DOI: 10.1111/dme.13854

Review Article

A systematic review of observational studies of the association between pioglitazone use and bladder cancer

E. Ripamonti[1], L. Azoulay[2,3], M. Abrahamowicz[2,4], R.W. Platt[2,3] and S. Suissa[2,3]

[1]IRCCS Fondazione Don Carlo Gnocchi, Milan, Italy, [2]Department of Epidemiology, Biostatistics and Occupational Health, McGill University, [3]Centre for Clinical Epidemiology, Lady Davis Research Institute, Jewish General Hospital and [4]Centre for Clinical Epidemiology, Royal Victoria Hospital, Montreal, QC, Canada

Accepted 29 October 2018

Abstract

Aim To conduct a systematic review of all observational studies on the effect of pioglitazone on the risk of bladder cancer.

Methods The MEDLINE and EMBASE databases were queried for papers published between 1 January 2000 and 30 October 2017. We took into consideration observational studies (both retrospective and prospective) that included participants with Type 2 diabetes prescribed anti-hyperglycaemic drugs.

Results While some studies reported an association, others did not, and meta-analyses of these studies showed a significantly increased risk; however, while meta-analysis is a powerful and practical statistical tool, its results should be considered with caution when applied to widely heterogeneous studies. We describe how many of these studies are affected by different types of bias, most notably time-related biases, which should preclude a pooled analysis that would result in biased estimation of the risk.

Conclusions Given existing data, it is not appropriate to pool the outcomes of highly heterogeneous studies and further rigorously conducted observational research is needed to clarify the role of pioglitazone use on the incidence of bladder cancer.

Diabet. Med. 36: 22–35 (2019)

Introduction

Pioglitazone is a drug of a class of chemical agents known as thiazolidinediones (or glitazones), which have been available for about 20 years. It is currently used as a second-line therapy for Type 2 diabetes, and has long been suspected to be associated with bladder cancer. Much of the epidemiological work in this area is relatively recent, having been conducted in the last decade only; this research has led to controversy and there is an ongoing debate on this issue involving scholars, policy-makers, and regulatory agencies.

In 1997, the market for oral anti-hyperglycaemic drugs was significantly renewed by the launch of troglitazone, the precursor of thiazolidinediones. Two years later, in 1999, two other drugs of this class were put on the market, namely rosiglitazone and pioglitazone. In 2000, the US Food and Drug Administration (FDA) withdrew troglitazone from the market because of hepatotoxicity and, in 2010, the FDA restricted use of rosiglitazone and the European Medicines Agency suspended it from the market because of safety concerns (heart failure). As of 2018, therefore, pioglitazone is the only glitazone available in most medium-/high-income countries for oral treatment of Type 2 diabetes.

Ten previous systematic reviews or meta-analyses on the association of pioglitazone use with bladder cancer have already been published [1–10]. Nevertheless, a final word on risk associated with pioglitazone use cannot be given because these studies are flawed by factors that range from improper use of analytical procedures to scarce consideration of time-related and selection biases. Indeed, time-related bias is a common source of error in pharmacoepidemiological studies that has a direct impact on the effect estimates. The different types of time-related bias include immortal time bias [11], time-window bias [12] and time-lag bias [13].

The aim of the present paper is to provide an up-to-date systematic review of current evidence on the role of pioglitazone as a risk factor for bladder cancer. Particular emphasis is put on the methodological choices of the authors, which could partially explain differences in their results.

Correspondence to: Robert Platt. E-mail: robert.platt@mcgill.ca

What's new?

- It is unclear from the literature whether or not pioglitazone is associated with bladder cancer.
- We addressed this issue, conducting a systematic review of observational studies.
- We found that some studies reported an association, others did not, but meta-analyses of these studies detected a significant risk.
- We show how the clinical relevance of most research studies in this area is hampered by the presence of bias in their design, most notably time-related bias.
- Further studies are needed to clarify the association, and pioglitazone should be prescribed with caution.

Methods

Eligibility criteria

We considered observational studies (cohort, case–control and hybrid designs; prospective or retrospective) involving people with Type 2 diabetes mellitus prescribed anti-hyperglycaemic drugs. To be eligible for the present review, the studies had to: (1) have taken into account participants exposed/unexposed to pioglitazone; (2) have used bladder cancer as an outcome, and (3) have recorded outcomes as counts, or as relative risk, odds ratio and hazard ratio, with a suitable measure of variance. Studies listing more than one cancer outcome, or investigating more than one thiazolidinedione as the exposure were also eligible for inclusion provided they included pioglitazone and bladder cancer.

Information source

We searched the EMBASE and MEDLINE databases systematically for papers published between 1 January 2000 and 30 October 2017, combining the keywords bladder cancer + pioglitazone, bladder cancer + thiazolidinedione, and cancer + pioglitazone, and evaluating the same combinations in the titles of the articles. We examined pertinent papers cited in the references of selected manuscripts, relevant reviews and previous meta-analyses, and we completed the retrieval of information checking the same keywords in Google Scholar. Only original papers published in English were considered. Two independent reviewers (E.R. and L.A.) searched papers eligible for the present review by screening the titles and the abstract. Disagreement was resolved by consensus or, when necessary, by a third independent evaluator (R.P.).

Data collection and reporting

From each paper, reviewers extracted information on study design, general setting, sample size, exposure definition, follow-up duration, comparator drugs, baseline and demographic characteristics of the participants. Selected baseline characteristics included age, sex, ethnicity, presence of other bladder conditions, urinary tract infection, smoking, alcohol use, obesity, duration of diabetes, and presence of other medical conditions. Outcomes are presented as count data, and adjusted risk estimates were obtained for each paper, as well as the covariates included in the multivariable models. In case a paper reported more than one outcome estimate, we turned to the model that was adjusted for the most potential confounders. The study was reported in accordance with the recommendations set out by the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) working groups [14]. The ROBINS-I tool [15] was used to detect sources of bias and confounding; a detailed protocol is available from the authors on request. The analyses presented here are mainly focused on time-related issues, which represent the most critical area of bias in this context.

