Interpreting Adverse Signals in Diabetes Drug Development Programs

CLIFFORD J. BAILEY, PHD

Detection and interpretation of adverse signals during preclinical and clinical stages of drug development inform the benefit-risk assessment that determines suitability for use in real-world situations. This review considers some recent signals associated with diabetes therapies, illustrating the difficulties in ascribing causality and evaluating absolute risk, predictability, prevention, and containment. Individual clinical trials are necessarily restricted for patient selection, number, and duration; they can introduce allocation and ascertainment bias and they often rely on biomarkers to estimate long-term clinical outcomes. In diabetes, the risk perspective is inevitably confounded by emergent comorbid conditions and potential interactions that limit therapeutic choice, hence the need for new therapies and better use of existing therapies to address the consequences of protracted glucotoxicity. However, for some therapies, the adverse effects may take several years to emerge, and it is evident that faint initial signals under trial conditions cannot be expected to foretell all eventualities. Thus, as information and experience accumulate with time, it should be accepted that benefit-risk deliberations will be refined, and adjustments to prescribing indications may become appropriate.

Development of a new pharmacotherapy typically follows a sequence of preclinical and clinical stages (Table 1). Completion of phase 3 marks the accumulation of clinical experience used to prepare an application for marketing authorization by a regulatory agency such as the Food and Drug Administration (FDA) or European Medicines Agency (EMA). This application should provide evidence for an adequately favorable benefit-risk balance for the intended use (1). Detection and interpretation of adverse properties during the development program are critical components of the authorization decision and guide the labeling and postmarketing obligations. The main safety questions raised about diabetes therapies are summarized in Table 2. This review examines the interpretation of adverse signals associated with recently approved blood glucose-lowering agents.

PREAPPROVAL CLINICAL TRIALS—The journey from molecule to medicine is likely to take at least 10 years for a new class of agent and to cost ~$500 million (1–3). If the many unsuccessful drug discovery studies and early development programs are factored in, the average approved drug probably reflects $1.3–1.8 billion of investment, so there is a strong incentive to secure marketing authorization once an agent has progressed through phase 3 (2,3). For a new diabetes medication, phase 3 usually takes ~3 years and consumes up to 90% of the total development costs; hence, an agent that is entered into phase 3 has invariably passed very thorough scrutiny of phase 2 efficacy and safety data and an assessment of marketing potential. Phase 3 customarily involves a minimum of 1,500 patients treated with the test agent (Table 3), generating 1,000–3,000 patient-years of exposure and a database from which to consider benefits and risks for a real-world setting (4,5).

DETERMINING BENEFITS AND RISKS—Although it is well recognized that early and sustained reductions of HbA1c markedly reduce microvascular complications and may contribute to reduced macrovascular risk, about one-half of diabetic patients still do not achieve or maintain sufficient glycemic control to avoid substantial long-term morbidity (6–10). New types of glucose-lowering agents are particularly required for type 2 diabetes because of the multivariable mix of genetic and environmental factors that conspire to create a progressive and highly heterogeneous natural history, with at least eight major organ systems being etiopathogenically implicated (11).

With regard to benefit, it is unlikely that an agent will be submitted for marketing authorization if it does not meet recognized approved efficacy criteria. The generally accepted efficacy surrogate is a reduction in HbA1c that is commensurate with the baseline HbA1c of the patient population studied (greater improvements of glycemic control expected if higher baseline hyperglycemia), taking into account the concomitant health issues within that population and additional benefits of the agent (12). The favorable–unfavorable boundary of benefits and risks will reflect whether efficacy is comparable with or better than existing therapies, achieved by a new mode of action that can be substituted for or add compatibly to these therapies, or assist subpopulations and comorbid conditions inadequately served by existing therapies. Set against this, we have the frequency and severity of apparent adverse effects, their predictability, potential avoidance, minimization or rectification, and practicalities of identifying and containing risk during real-world usage. Beyond this, theoretical adverse effects might be contemplated with regard to mode of action or other known pharmacodynamic or pharmacokinetic properties (4,5,12).

