbA1c level was in the lowest quartile (<7.3%) was responsible for 70% of its increase above normal levels, while for those in the highest quartile (≥9.3%) the influence of PPG was decisive. Between those two extremes, the effects of FPG and PPG on HbA1c level were similar [29, 30]. These findings were confirmed in the latest research published by Woerle et al. [31]. They showed that the impact of PPG was about 80% on HbA1c when HbA1c was below 6.2%, and only about 40% when HbA1c was >9%.

In many diseases, establishing a relationship between surrogates and patient-important endpoints remains a challenge. Providing such proof may bring benefit to both patients and physicians, as both are interested in finding the most effective therapy. Introducing a new intervention into practice may be much faster if the effect of the intervention on a surrogate is measured and the relationship between the surrogate and a patient-important endpoint is unequivocally proven.

In case of diabetes mellitus, less than twenty large randomized clinical trials evaluating patient-important endpoints have been published. Frequently, there are no clinical trials regarding patient-important endpoints for some widely used interventions; for some interventions, trials of this kind appeared years after the intervention was introduced onto the market.

Could, or what is more important, should the results of clinical trials designed to evaluate surrogate endpoints be used in formulating clinical decisions? Metformin may serve as an example of a medication that was used in diabetic clinical practice for many years despite lack of proven beneficial effect on patient-important outcomes.

The example of metformin – surrogates and patient-important endpoints

Metformin belongs to a group of drugs called biguanides. It is currently one of the most popular medications used in treatment of type 2 diabetes mellitus. It has been in use for more than 50 years [32]. Metformin was banned in USA in 1977 and was removed from US market for the subsequent 2 decades because of uncertainty as to whether it was associated
with lactic acidosis. The drug reappeared on the US pharmaceutical market in mid-1990s [33].

In the nineties, use of this drug was restricted by numerous precautions, not only because of the risk of adverse effects, but also because the efficacy of metformin was proven with respect to typical surrogates only. The situation changed dramatically with the publication of the results of the UKPDS study in 1998 [20]. It may be interesting, therefore, to look at the evidence available prior to publication of the UKPDS results and afterwards.

In 1999 a meta-analysis of the results of all studies published up until 1995 where metformin was compared with placebo or sulphonylurea derivatives was published by Johansen [34]. This meta-analysis was based on a valid systematic review of clinical trials, and its results may be considered reliable and up-to-date at the time of publication. A review of the literature published before the mid-1990s identified 9 randomized controlled clinical trials, in which metformin was compared to placebo, and 10 randomized trials, in which metformin was evaluated in comparison with the sulphonylureas. The results of these studies were included in the meta-analysis.

The pooled results of the 9 studies led to the conclusion that metformin, in comparison to placebo, decreased FPG and the HbA1c level and caused no changes in body weight. The specific data are shown in Table 2.

In the same meta-analysis of 10 randomized controlled trials, no differences were found between metformin and sulphonylureas with respect to the change in glycated hemoglobin level or fasting glycemia. At the same time, it was demonstrated that the drugs investigated had the opposite effects on body weight. Use of metformin contributed to reduction of body weight, while use of sulphonylureas increased it (Table 3). The meta-analysis of Johansen, published before the UKPDS study results were available, demonstrated the efficacy of metformin, against placebo, measured by laboratory parameters and suggested a lack of unfavorable effect on body weight. As compared to sulphonylureas, metformin demonstrated similar effects on metabolic control, but its influence on body weight was more favorable.

The UKPDS study results published in 1998 evaluated the effect of metabolic control on clinical outcome in type 2 diabetes mellitus and demonstrated that the risk of death or diabetes-related complications was lower in the group of obese patients treated with metformin as compared to such patients treated with insulin or sulphonylurea derivatives. At the same time, patients treated with metformin (as compared to conventional treatment) were at lower risk of diabetes-related complications, myocardial infarction, cardiovascular death and death regardless of the cause [20].

In 2005, Saenz et al. [35] published a Cochrane Collaboration systematic review, including (among others) the UKPDS study results. In this paper, the favorable effect of metformin versus placebo on metabolic control was confirmed (Table 2). Similarly to Johansen’s meta-analysis [34], no differences between metformin and placebo were demonstrated with respect to changes in body weight (Table 3) [35].