Statistical analysis

To evaluate heterogeneity we calculated I^2 statistics [16] (using the Sidik–Jonkman estimator). We also employed funnel and radial plots. Funnel plots show on the horizontal axis the observed effect measure (y_i) against the corresponding standard error ($\sqrt{v_i}$) on the vertical axis. In radial plots the inverse of each standard error properly adjusted for τ^2, i.e. a constant indicating residual heterogeneity among the true effects that is estimated by random effects modelling ($x_i = 1/\sqrt{v_i + \tau^2}$) on the horizontal axis, is projected on the vertical axis into the observed effect measure standardized by its corresponding standard error (i.e. $z_i = y_i/\sqrt{v_i + \tau^2}$; see Viechtbauer [17]). For each individual study we estimated a dose–response model in which the log-linear relation assumption was relaxed [18]. The variables indicating exposure definition were transformed, introducing restricted cubic spline models with four knots (corresponding to the 5th, 35th, 65th and 95th percentiles of the exposure distribution) using the RSM package [19] of R software. Such variables, together with exposure in the original metric, were included as regressors in a fixed-effects non-linear regression model for dose–response association; because error terms were not independent, the variance–covariance matrix was approximated [20]. Analyses were conducted using the METAFOR [17] and DOSRESMETA [21] R packages.

Results

Search results and study inclusion

A total of 180 publications were initially retrieved from the EMBASE/MEDLINE databases. Of these, 114 were excluded as these were abstracts, conference proceedings, research reports, commentaries, letters, replies or brief reports only, 37 were discarded for not considering bladder cancer as an

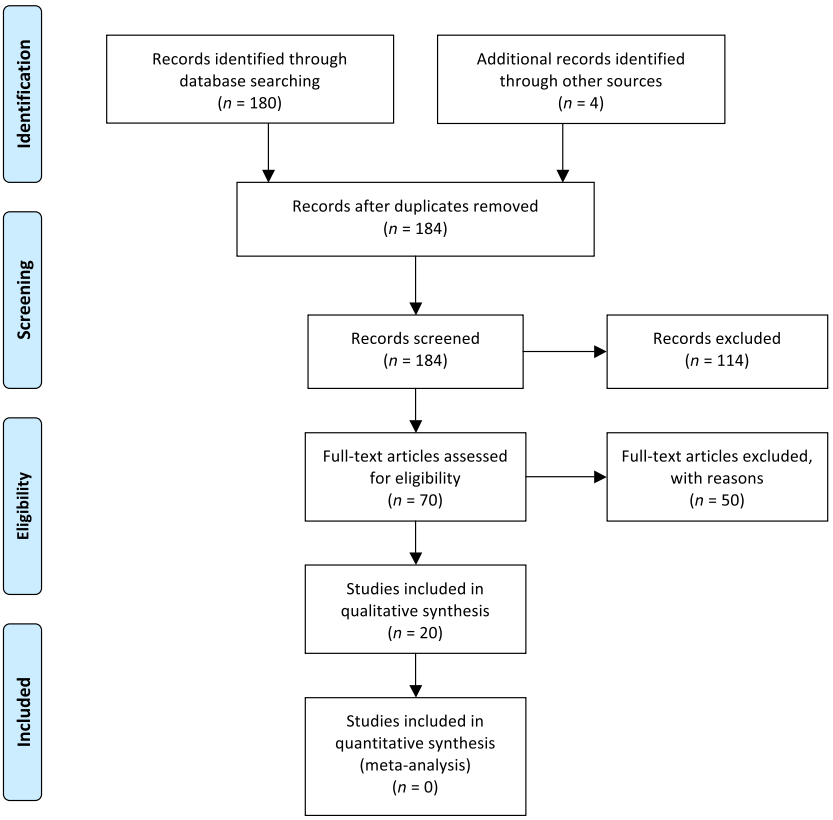

FIGURE 1 Flowchart of the studies included in the systematic review.

outcome, six referred to previous reviews/meta-analyses, and two did not focus on pioglitazone. This left a total of 21 studies the full text of which was downloaded and analysed by the examiners. Four of these studies did not meet eligibility criteria, and one paper was a descriptive adverse event reporting investigation (Fig. 1). Four additional papers were selected through our hand search of references from published studies, relevant reviews and previous meta-analyses. A final total of 20 studies (all published between 2011 and 2018) were thus included in the present review.

Descriptive findings

Out of the 20 included studies, 15 used a cohort approach [22–36]. The study by Lewis *et al.* [28] presents a re-analysis of data from a previous article [30]. Fourteen of the studies were conducted using electronic/administrative databases, while only one study referred to in-hospital data [35]. Two papers reported the findings of international multi-population studies [32,37]. The total sample size varied from 21 335 to 1 491 060 people and the follow-up period ranged from 40 to 180 months (mean 109.2 months; median 120 months). Nine studies considered the effect of differential cumulative exposure and six considered the effect of differential cumulative dose. In the primary analysis, adjusted for the greatest number of putative confounders, a risk effect of pioglitazone was obtained in four out of 12 studies [24,26,31,34] (Table 1 and Fig. 2).

We included six nested case–control studies [29,33,38–41], conducted using administrative/electronic records. We also found one non-nested case–control study [42], conducted using an in-hospital database, but we excluded this from the analyses as this type of design is likely to present biases. The total sample size ranged from 113 193 to 606 683 people and the follow-up period ranged from 55 to 180 months. Two studies showed results for differential cumulative exposure [38,40] and one study for differential cumulative dose [38]. In the primary analysis, a risk effect of pioglitazone was reported in two studies [38,40] (Table 2 and Fig. 2).

Assessment of heterogeneity

To diagnose the presence of heterogeneity, we calculated I^2 statistics (with the Sidik–Jonkman estimator), which were 51.61% for cohort studies and 11.58% for case–control studies. These measures indicate a low degree of heterogeneity for case–control studies and a moderate degree of heterogeneity for cohort studies [43]. In addition, funnel and radial plots showed that both cohort studies and case–control studies were characterized by a certain amount of variability in the effect measure estimates, which was more pronounced in the case of cohort studies (Fig. 3).