Preregistration trials are mostly designed and powered for a primary efficacy end point, customarily a reduction in HbA1c over 6 months at two or more dosage strengths (Table 3). As durability of efficacy receives more attention, extended randomized controlled trials are envisioned, with greater use of add-on rescue therapy to enable prolonged double blinding, and greater acceptance of post hoc analyses may be required. However, there will always be a desire for more drug exposure to assess risk, especially that...
Adverse signals in diabetes drug development

Table 1—Stages in the development of a new drug

<table>
<thead>
<tr>
<th>Preclinical stages</th>
<th>Clinical stages</th>
</tr>
</thead>
<tbody>
<tr>
<td>New chemical entity (NCE)</td>
<td>Investigational new drug (IND) application: permission to begin clinical studies</td>
</tr>
<tr>
<td>Preclinical phases</td>
<td>First administration to a small number of healthy human volunteers. Dose ranging, vital signs, pharmacodynamics, pharmacokinetics, drug interactions, safety</td>
</tr>
<tr>
<td>Preclinical pharmacology, mode of action, pharmacodynamics (activity, safety, tolerance), pharmacokinetics (bioavailability, distribution, metabolism, elimination), and toxicity in two or more mammalian species</td>
<td>Phase 1</td>
</tr>
<tr>
<td></td>
<td>First trials in small numbers of patients. Dose ranging, efficacy, further pharmacodynamics, pharmacokinetics and safety</td>
</tr>
<tr>
<td></td>
<td>Phase 2</td>
</tr>
<tr>
<td></td>
<td>Trials in larger numbers of patients. Multicenter trials, comparative trials with other treatments, efficacy, further pharmacodynamics and safety (including meta-analysis of CV outcomes)</td>
</tr>
<tr>
<td></td>
<td>Phase 3</td>
</tr>
<tr>
<td></td>
<td>New drug application (NDA): permission to market as a drug</td>
</tr>
<tr>
<td></td>
<td>Phase 4</td>
</tr>
<tr>
<td></td>
<td>Use in medical practice. Additional trials (similar to phase 3), postmarketing surveillance, adverse events reporting, use in special subgroups (e.g., elderly people)</td>
</tr>
</tbody>
</table>

pertaining to more vulnerable subpopulations. This raises the issues of preregistration meta-analyses of trial data and the role of postmarketing studies.

Relative risk against a placebo or comparator (or sometimes compared with a separate study of similarly disposed patients) provides a convenient indicator for interpreting frequency and severity of adverse signals, but absolute risk is perhaps more likely to interest the patient and prescriber. Figure 1 illustrates the

Table 2—Safety cautions and questions raised about current diabetes therapies

<table>
<thead>
<tr>
<th>Agent</th>
<th>Cautions and questions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Metformin</td>
<td>Lactic acidosis</td>
</tr>
<tr>
<td>Sulfonylureas</td>
<td>Hypoglycemia, CV risk</td>
</tr>
<tr>
<td>Thiazolidinediones</td>
<td>Aggravation of congestive heart disease and risk of fractures with rosiglitazone and pioglitazone; bladder cancer with pioglitazone; CV events with rosiglitazone</td>
</tr>
<tr>
<td>DPP-4 inhibitors</td>
<td>Pancreatitis</td>
</tr>
<tr>
<td>GLP-1 receptor agonists</td>
<td>Pancreatitis, thyroid C cells, heart rate/QTc</td>
</tr>
<tr>
<td>Insulin</td>
<td>Hypoglycemia, tumor growth</td>
</tr>
</tbody>
</table>

This table notes some safety cautions and questions raised with regard to current diabetes therapies. As noted in the text, some of these concerns are unproven or do not necessarily apply to all agents within a class. For example, data available at the time of writing have not supported an increased risk of pancreatitis with DPP-4 inhibitors or an increased risk of pancreatitis, thyroid C-cell cancer, or QTc changes with GLP-1 receptor agonists. Evidence associating sulfonylureas with increased CV risk remains equivocal, and data pertaining to pioglitazone and bladder cancer are open to variable interpretations. Consideration of benefit-risk balance for insulin therapy is discussed in the text.

Unfortunately, the quantification of hypoglycemic risk in clinical trials

HYPOGLYCEMIA—Treatment with antidiabetic drugs commonly produces mild hypoglycemic symptoms, especially with insulin, sulfonylureas, or combinations involving these classes (29). Signs of moderate or severe hypoglycemia and extra risks associated with patient unawareness or use in vulnerable groups, such as the elderly or renally impaired, are well rehearsed in product labeling and education packages. Nevertheless propensity for hypoglycemia forms an integral part of the calculation of acceptable risk for all diabetes pharmacotherapies (12). Where clinically significant risk is anticipated with combination therapy, this is typically minimized by recommending down-titration of the existing agent to coincide with the addition of the new agent. The interpretation of hypoglycemic events is usefully assisted by mode-of-action studies, particularly for agents that raise insulin concentration, to check whether the already-impaired counterregulatory capability in diabetic hypoglycemic states is further compromised (30).