In addition, this meta-analysis demonstrated more favorable effect of metformin on metabolic control in comparison to those of sulphonylureas. The difference between the two drugs was small but statistically significant. It also confirmed that metformin had a more favorable influence on changes in body weight.

### Table 2. Two meta-analyses on the comparison of metformin vs. placebo in type 2 diabetes mellitus

<table>
<thead>
<tr>
<th>Publication date</th>
<th>1999 [34]</th>
<th>2005 [35]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data of last search</td>
<td>1995</td>
<td>2003</td>
</tr>
<tr>
<td>FPG (mmol/l)</td>
<td>-2.0 (-2.4; -1.7)</td>
<td>-0.87 (-1.13; -0.61)</td>
</tr>
<tr>
<td>HbA1c (%)</td>
<td>-0.9 (-1.1; -0.7)</td>
<td>-0.97 (-1.25; -0.69)</td>
</tr>
<tr>
<td>Body weight (kg)</td>
<td>0.8 (-1.0; 2.5)</td>
<td>0.00 (-0.12; 0.12)</td>
</tr>
<tr>
<td>Any diabetes-related complications</td>
<td>-</td>
<td>0.74 (0.60; 0.90)</td>
</tr>
<tr>
<td>Diabetes-related deaths</td>
<td>-</td>
<td>0.61 (0.40; 0.94)</td>
</tr>
<tr>
<td>All cause mortality</td>
<td>-</td>
<td>0.68 (0.49; 0.93)</td>
</tr>
<tr>
<td>Myocardial infarction</td>
<td>-</td>
<td>0.64 (0.45; 0.92)</td>
</tr>
</tbody>
</table>

**Legend:** 1 Metformin compared with placebo. Data are weighted mean difference (95% CI in parentheses). 2 Metformin compared with conventional therapy. Data are relative risk (95% CI in parentheses).
Most importantly, however, the UKPDS study results proved its benefits with respect to patient-important endpoints (Table 2 and 3) [35].

Table 3. Two meta-analyses on the comparison of metformin vs. sulphonylureas in type 2 diabetes mellitus

<table>
<thead>
<tr>
<th>Publication date</th>
<th>1999 [34]</th>
<th>2005 [35]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data of last search</td>
<td>1995</td>
<td>2003</td>
</tr>
<tr>
<td>FPG (mmol/l)</td>
<td>-0.29 (-0.58; 0.04)</td>
<td>-0.16 (-0.27; -0.05)</td>
</tr>
<tr>
<td>HbA1c (%)</td>
<td>0.1 (-0.2; 0.4)</td>
<td>-0.14 (-0.28; -0.01)</td>
</tr>
<tr>
<td>Body weight (kg)</td>
<td>-2.9 (-4.4; -1.1)</td>
<td>-0.45 (-0.80; -0.10)</td>
</tr>
<tr>
<td>Any diabetes-related complications</td>
<td>-</td>
<td>0.78 (0.65; 0.94)</td>
</tr>
<tr>
<td>All cause mortality</td>
<td>-</td>
<td>0.73 (0.55; 0.97)</td>
</tr>
</tbody>
</table>

Legend: ¹ Metformin compared with sulphonylurea (SU). Data are weighted mean difference (95% CI in parentheses). ² Metformin compared with SU or insulin. Data are relative risk (95% CI in parentheses). ³ Approximate data, read from the graph.

The above theoretical and historical considerations on the role and position of metformin demonstrate that, in this case, the favorable effects of the drug with respect to surrogates and body weight were confirmed in randomized clinical trials using patient-important outcome measures, long after the drug was introduced on the basis of its effect on metabolic surrogates.