Table 1 Cohort studies included in the review

Study	Database	Ntot	Exposed	Unexposed	Follow-up, months	Cumulative exposure, months	Cumulative dose, mg	HR (95% CI)
Lewis *et al.* [30]	KPNC	193 099	30 173	162 926	136	<12		0.8 (0.6–1.3)
						12–24		1.4 (0.9–2.1)
						>24		1.4 (1.03–2.0)
							1–10 500	1.0 (0.75–1.5)
							10 501–28 000	1.2 (0.8–1.8)
							>28 000	1.4 (0.96–2.1)
Mamtani *et al.* [27]	THIN	52 296	10 900	41 396	120			1.14 (0.79–1.66)
						12–24		1.40 (0.40–4.90)
						25–36		4.21 (1.41–12.06)
						37–48		2.94 (0.85-10.04)
						49–60		5.15 (1.52–17.4)
						>60		6.97 (2.28–21.3)
Neumann *et al.* [26]	SNIIRAM-Fr	1 491 060	155 535	1 335 525	40			1.22(1.05-1.43)
						<12		1.05 (0.82–1.36)
						12–23		1.34 (1.02–1.75)
						>24		1.36 (1.04–1.79)
							<10 500	1.12 (0.89–1.40)
							10 500–27 999	1.20 (0.93–1.53)
							>28 000	1.75 (1.22–2.50)
Tseng *et al.* [25]	Taiwan	54 928	2545	52 383	48			1.15 (0.60–2.19)
						<12		1.54 (0.72–3.26)
						>12		0.81 (0.19–3.34)
							1–10 500	1.45 (0.68–3.06)
							>10 500	0.93 (0.22–3.84)
Fujimoto *et al.* [35]	In-hospital	21 335	663	20 672	144			1.75 (0.89–3.45)
Vallarino *et al.* [23]	InVision	56 535	38 588	17 948	120			0.92 (0.63–1.33)
Wei *et al.* [22]		207 714	23 548	184 166	108			1.16 (0.83–1.62)
						>12		1.12 (0.76–1.64)
						>24		1.20 (0.74–1.93)
Jin *et al.* [33]	Korea	113 193	11 240	101 953	68			1.13 (0.77–1.67)
Lee *et al.* [31]	Taiwan	34 970	3497	31 473	60			1.59 (1.32–1.91)
						<12		1.68 (1.36–2.08)
						>12		1.39 (1.09–1.76)
							1–10 500	1.69 (1.37–2.07)
							>10 500	1.34 (1.04–1.73)
Lewis *et al.* [28]	KPNC	193 099						1.00 (0.82–1.22)
						<18		0.69 (0.51–0.93)
						18-48		1.05 (0.79–1.38)
						>48		1.26 (0.88–1.78)
							1–13 000	0.78 (0.59–1.05)
							13 001–35 000	0.89 (0.65–1.22)
							>35 000	1.21 (0.88–1.65)
Levin *et al.* *[36]	Scotland	135 306						1.36 (0.73–2.54)
	Finland	215 757						0.59 (0.29–1.20)
	British Columbia	82 280						1.63 (0.73–3.66)
	CPRD	84 922						1.34 (0.81–2.21)
	Manchester	6456						0.53 (0.02–14.38)
Lewis *et al.* [29]	KPNC	193 099	34 181	158 918	180			1.06 (0.89–1.26)
						<18		0.88 (0.68–1.16)
						18–48		1.03 (0.80–1.33)
						>48		1.16 (0.87–1.54)
							<14 000	0.90 (0.69–1.16)
							14 000–40 000	1.10 (0.85–1.42)
							>40 000	1.07 (0.79–1.44)
Tuccori *et al.* [24]	CPRD	145,806	921	142 758	163			1.63 (1.22–2.19)
						<12		1.33 (0.73–2.40)
						12–24		1.66 (0.97–2.84)
						>24		1.78 (1.21–2.64)
							<10 500	1.63 (1.02–2.60)
							10 500–28 000	1.58 (0.98–2.55)
							>28 000	1.70 (1.04–2.78)
Korhonen *et al.* [32]	Finland	373 446	55 337	317 109	115.2			0.99 (0.75–1.30)
	Sweden				60			

Table 1 (*Continued*)

Study	Database	Ntot	Exposed	Unexposed	Follow-up, months	Cumulative exposure, months	Cumulative dose, mg	HR (95% CI)
	Netherlands				115.2			
	UK (general practice)				139.2			
						<18		1.10 (0.82–1.48)
						18–48		0.78 (0.52–1.19)
						>48		0.86 (0.44–1.66)
							1–14 000	1.05 (0.77–1.42)
							14 001–40 000	0.99 (0.66–1.46)
							>40 000	0.65 (0.33–1.26)
Garry *et al.* [34]	Medicare	247 356	38 700	208 656	96			1.32 (1.02–1.70)
						<24		1.32 (0.98–1.78)
						>24		1.29 (0.76–2.18)

CPRD, Clinical Practice Research Datalink; KPNC, Kaiser Permanente Northern California; HR, hazard ratio.
*Estimates are expressed as risk ratio (Poisson regression models).
Notice that in the study by Levin *et al.* [36] HR estimates are reported separately for men and women, and here we have only reported those for men (no risk also emerged in the samples of women).

Impact of cumulative exposure and cumulative dose

A cursory glance at Tables 1 and 2 suggests that 20 included studies present very heterogeneous cumulative exposure outcomes and cumulative dose-related effects. We estimated, for each individual study, fixed-effects non-linear regression models for dose–response association. These models can be estimated only if (1) the median or mean of each exposure level is shown in the article (or can be estimated); (2) the number of deaths and person-time is reported for each level; and (3) hazard ratio estimates and their standard errors are reported (or can be estimated) for at least three exposure categories. This information was available for four studies [24,25,30,32] (Fig. 4).

We observed that only one study [24] showed a relatively increasing trend between a cumulative dose of 0–10 000 mg, whereas from the other graphs a clear relationship did not emerge for dose–response association. The pooled (random-effects) model has not been calculated given the high heterogeneity and different sources of bias among studies.

Assessment of bias and confounding

We assessed the presence of bias and adequate control for putative confounders in all papers included in the review. Most studies are characterized by relatively poor adjustment for covariates in the statistical models, which could leave open the possibility of residual confounding. This is partly attributable to the use of in-hospital or administrative databases with lack of information or scarce possibility of merging data with those of other databases, as we will discuss in the next section. A minimum set of confounders to consider for full adjustment should comprise: age; gender; year of cohort entry; follow-up duration; concurrent use of other diabetes medications; diabetes duration; presence of bladder conditions, such as bladder polyp or schistosomiasis (but not a history isolated bacterial infection episodes); current or past occupations that may be associated with raised risk of bladder cancer, such as employment in the rubber industry before 1990 or the painting industry. Only few studies did include these controlling variables (Table 3).