Because diabetes is for life, comorbid conditions are rife, and therapies mostly last for much longer than the trial periods, the rigorous safety cautions for a new drug approval often require a comprehensive postmarketing risk management program to accompany standard pharmacovigilance monitoring (5). These can include small studies in special populations and crucial attention to issues of suspicion from preregistration signals, such as trends in biomarkers together with safety-first labeling restrictions. Accepting a caveat for the unpredictable, black box warnings can serve a valuable precautionary role to minimize misuse.
Table 3—Typical phase 3 development program of clinical trials for an oral blood glucose-lowering medication for type 2 diabetes

<table>
<thead>
<tr>
<th>Test agent</th>
<th>Versus</th>
<th>Duration (weeks)</th>
<th>Dosages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monotherapy in drug-naïve patients</td>
<td>Placebo</td>
<td>24</td>
<td>2–4</td>
</tr>
<tr>
<td>Add-on therapy in patients already treated with one or two glucose-lowering agents</td>
<td>Add-on placebo</td>
<td>24</td>
<td>1–4</td>
</tr>
<tr>
<td>Add-on therapy in patients already treated with one or two glucose-lowering agents</td>
<td>Add-on comparator glucose-lowering agent</td>
<td>24–104</td>
<td>1–2</td>
</tr>
<tr>
<td>Mono- or add-on therapy to existing therapy in patients with special requirements, e.g., renal or hepatic impairment</td>
<td>Placebo or comparator</td>
<td>12–24</td>
<td>1–4</td>
</tr>
</tbody>
</table>

These studies are normally prospective randomized double-blinded controlled trials, and the primary end point is likely to be a change in the HbA1c value compared with baseline, placebo, or comparator agent. Numbers will vary depending on whether superiority or noninferiority is sought but will typically be in the range 80–150 subjects per arm. Secondary end points are likely to include changes in fasting plasma glucose level and body weight. Studies involving add-on therapy to existing agents, especially insulin, may involve adjustments to the dosage of the existing agent before and/or during the trial. Safety data will be collected from all studies. Studies comparing injected agents may not be blinded, and extension studies may be unblinded or use strict rescue criteria to remain blinded.

Table: Risk of diabetes complications and adverse events associated with diabetes drug therapies expressed as absolute risk per 1,000 patient-years. Black bars show representative values for risk in diabetic patients, and gray bars indicate excess risk of events in patients treated with specific glucose-lowering medications relative to diabetic patients receiving comparator therapies. Incidence rates are subject to considerable variation with age, duration of diabetes, and concomitant morbidity. Adapted with permission from Bergenstal et al. (13). Data were derived from references 13–28. ACCORD, Action to Control Cardiovascular Risk in Diabetes; CHF, congestive heart failure; PVD, peripheral vascular disease; UKPDS, U.K. Prospective Diabetes Study.

Figure 1—Risk of diabetes complications and adverse events associated with diabetes drug therapies expressed as absolute risk per 1,000 patient-years. Black bars show representative values for risk in diabetic patients, and gray bars indicate excess risk of events in patients treated with specific glucose-lowering medications relative to diabetic patients receiving comparator therapies. Incidence rates are subject to considerable variation with age, duration of diabetes, and concomitant morbidity. Adapted with permission from Bergenstal et al. (13). Data were derived from references 13–28. ACCORD, Action to Control Cardiovascular Risk in Diabetes; CHF, congestive heart failure; PVD, peripheral vascular disease; UKPDS, U.K. Prospective Diabetes Study.

**WEIGHT CHANGES**—Although increased adiposity with insulin, sulfonylureas, and thiazolidinediones is sometimes overlooked as a risk factor, the escalation in vascular susceptibility that accompanies coexistent diabetes and obesity is a quantifiable and important component of overall risk (31). Additionally, rapid weight gain may signal issues other than adiposity, such as fluid retention with thiazolidinediones (32). The trade-off between improved glycemic control and adipose weight gain is a perennial treatment conundrum for which only individualized advice can be recommended, bearing in mind that dietary measures to counter the weight gain can aggravate susceptibility to hypoglycemia (33). Rapid initial weight loss during a clinical trial may lead to a detection bias for certain tumors, as discussed in a subsequent section of this review (34,35).