Further comments on surrogate endpoints, epidemiological studies and other classes of diabetic drugs

Our review of the literature indicates that clinical trials using patient-important endpoints are scarce in diabetology. The Cochrane Collaboration, an organization which develops valid and up-to-date systematic reviews, has published dozens of these on diabetes mellitus. Analysis of these and other credible systematic reviews dealing with diabetic management indicates that the efficacy of most interventions, both pharmacological and behavioral, is proven with respect to surrogates, while the effects of those interventions on patient-important endpoints remain unknown or unproven. For example, no studies unequivocally confirming the favorable effect of sulphonylureas on the risk of diabetes-related complications have been published so far. Although sulphonylurea derivatives are used as comparators for new drugs (e.g. the ADOPT and RECORD studies for rosiglitazone) [36, 37, 38], the effect of this group of drugs on primary endpoints as compared to placebo has not been demonstrated [39]. In June 2008, the results of the ADVANCE study, which indirectly suggested that gliclazide has a favorable effect on a combination of macrovascular and microvascular endpoints (or, at least, on new or worsening nephropathy), was published. However, the objective of that study was to assess the effect of strict glycemic control (target HbA1c level ≤ 6.5%) not that of the drug itself on reducing the risk of diabetic complications and, therefore, any conclusions regarding sulphonylureas are indirect [40].

Other anti-diabetic medications, such as alpha-glucosidase inhibitors or meglitidines, were also proven to have a favorable effect on metabolic control, while their effect on patient-important endpoints in type 2 diabetes mellitus remains unknown, unproven or disputable [41, 42]. In 2004, Hanefeld et al. published a meta-analysis [43] which indicated that acarbose exerts a favorable effect on reduction of the risk of cardiovascular complications in type 2 diabetes mellitus; however, that analysis was not based on systematic review. The criteria for considering studies were not clearly specified, only trials from Bayer Acarbose clinical database were included (no systematic searching), no critical appraisal of included studies were performed and only pooled results of the meta-analysis were presented but not data from single studies [44, 45]. However, a systematic review concerning acarbose published by the Cochrane Collaboration one year later [41], failed to confirm the optimistic report made by Hanefeld et al. Publication of the recent ACCORD study casts more doubt on the connection of surrogate endpoint HbA1c (6.4% versus 7.5%) and patient-important outcomes, as the study was interrupted because of excessive mortality in the intensively treated group [46].

Final comments and our conclusions

The facts discussed above indicate that, when dealing with diabetes mellitus, individual physicians, opinion leaders and regulatory bodies have often based
(and still frequently base) their decisions or recommendations on the results of clinical trials evaluating surrogate endpoints. Such decisions, albeit difficult and associated with the risk of erroneous clinical reasoning, were made in the past, as illustrated by the example of metformin. Studies evaluating the effect of drugs on metabolic control predominate in diabetology, which is why it was so important that, in pivotal trials, surrogates were found to be associated with patient-important endpoints, at least in the form of microvascular complications and in relation to specific drugs. In the UKPDS study, the relationship between lower HbA1c and a lower risk of developing microvascular complications in type 2 diabetes mellitus was demonstrated. The DCCT study confirmed the favorable effect of metabolic control as measured by this surrogate endpoint on reducing the risk of microangiopathies in type 1 diabetes mellitus.

Additional support for the relationship between surrogate and patient-important outcomes was provided by Selvin et al. [47], who demonstrated in a meta-analysis of observational studies that a 1 percentage point increase in HbA1c level is related to a statistically significant increase in the risk of cardiovascular complications. Along the same lines, Groeneveld et al. reviewing the literature demonstrated that in 23 out of 27 identified studies there was a relatively weak but statistically significant relationship between the metabolic control level (measured using FPG or HbA1c) and mortality [48]. However, it is impossible to ignore opposite findings, including the potential for increased mortality observed in the ACCORD trial. Another finding that cannot be ignored is that from other areas of medicine, where correlation between surrogates and patient-important outcomes hold for some but not all drugs influencing any given surrogate [2].

Considering all of the above, we conclude that both pivotal clinical studies as well as meta-analyses (which are scarce and suffer from limitations) indicate a correlation and association between the improvement in HbA1c level and changes in certain selected patient-important endpoints, best proven for micro-vascular complications. Evidence for the relationship between PPG—used as a surrogate in clinical trials—and the occurrence of patient-important endpoints is lacking. Formulating definite and complete answers to questions concerning the relationship between all surrogates widely used as metabolic control measures and patient-important endpoints seems to be an important challenge in contemporary diabetology. After new data from clinical studies currently underway becomes available and a complete systematic review of both new data and already published reports is performed, we may well be close to an answer.

### References