With regard to selection bias, most of the studies were not planned as new-user designs, but may have inserted prevalent users of pioglitazone. Prevalent user designs can lead to biased and inconsistent estimation of risk effect.

Immortal time bias [11] consists of a misclassification of exposure time, when, at least for some subjects in the 'exposed' group, cohort entry does not coincide with first prescription of pioglitazone, and the exposure is modelled by a time-invariant variable assigned prior to the first prescription. The bias is induced when the person-time between cohort entry and the actual start of the exposure is classified as 'exposed' instead of 'unexposed', but certainly the outcome could not have occurred within this interval, as the patient had to survive until the first prescription. Hence, such time is misclassified and is referred to as 'immortal time'. This misclassification typically leads to a systematic bias toward a 'protective effect' of exposure. We found potential for this bias in seven of the 20 studies (35%) we reviewed [22,23,26,27,31,33,35]. In the case of comparison of two drugs (e.g. pioglitazone vs a sulfonylurea [27,34]) there may again be an immortal time bias, but its strength and direction might be unpredictable based only on information provided in the article. In fact, in this case, the bias depends on the mean time to first prescription because the group starting the first prescription later will obtain some spurious 'protective' effect as the denominator (person-years) for this group will be more greatly over-estimated than the denominator for the other group (Table 4).

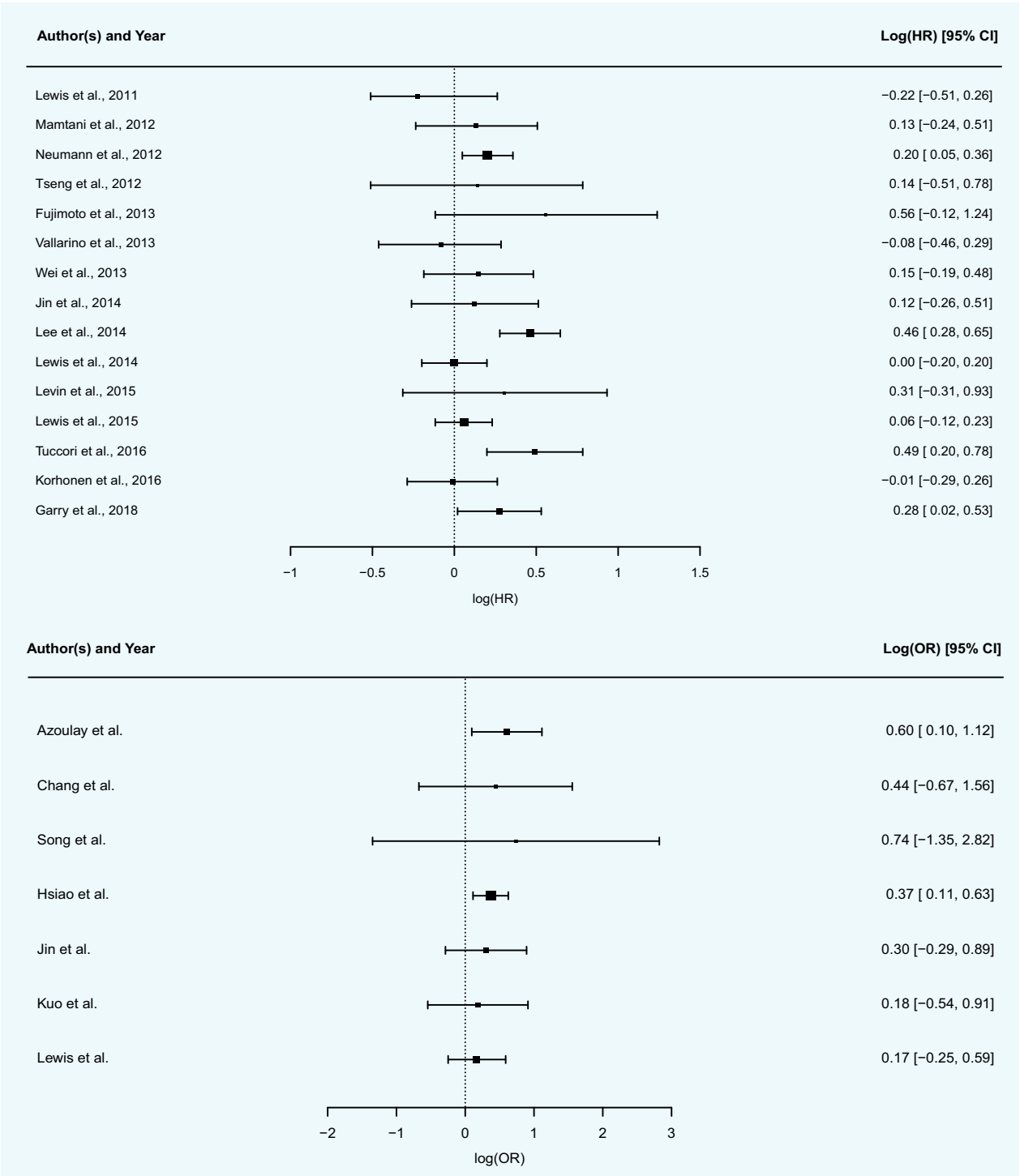

FIGURE 2 Forest plots illustrating the main results obtained in cohort studies (top) and case–control studies (bottom). HR, hazard ratio; OR, odds ratio.