**CARDIOVASCULAR RISK**—Because cardiovascular (CV) diseases are highly prevalent and the most common cause of premature death among individuals with diabetes (Fig. 1), considerable attention has been directed toward CV risk assessment and minimization. However, being able to attribute causality to events that are not unexpected in the diabetic population is not a precise science (7). To illustrate the difficulties, we can reflect on the University Group Diabetes Program (UGDP) study of the 1960s. Although this brought to attention the lactic acidosis issue with phenformin (prompting further investigation that led to the first diabetes drug withdrawal by the FDA in 1978), it also initiated concern about possible detrimental CV effects of sulfonylureas (36). Despite a further 4
Adverse signals in diabetes drug development

decades of very extensive usage and endless clinical trials and database analyses, the overall balance of evidence is still equivocal (13). Previous accusations that insulin is atherogenic have since been dismissed, but they complicated risk assessments in a bygone era (37). They also remind us that treatments that sustain life expectancy among patients in poor general health can result in emergent disease-related morbid conditions that should not be confused with iatrogenic morbid conditions.

Interpreting adverse CV signals was brought to the fore with rosiglitazone, which after almost 10 years of postmarketing experience, created sufficiently ambiguous adverse event data that the drug was withdrawn in Europe, but the FDA settled for tighter labeling (38,39). Although relevant CV event data were available in the GlaxoSmithKline clinical study register website and received updated evaluation by regulatory agencies (40), it was a meta-analysis in the New England Journal of Medicine that attracted attention (26). With data from 42 randomized trials, excluding six studies in which there were no CV events, this analysis reported that rosiglitazone was associated with an increased odds for myocardial infarction (MI) (odds ratio 1.43, 95% CI 1.03–1.98) and CV death (1.64, 0.98–2.74). Expressed as events/total patients for rosiglitazone versus control groups, these odds ratios corresponded to 86/14,372 versus 72/11,634 for MI and 39/10,936 versus 22/9,509 for CV deaths. The analyses have been repeated and debated extensively elsewhere (41,42) and will not be reiterated further here, but some generic lessons for interpreting signals during development programs are highlighted. When the FDA advisory committee initially deliberated the new drug application for rosiglitazone, there were data on CV events from five randomized trials. The occurrence of MI was 6/1,967 and 3/793 for the rosiglitazone and control groups, respectively, corresponding to 0.30 and 0.37%, respectively. The occurrence of CV deaths was 2/1,967 and 0/793 for the rosiglitazone and control groups, respectively. So there was no obvious signal, and biomarker data tended to be positive, although fluid retention, edema, risk of heart failure, and weight gain were dutifully considered. At the time, attention was distracted toward the liver because another thiazolidinedione, troglitazone, was in the spotlight for idiosyncratic hepatotoxicity (43). With regard to the approval of thiazolidinediones, Europe evaluated these agents a little later and made approvals for second-line indication (unless metformin was not appropriate) and excluded New York Heart Association class I–IV (vs. first-line approval and exclusion of New York Heart Association class III and IV by the FDA).

The rosiglitazone experience prompted the FDA to issue new guidance about CV risk (44). This requires a meta-analysis of important CV events in phase 2/3 to achieve an upper 95% CI of <1.2 to qualify for approval without requiring a postmarketing CV trial, provided that overall benefit and risk support approval (Fig. 2). If the upper 95% CI is >1.8, additional phase 3 safety studies are required before resubmission for marketing authorization. If the upper CI lies between 1.3 and 1.8 and approval is otherwise appropriate, then a postmarketing CV events study generally will be necessary and required to show an upper 95% CI of <1.3. In practice, each sponsor of recently approved drugs has elected to undertake such a study (or have been encouraged to do so by the FDA), even if the phase 2/3 CV events conform to an upper 95% CI <1.3. Indeed, such postmarketing studies appear to be almost obligatory because no sponsor company would wish for its product to be disadvantaged in years to come if its competitors can claim hard end point CV data from large purpose-designed trials. These studies, which are currently ongoing, are summarized in Table 4 (45). Although they are mostly event driven, differences in estimated event rates and duration, selection of major CV events, power calculations, types of statistics (superiority vs. noninferiority), and use of placebo, an active comparator, or both make direct comparisons between studies difficult.