An example of a study affected by immortal time bias is that of Wei *et al.* [22]. This was a cohort study using the General Practice Research Database in which 23 548 participants exposed to pioglitazone and 184 166 participants exposed to 'other medication but not pioglitazone' were followed between 2001 and 2010. Cohort entry does not coincide with filling the first prescription (but rather was set considering the presence of 'at least one prescription' during the study period, which might include drugs other than pioglitazone that are not specified), but the authors

Table 2 Nested case–control studies included in the review

Study	Database	Total sample size	Cases	Controls	Follow-up, months	Cumulative exposure, months	Cumulative dose, mg	OR (95% CI)
Azoulay *et al.* [38]	GPRD	115 727	376	6699	55			1.83 (1.10–3.05)
						≤12		0.56 (0.07–4.42)
						13-14		3.03 (0.63–14.52)
						>24		1.99 (1.14–3.45)
							1–10 500	1.58 (0.69–3.62)
							10 501–28 000	1.66 (0.70–3.94)
							>28 000	2.54 (1.05–6.24)
Chang *et al.* [39]	Taiwan	606 683			84	≥36		1.56 (0.51–4.74)
Song *et al.* [41]	In-hospital		329	658	72			2.09 (0.26–16.81)
Hsiao *et al.* [40]	Taiwan		3412	17 060		<12		1.45(1.12–1.87)
						12–24		1.74 (1.05–2.90)
						>24		2.93 (1.59–5.38)
Jin *et al.* [33]	Korea	113 193	208	620				1.35 (0.75–2.44)
Lewis *et al.* [29]	KPNC	193 099	464	464	180			1.18 (0.78–1.80)

GPRD, General Practice Research Database; KPNC, Kaiser Permanente Northern California; OR, odds ratio.

performed a Cox proportional hazard model without adjusting for time-dependent exposure; this could have led to immortal time bias. The paper by Neumann *et al.* [26], a large-scale study which led the French Regulatory Agency to withdraw pioglitazone from the market because of safety concerns, may also have been subject to immortal time bias. The study was conducted based on the French national health insurance information system, linked with the French hospital discharge database. The cohort, followed starting from 2006, included 1 491 060 participants with diabetes, 155 535 of whom were exposed to pioglitazone. In principle, the authors referred to a dynamic definition of exposure as 'having filled at least two prescriptions over a 6-month period'. By contrast, they did not adjust drug exposure dynamically in the Cox model, which causes the bias. Since there is no 'unexposed' group, but the control group includes participants taking other glucose-lowering medications, it is not possible to further specify which would be the strength and the direction of the bias. In two out of seven studies with potential for immortal time bias we reported significant risk of bladder cancer [26,31]. While the direction of the bias is unpredictable in the study by Neumann *et al.* [26], it should be protective in the study by Lee *et al.* [31] because the authors used fixed cohort entry with unexposed participants as the reference group. Nevertheless in the latter study there is also potential for time-lag bias, inclusion of prevalent users and no consideration of latency. Taken together, the direction of these effects might be also unpredictable.

Time-lag bias [13] is a source of error that might emerge when a second-line therapy, such as pioglitazone in Type 2 diabetes, is compared with a first-line therapy, such as metformin, or with a third-line agent. The problem is that people prescribed with different lines of treatment are unlikely to be at the same stage of disease and this, by itself, may confound the relationship with the outcome. Furthermore, there is a general trend for an increasing number of treatment switches to be a marker for a more 'difficult' (less responsive to treatment) person with diabetes [44]. We found potential for this bias in six studies [23–26,31,33]. One of these studies is that of Vallarino *et al.* [23]. In their retrospective cohort study, follow-up took place from 2000 to 2010 and 56 536 participants with diabetes were extracted from the InVision database; participants were users of pioglitazone and 17 948 were users of insulin, which was chosen as the active comparator. Insulin is not commonly considered as a second-line therapy [45], hence the reference group might have included people with more advanced disease, causing time-lag bias and confounding by disease severity.

Time-window bias [12] is an error peculiar to case–control studies, which consists in failure to match cases and controls on duration of time-exposure opportunity during the follow-up. We detected the presence of time-window bias in the study by Hsiao *et al.* [40], a case–control study conducted using the Taiwan's National Health Insurance Research Database, in which the authors identified 3412 cases of newly diagnosed bladder cancer and 17 060 controls. To define exposure the authors adopted ordinal labels according to which use of pioglitazone was defined as 'current' if the prescription duration covered the index date or ended during the previous 90 days, as 'recent' if it ended 91–180 days before the index date, or as 'past' if the last prescription ended >180 days before. In addition, controls were not matched to cases with regard to possibly being exposed to pioglitazone for a comparable amount of time. In other words, because cases and controls were not matched for follow-up duration, it is possible that in one or the other group there was a higher probability of receiving additional prescriptions.

Most papers reviewed excluded participants with an incident diagnosis of bladder cancer at cohort entry, but did not mention any consideration of latency period, which

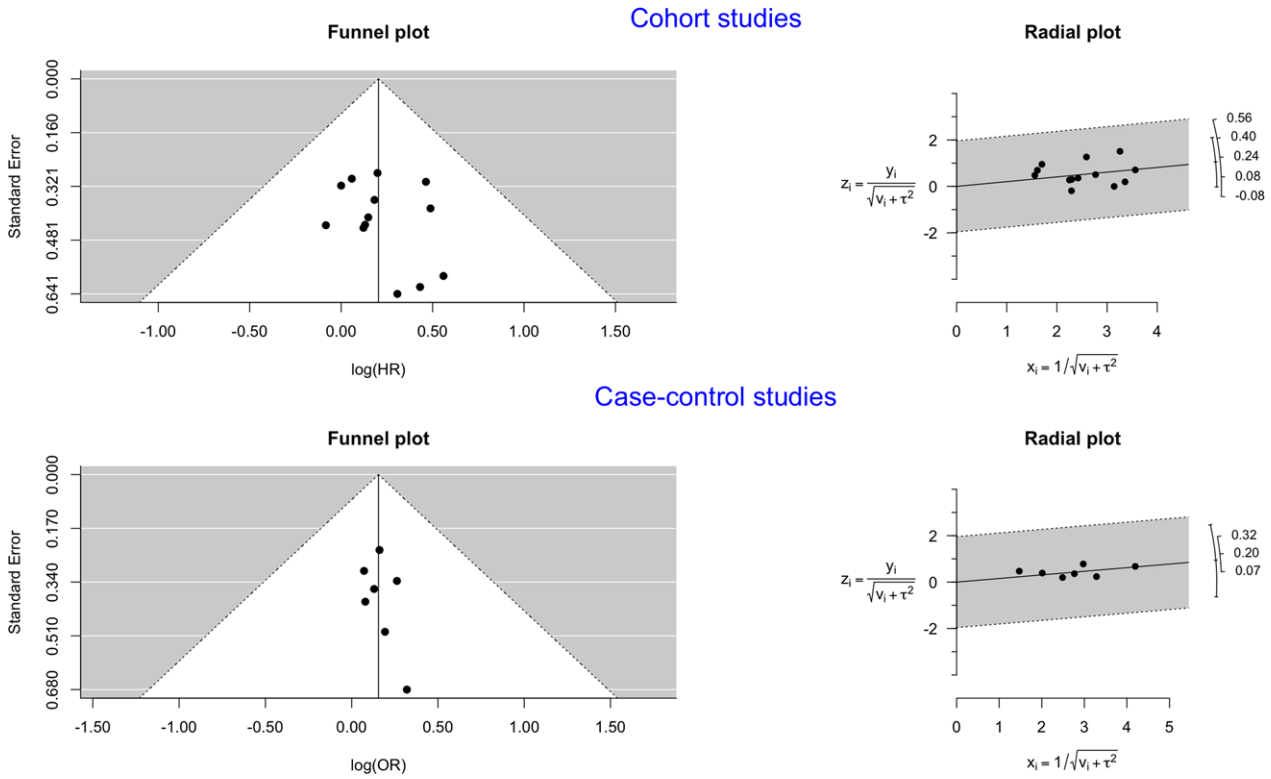

FIGURE 3 Funnel and radial plots for the observational studies considered in this review. HR, hazard ratio; OR, odds ratio.

would imply excluding people with less than a certain amount of follow-up time (e.g. 1 year) from cohort entry, as short-term therapy with pioglitazone may not increase risk of bladder cancer.

Finally, most papers did not present any clear statement about the availability of outcome/covariate data for all participants, the exclusion of subjects because of missing data and the techniques used to deal with missing data.

Discussion

Heterogeneous studies that are affected by several sources of bias, in particular time-related bias, lead to uncertain and inconclusive results. First, it is worth mentioning that previous meta-analyses have included biased studies or studies with poor adjustment. In particular, Ferwana *et al.* [4] pooled the results of six observational studies and one randomized controlled trial (RCT) [46], thus including a total of 215 142 participants; a random-effects model was run, leading to the detection of a significantly increased risk of bladder cancer in people using pioglitazone (hazard ratio 1.23, 95% CI 1.09–1.39). Even though the authors used a sensitivity analysis to study the robustness of the choice of stirring together different study designs, it seems imprudent to include RCTs and observational studies in the same pooled model; on the one hand, unlike observational studies, properly conducted RCTs allow both measured and unmeasured confounders to be fully accounted for, and on the other hand, these studies, in most cases, have a reduced degree of generalizability to particular segments of the population (such as elderly, pregnant women, children). Hence, as they present with different strengths and limitations which cannot be accounted for by statistical adjustment, RCTs and observational studies should be kept separate in the primary analysis. In addition, Ferwana *et al.* pooled the results of three studies [22,25,26] including prevalent users (i.e. participants taking the therapy for some time before the follow-up was started), which could have led to substantial bias. We found a similar problem in the review by Bosetti *et al.* [5], who pooled the estimates of several studies with bias, such as immortal time bias. Turner *et al.* [1] separately pooled the estimates from five RCTs (7878 participants) and 13 observational studies (>2.6 million people). Both meta-analyses showed increased risk for pioglitazone users (RCTs: odds ratio 2.51, 95% CI 1.09–6.80; observational studies: odds ratio 1.21, 95% CI 1.09–1.35). Even though studies with very different characteristics and limitations were included in the analysis, the authors adopted a fixed-effects approach, which cannot utterly capture the heterogeneity underpinning different designs. They also considered in the meta-analyses studies with prevalent users of pioglitazone, and studies affected by rather inadequate methodological choices, which could threaten the identification of a genuine effect. It might also be problematic to pool the results of studies with different lengths of follow-