Effective use of antihypertensive and lipid-controlling medications has reduced the rates for major adverse CV events among diabetic patients (46), which in turn will affect the spread of CIs in an analysis of CV events, especially in phase 2/3 studies, where the hypothetical wonder drug with no events would defy the statistical analysis. To boost event rates, the recruitment of extra patients at high CV risk may alter the primary–secondary preventive balance of the risk assessment. The importance of time in the generation of CV events must also be borne in mind because events occurring early in trials may more strongly reflect the accumulated risk in the months and years before trial entry rather than reflect the initial effects of the test drug. For example, in a more recent analysis of the rosiglitazone CV event data, it is evident that short-term trials give very different and far-worse outcomes than longer-term studies (47). Moreover, several large trials of glycemic intensification have found that CV event rates are initially worse before they subsequently show benefit (48,49), whereas others suggest a possible association between CV events and the extent and speed of intensification (7).

Dealing with CV changes that are not necessarily included as major adverse cardiac events requires a different approach and is illustrated for some GLP-1 receptor agonists, which tend to slightly increase heart rate (sharp cardiological intake of breath) but decrease blood pressure slightly (possible cardiological sigh of relief). GLP-1 receptors were identified in the vasculature and myocardium, so small bespoke safety studies

---

**Figure 2**—FDA guidance for industry for a meta-analysis of phase 2/3 data to evaluate CV risk with new glucose-lowering therapies to treat type 2 diabetes. Based on ref. 44. *If overall benefit-risk analysis of the drug supports approval. RR, relative risk.*
Table 4—Postmarketing CV outcome studies with diabetes therapies

<table>
<thead>
<tr>
<th>Trial</th>
<th>Test agent</th>
<th>Started</th>
<th>Years</th>
<th>No. pts</th>
<th>Primary end point</th>
</tr>
</thead>
<tbody>
<tr>
<td>TECOS</td>
<td>Sitagliptin</td>
<td>2008</td>
<td>~4</td>
<td>14,000</td>
<td>Composite of CV death, NFMI, NFS, Angh</td>
</tr>
<tr>
<td>EXAMINE</td>
<td>Alogliptin</td>
<td>2009</td>
<td>4.75</td>
<td>~5,400</td>
<td>Composite of CV death, NFMI, NFS</td>
</tr>
<tr>
<td>CANVAS</td>
<td>Canagliflozin</td>
<td>2009</td>
<td>~4</td>
<td>~4,400</td>
<td>Composite of CV death, NFMI, NFS</td>
</tr>
<tr>
<td>ACE</td>
<td>Acarbose</td>
<td>2009</td>
<td>~4</td>
<td>~7,500</td>
<td>Composite of CV death, NFMI, NFS</td>
</tr>
<tr>
<td>LEADER</td>
<td>Liraglutide</td>
<td>2010</td>
<td>~5</td>
<td>~9,400</td>
<td>Composite of CV death, NFMI, NFS</td>
</tr>
<tr>
<td>EXCEL</td>
<td>Exenatide QW</td>
<td>2010</td>
<td>~5.5</td>
<td>9,500</td>
<td>Composite of CV death, NFMI, NFS</td>
</tr>
<tr>
<td>SAVOR-TIMI 53</td>
<td>Saxagliptin</td>
<td>2010</td>
<td>~4</td>
<td>16,500</td>
<td>Composite of CV death, NFMI, NFS</td>
</tr>
<tr>
<td>CAROLINA</td>
<td>Linagliptin</td>
<td>2010</td>
<td>8</td>
<td>6,000</td>
<td>Composite of CV death, NFMI, NFS</td>
</tr>
<tr>
<td>ELIXA</td>
<td>Lixisenatide</td>
<td>2010</td>
<td>~4</td>
<td>~6,000</td>
<td>Composite of CV death, NFMI, Angh</td>
</tr>
<tr>
<td>NCT01131676</td>
<td>Empagliflozin</td>
<td>2010</td>
<td>~4</td>
<td>~7,000</td>
<td>Composite of CV death, NFMI, NFS</td>
</tr>
<tr>
<td>NCT01042769</td>
<td>Alogliptar</td>
<td>2010</td>
<td>~4.5</td>
<td>~7,000</td>
<td>Composite of CV death, NFMI, NFS</td>
</tr>
<tr>
<td>REWIND</td>
<td>Dulaglutide</td>
<td>2011</td>
<td>~6.5</td>
<td>~9,600</td>
<td>Composite of CV death, NFMI, NFS</td>
</tr>
<tr>
<td>DECLARE-TIMI 58</td>
<td>Dapagliflozin</td>
<td>2013</td>
<td>~6</td>
<td>17,150</td>
<td>Composite of CV death, NFMI, NFS</td>
</tr>
</tbody>
</table>