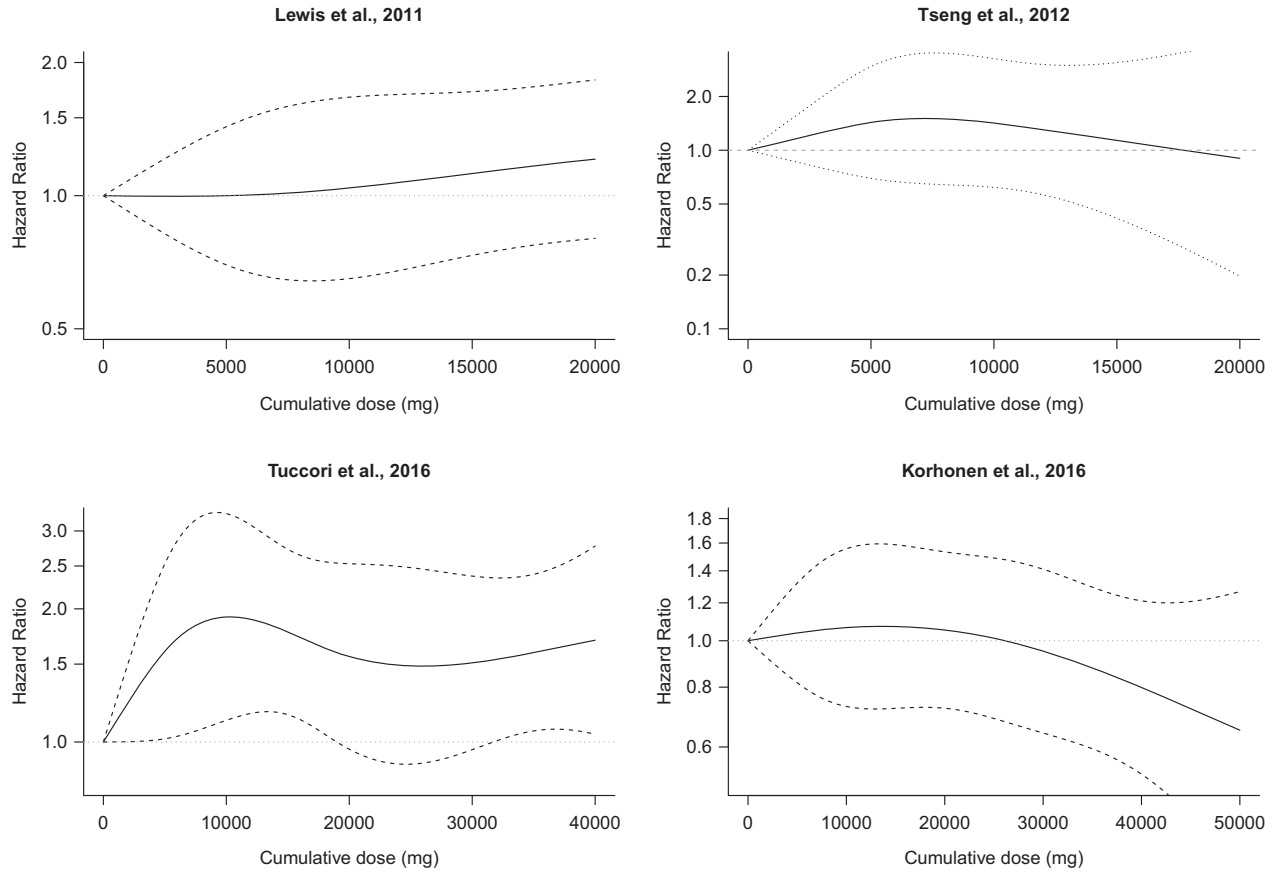

FIGURE 4 Estimation, for four individual cohort studies [24,25,30,32], of fixed-effects non-linear regression models for dose–response association.

up. In fact, effect size (and even its direction) could potentially vary at different time points of follow-up and, for each study, it should be checked whether the assumption of hazards proportionality is verified, and, if not, *ad hoc* sensitivity analyses should be run, e.g. stratifying by duration of follow-up.

In summary, previously published systematic reviews put forward meta-analyses that support the finding of increased bladder cancer risk for pioglitazone users; however, important concerns stand out, primarily because of the inclusion of biased studies; in fact such sources of bias (either time-related or not) could not be accounted for by standard meta-analytical techniques. In the present review, we found that seven studies were potentially affected by immortal time bias, one study included time-window bias and five studies did not consider time-lag bias. All of these errors can be fundamentally attributed to misclassification bias, i.e. failure to properly consider and label exposure/non-exposure time in the analysis. In general, the introduction of these biases leads to a reduction in the risk effect of a drug, exaggerating downward the effect estimate.

It is not unusual to come across immortal time bias in the context of cohort studies using administrative/electronic databases. In general, the bias can be avoided using an intention-to-treat approach, which, by definition, would have considered exposure according to the first pioglitazone prescription, without taking into account switching or gaps. An alternative is to model hazard ratios using a statistical approach, taking into account time-varying exposures, such as Cox proportional hazards models with time-dependent exposure; however, this approach requires a careful choice of the aetiologically relevant time-varying exposure metric, as different metrics, for example, current use vs cumulative duration of past use, may lead to very different relative risk estimates and qualitatively different conclusions [47]. A second cautionary note is that, in pharmaco-epidemiological studies conducted using large databases, a time-varying approach for exposure to avoid immortal time bias might require long computation times because accounting for within-subject measurement over time requires some sort of clustering. In addition, marginal structural models [48] are a valid option to account for time-varying exposures. A less frequent bias is time-lag bias, which could also be prevented, however, by matching on disease duration.

A common limitation of case–control studies is time-window bias, which is another source of imbalance, as cases and controls are not guaranteed to have the same

Table 3 Assessment of bias using the ROBINS-I tool.

		Pre-intervention										Intervention	Post-intervention			
		Bias attributable to confounding						Bias attributable to selection/time				Bias attributable to classification				
Study	Demographics	SES	Lifestyle	Comorbidities	Environment*	Diabetes†	TRV	ITB	TWB	TLB	Prevalent users		Deviations	MD	MO	RR
Lewis *et al.* [30]			+	+	+		+			+						
Mamtani *et al.* [27]		+	+	+	+		+	+								
Neumann *et al.* [26]		+	+	+		+	+	+		+	+	+				+
Tseng [25]	+	+	+	+	+	+	+	+		+	+	+	+	+	+	+
Fujimoto *et al.* [35]	+	+	+	+	+	+	+	+			+	+				
Vallarino *et al.* [23]		+		+	+	+	+	+		+		+	+			
Wei *et al.* [22]				+	+		+	+		+	+	+	+		+	+
Jin *et al.* [33]	+	+	+	+	+	+	+	+		+	+	+	+		+	
Lee *et al.* [31]		+	+	+	+		+	+		+	+					+
Lewis *et al.* [28]		+	+	+	+	+										
Levin *et al.* [36]		+		+	+	+		+			+	+				+
Lewis *et al.* [29]				+	+											
Tuccori *et al.* [24]		+		+	+					+						
Korhonen *et al.* [32]			+	+	+											
Garry *et al.* [34]				+	+											
Azoulay *et al.* [38]		+		+	+											
Chang *et al.* [39]	+		+	+	+	+	+					+				
Song *et al.* [41]		+		+	+	+	+				+	+				
Hsiao *et al.* [40]		+	+	+	+		+		+		+	+				
Jin *et al.* [33]		+	+	+	+		+				+	+				
Lewis *et al.* [29]				+	+											

ITB, immortal time bias; MD, bias attributable to missing data; MO, bias in measurement of outcomes; RR, bias in selection of reported results; SES, socio-economic status; TLB, time-lag bias; TRV, time-related variables; TWB, time-window bias.
*Enviromental factors. †Diabetes-related covariates.
Pre-intervention, bias attributable to confounding: we report the domains not adequately controlled for in the analyses, which could have led to bias. We listed six domains in the protocol: demographics (e.g. age, gender); SES (e.g. household income); lifestyle (e.g. smoking, alcohol consumption) and anthropometric variables (e.g. BMI); comorbidities, i.e. presence of bladder conditions such as bladder polyp or schistosomiasis (but not a history isolated bacterial infection episodes); environmental factors linked to current or past occupations that may be associated with raised risk of bladder cancer; diabetes-related covariates (e.g. diabetes duration, use of other diabetes medications); TRV (e.g. follow-up duration, year of cohort entry).
Bias in selection of participants into the study: this is an umbrella term in ROBINS-I, including different sources of bias described with different labels in the literature. We report here ITB, TWB, TLB, and inclusion of prevalent users. At intervention. bias due classification of interventions.
Post-intervention, deviations: bias attributable to deviations from the intended interventions,