Some studies started during phase 3. All are randomized and double blind in patients with type 2 diabetes receiving usual care plus test agent and compared with placebo (or to comparator in CAROLINA). Recruited patients include a high proportion with high CV risk (e.g., history of CV events) or, specifically, recently experienced an acute coronary syndrome event. The primary end point is a composite of (time to) major cardiovascular events. ACE, Acarbose Cardiovascular Evaluation Trial; Angh, unstable angina requiring hospitalization; CANVAS, Canagliflozin Cardiovascular Assessment Study; CAROLINA, Cardiovascular Outcome Study of Linagliptin Versus Glimipride in Patients with Type 2 Diabetes; DECLARE-TIMI38, Multicenter Trial to Evaluate the Effect of Dapagliflozin on the Incidence of Cardiovascular Events; ELIXA, Evaluation of Cardiovascular Outcomes in Patients with Type 2 Diabetes After Acute Coronary Syndrome During Treatment with AVE0010 (Lixisenatide); EXAMINE, Examination of Cardiovascular Outcomes: Alogliptin versus Standard of Care; EXCEL, Exenatide Study of Cardiovascular Event Lowering; LEADER, Liraglutide Effect and Action in Diabetes: Evaluation of Cardiovascular Outcome Results; NCT, clinical trials identifier number; NFMI, nonfatal MI; NFS, nonfatal stroke; REWIND, Researchers Cardiovascular Events with a Weekly Incretin in Diabetes; SAVOR-TIMI 53, Saxagliptin Assessment of Vascular Outcomes Recorded in Patients with Diabetes Mellitus; TECOS: Trial to Evaluate Cardiovascular Outcomes with Sitagliptin.

were warranted. These studies focused on corrected QT (QTc) interval duration, which reassuringly showed no clinically significant alterations (50,51).

CANCER—Whereas CV events are common in diabetes, cancers are uncommon or rare but modestly increased overall by ~40% (52). Detection is often delayed, and the attributing cause is complicated by familial susceptibility; present and prior obesity; history of exposure to carcinogens, including smoking; and comorbid conditions. Allocation factors can also confuse the evidence surrounding malignancies; for example, an uneven randomization for ethnicity, gender, socioeconomic status, and educational or geographical background can significantly influence the occurrence of cancers. Many patients who enter clinical trials will not have previously received the type of vigilant, frequent attention to ascribe to a potential danger signal. Reporting rates are variable for many reasons, events are not confirmed, and confounding factors may not be taken into account (5,58).

An appreciation that pancreatitis presents a risk for pancreatic cancer (59), coupled with an animal study suggesting that incretins might cause pancreatic duct cell proliferation, sparked interest in a potential link between incretins and pancreatic cancer (55,60). Additional studies in several animal species could not confirm the ducal proliferation (61), but the FDA database showed more cases of pancreatic cancer in patients who had received certain incretins compared suggesting that the combination of weight loss and attentive clinical care in a trial setting conferred a detection bias for breast cancer—a valuable benefit for the patients but a substantial delay for drug development (34,35).

Pancreatic cancer

Deciding the proportionate level of caution to ascribe to a potential danger signal is never straightforward. Take for example the evidence for a putative link between incretin therapies (GLP-1 receptor agonists and DPP-4 inhibitors) and pancreatitis, pancreatic ductal metaplasia, and pancreatic cancer (28,53–56). Reports of cases of acute pancreatitis in patients receiving incretin therapies generated awareness, which may have prompted further similar reports and an accumulation of cases in the pharmacovigilance databases (57,58). However, the diagnosis of pancreatitis was not always confirmed; severity was highly variable; and although the condition is known to be more common among diabetic patients than among the nondiabetic population, estimated incidence rates have varied widely from <1 to ~5 per 1,000 patient-years (13,17,57). Given a need to await additional retrospective interrogation of large databases, the labeling advice to discontinue incretins in patients where pancreatitis is suspected would seem proportionate.

It is relevant to note here that pharmacovigilance is an ongoing process to monitor side effects, but a reported event is not necessarily caused by the medicines to which the report relates and should not be interpreted as meaning that a medicine is unsafe to use (57,58). Spontaneous reporting of suspected side effects should not be used as a basis for estimating incidence rates because they provide a numerator without knowing the denominator. Reporting rates are variable for many reasons, events are not confirmed, and confounding factors may not be taken into account (5,58).

Bailey
with other classes of diabetes drugs (62). The problem of attempting numerical analyses from spontaneous adverse event reporting has already been noted, but could this be a signal? Pancreatic cancer is rare at <0.1 per 1,000 patient-years. It is slow to develop to detection, and unlike other cancers, its occurrence tends to decrease with time after diagnosis of diabetes (62). So, numbers and time may preclude an early resolution from post-marketing trials. In general, clinical experience to date does not appear to support an association of incretins with pancreatic cancer (63,64), but the issue has emphasized that procedures for extrapolating preclinical signals to clinical situations and projecting possible clinical signals into future clinical events are still tenuous.

**Thyroid cancer**

An example of scrambled signals is that of a purported link between liraglutide and thyroid C-cell medullary carcinoma. Thyroid C cells of rats and mice express high concentrations of GLP-1 receptors and respond to high concentrations of GLP-1 receptor agonists with excess calcitonin secretion and C-cell hyperplasia (65). In clinical studies, liraglutide slightly raised calcitonin concentrations, but human thyroid C cells have a low expression of GLP-1 receptors, and calcitonin concentrations were not elevated to an extent that is considered a signal for medullary C-cell cancer (66). Bearing in mind that such cancers are rare, ~600–1,000 annually in the U.S., a clinical trial is not an option. Sensibly, the FDA explained its reasoning in a high-profile journal, a move that was similarly useful with regard to metformin and lactic acidosis some 15 years ago (67).

**Bladder cancer**

Interpreting preclinical and clinical signals suggestive of a bladder cancer risk with pioglitazone is still difficult after more than one decade of accumulated evidence. Animal studies suggested a metabolite that could irritate the urothelium and cause localized hyperplasia (68), and a major clinical trial (PROactive [Prospective Pioglitazone Clinical Trial in Macrovascular Events]) noted an increased occurrence of bladder tumors (25), but no excess was observed after a 6-year observational follow-up (69). Several meta-analyses based around five large cohort studies totaling >2.3 million diabetic patients have shown a small increased risk of bladder tumors (1.17–2.22 fold), with a slightly higher risk (1.3–1.4 fold) in those exposed for >2 years (70,71). Given a bladder cancer incidence of ~0.5 per 1,000 patient-years in diabetes, these studies indicated an increased risk of ~0.1 per 1,000 patient-years. However, within these studies, there were difficulties in assessing prior exposure to key risk factors, including smoking, the time taken for these cancers to develop, and variability in detection. Additionally, the risk of some other cancers may have been reduced in patients receiving pioglitazone, leaving no apparent overall cancer imbalance (72). Some countries in Europe decided to discontinue pioglitazone, but the EMA has not endorsed this, preferring to modify the label to recommend avoidance of the drug in patients with active bladder cancer and in those at high risk (73). Thus, we have another example of varying interpretations by different expert groups assessing the same information and weighing the extent of adverse signals and events against the benefits.

Another cancer question has been raised for the sodium glucose cotransporter-2 inhibitor dapagliflozin, which was recently approved in Europe and Australia. Although there was no evidence of mutagenicity or carcinogenicity in preclinical trials and no overall imbalance in malignancies during the development program, there were increases in breast, prostate, and bladder cancers (74). Given the weight loss associated with this agent, the detection bias discussed for orlistat may apply, and the prostate cancers (mostly well advanced) were generally discovered early in the trials when scrutiny of urinary phenomena brought to attention hematuria and urination difficulties. A history of hematuria recently before or at randomization was recorded for 7 of 10 cases of bladder cancer; these cancers were mostly detected too early and were too advanced to have originated after treatment with dapagliflozin. Whether persistently high glucose in the urine could promote growth of preexisting tumors is not entirely excluded, but extensive preclinical studies have not shown a signal, and patients with familial renal glucosuria do not appear to experience detrimental effects of lifelong glucosuria (75).

**Insulin and cancer**

The evidence with regard to insulin and cancer is far too extensive and controversial to dissect here (52,76–79). There are so many confounders in most of the studies that it is possible to question and criticize any interpretation. Thus are the shaky foundations upon which adverse events have to be assessed. Insulin is not of itself carcinogenic, but through its interactions with the different isoforms of the insulin receptor and the IGF1 receptor, it stands under suspicion of promoting tumor growth. By way of summary and excluding less solid evidence, it appears that replacement amounts of human insulin do not increase risk. However, this conclusion cannot yet be extended to the massive pharmacological doses needed for highly insulin resistant individuals who likely carry an abundance of other risks and for whom alternative glucose-lowering therapies are not available. Analog insulins in present use do not appear to be any less safe in this context, as most recently reported by the ORIGIN (Outcome Reduction with Initial Glargine Intervention) trial (80). For the record, most circulating glargine is dearginated to a low-affinity binding metabolite (M1) before it meets tissue receptors (79,81). Also recall that the highest endogenous exposure to insulin is in the liver, and this is likely to have been exaggerated during a hyperinsulinemic prediabetes period, whereas subcutaneous administration of insulin alters the disposition of insulin in favor of other tissues. A consensus report of the American Diabetes Association and American Cancer Society reasonably advises that on the balance of current evidence, cancer risk should not be a major factor in choosing between available diabetes therapies for the average patient, although patients with a very high risk for cancer occurrence or recurrence may require more careful consideration (82).

**Bone fractures**—Thiazolidinediones have been associated with declining bone mineral density and increased risk of bone fractures, especially in older women (13,83–85). Although thiazolidinediones continue to receive careful clinical investigation, they provide a further example of the importance of preclinical mechanistic studies. In vitro and animal data have shown effects of peroxisome proliferator-activated receptor-γ (PPARγ) agonists on the differentiation pathways of various cell types populating bone (86). Although the various biomarkers of bone health can be used effectively for diagnosis and therapeutic monitoring, there are considerable
limitations to their predictive value for the individual, and interpretation of small changes during preregistration trials remains uncertain (87,88).

CONCLUSIONS—A particular challenge when interpreting adverse signals in clinical trials is the need to extrapolate to a real-world environment where patient populations may be more diverse, prescribers less specialized, and monitoring less conscientious. Faint signals from preregistration trials can take a decade or more to reveal their clinical consequences and often are unpredictable and confounded by allocation and detection bias, prior exposure to risk factors (known and unknown), and limitations of time and numbers.

The need to revise the label or even withdraw a drug should not be viewed as a failure of foresight but, rather, as the mark of a vigilant and responsive regulatory process. To minimize risk, new medicines have explicit exclusions reinforced with temporary (and sometimes permanent) black box warnings that permit availability to defined patient populations while allowing time for separate controlled studies with more vulnerable or different subpopulations. Mounting pressures to make all clinical trial information available and all databases accessible to everyone will introduce opportunities for further misinterpretation. Transparency requires responsibility, but benefit and risk often are a judgment of probability, and revisions have to be accepted without prejudice as new information emerges. The teachings of the 16th century Swiss-German physician Paracelsus remind us that no chemical is absolutely safe, and the difference between a medicine and a poison can be the dose.

Acknowledgments—C.J.B. has undertaken ad-hoc consultancy for Bristol-Myers Squibb, AstraZeneca, Merck Sharp & Dohme, Novo Nordisk, sanofi-aventis, Janssen, Eli Lilly, Roche, and Takeda; delivered continuing medical education programs sponsored by Bristol-Myers Squibb, AstraZeneca, Glaxo-SmithKline, Merck Serono, Merck Sharp & Dohme, Eli Lilly, and Boehringer Ingelheim; and received travel or accommodation reimbursement from AstraZeneca and Bristol-Myers Squibb. No other potential conflicts of interest relevant to this article were reported.

C.J.B. was sole author, researched the data, and prepared the manuscript.

This article is based on a State-of-the-Art lecture delivered at the 72nd Scientific Sessions of the American Diabetes Association, Philadelphia, Pennsylvania, 8–12 June 2012.

C.J.B. served as an expert witness for the FDA, EMA, and Medicines and Healthcare Products Regulatory Agency (MHRA), providing benefit-risk assessments and testimony. He also served as a representative of the European Association for the Study of Diabetes to EMA. His research has contributed to the development of diabetes and obesity medicines with preclinical and clinical studies, and he is well known for his early work on metformin.

References
Adverse signals in diabetes drug development

22. Hampton T. Diabetes drugs tied to fractures in women. JAMA 2007;297:1645
43. Bailey CJ. The rise and fall of troglitazone. Diabetes Care 2010;33:428–433
58. European database of suspected adverse drug reaction reports [Internet], 2012.


77. Gerstein HC. Does insulin therapy promote, reduce, or have a neutral effect on cancers? JAMA 2010;303:446–447