Table 4 Illustration of immortal time bias in the cohort studies included in the review

Study	Comparator	Cohort entry	What might originate the bias?	Direction of the bias
Neumann *et al.* [26]	Other than pioglitazone (metamorfin, sulfonylurea, rosiglitazone, other oral hypoglycaemic agent and/or insulin)	Date of the first prescription of a glucose-lowering drug	Use of time invariant exposure in Cox proportional hazard model	Unpredictable
Wei *et al.* [22]	Other anti-diabetic drug	Date of first prescription for pioglitazone or other oral hypoglycaemic drug	Use of time invariant exposure in Cox proportional hazard model	Unpredictable
Mamtani *et al.* [27]	Sulfonylurea	Not clear	Use of time invariant exposure in Cox proportional hazard model	Unpredictable
Fujimoto *et al.* [35]	Possibly unexposed, but not specified	Not clear	Use of time invariant exposure in Cox proportional hazard model and comparison with unexposed group	Protective
Vallarino *et al.* [23]	Insulin	Not clear	Use of time invariant exposure in Cox proportional hazard model	Unpredictable
Jin *et al.* [33]	Unexposed	Not clear	Use of time invariant exposure in Cox proportional hazard model and comparison with unexposed group	Protective
Lee *et al.* [31]	Unexposed	Fixed	Use of time invariant exposure in Cox proportional hazard model and comparison with unexposed group	Protective

opportunity for exposure. It is also worth remarking that, in case–control studies, effects are generally reported in terms of odds ratio; however, this metric should be interpreted with caution because of the non-collapsibility issue. An effect measure is said to be collapsible when an unconditional effect (i.e. a marginal population effect) can be expressed in terms of a weighted average of conditional effects (i.e. strata-specific effects). Thus, collapsibility is a natural and desirable property, which might however be violated by odds ratios [49].

Another relevant problem underpinning several of the reviewed studies [22,25,26,31,40,41] is the inclusion of prevalent users in the analyses. Prevalent users are people (in principle, in either treated and untreated groups) initiating pioglitazone use before cohort entry; such a period of therapy might not be registered in the database or, more frequently, is attributable to the fact that the drug entered the marked before of start of follow-up. Inclusion of prevalent users may create bias, as 'prevalent users' are 'survivors' of the early period of pharmacotherapy. This can introduce substantial bias if risk varies over time, just as in studies of operating procedures that enrol patients after they have survived surgery [50]. In addition, long-term users are generally more adherent to therapy than short-term users, and this might reflect another type of bias, adherence bias. New-user designs avoid the introduction of these chronology biases because, in such designs, the study is restricted to participants with a minimum period of washout or never-use prior to cohort entry. Most of the studies included in the present review are affected by potential inclusion of prevalent users (Table 3), for instance, the study by Wei *et al.* [22] in which participants entered the cohort having filled 'at least one prescription between January 2001 and December 2010', but without a proper look backward to exclude past users of pioglitazone. It should also be observed that, in principle, adjusting for a 'prevalent' condition does not only refer to drug exposure, but also to the presence of other previous events (such as past admissions to hospital) at cohort entry, which are virtually present in all epidemiological analyses. A limitation of most pharmaco-epidemiological studies is that of neither considering nor modelling the time window between previous events related to exposure/outcome (e.g. a previous acute condition) and start of follow-up. Recent events might have a more relevant impact than relatively old events, for instance, in terms of confounding by indication.

We also remark that it is fundamental to take into account a suitable period of latency when evaluating the effect of a drug on cancer occurrence, as very short-term use is unlikely to be associated with incident cancer. Latency can be defined as that time period between irreversible disease occurrence and the detection of the disease [51]. It is particularly important to consider latency in cancer research, as different types of tumours might be characterized by different durations of the latency period. Thus, it is reasonable to set an interval of time after cohort entry during which incident cases of cancer are excluded from the count, as these are probably not related to drug exposure. Non-consideration of proper latency periods is an important issue in this area of research, and we detected this shortcoming in 11 out of 15 cohort studies [22,23,25–31,33,35] and in five out of six case–control studies [29,33,39–41].

A major problem in contemporary pharmaco-epidemiology also lies in the quality of data provided by automated databases. Observational studies reviewed in the present paper include investigations conducted using in-hospital databases [35,41], with relatively poor ability to adjust for confounders. Other studies used national administrative data [25,31,40,42], with limited ability to link information with other data sources. Few studies used high-quality registries [28–30] or electronic records [24,32,34,38], which, are currently characterized, however, by a relatively high rate of missing values, and ideally should be handled using proper statistical techniques for non-randomly occurring missing data to avoid miss-classification.

Finally, as shown in Tables 2 and 3, we found six cohort studies that were relatively unaffected by major sources of bias or confounding [24,29,30,32,34,36], modelling exposure as a time-variant variable; however, within this subsample, results are discordant, as four studies reported no association [29,30,32,36] and two found increased risk [24,34]. The same was also true for case–control studies, where only two studies were free of bias [29,38], but only one reported association [38]. As to cohort studies, Garry *et al.* [34] found increased risk for exposure to pioglitazone within 2 years; Tuccori *et al.* [24] reported a similar result in cases exposed to the drug for >2 years, but time of exposure of 1 to 2 years led to a near-significant result (hazard ratio 1.66 95% CI 0.97–2.84). These studies are free of immortal time bias, which, for cumulative exposures of this magnitude, is a relevant issue, as proven by empirical calculation targeting the association between exposure to metformin and different types of cancer [13]. It is also worth remarking that the study by Garry *et al.* [34] used an as-treated exposure definition for the main analysis, which may be problematic as it assumes that the drug has a reversible effect on the outcome after discontinuation (this is theoretically possible, but not observable in a short follow-up period). In addition the authors compared pioglitazone users with dipeptidyl peptidase-4 inhibitor or sulfonylurea users, but the latter assessment may be problematic as in the design people who shifted from sulfonylureas to pioglitazone are excluded (possible selection bias).

In conclusion, it remains a challenge to ascertain in future research whether pioglitazone plays a drug-specific role in the occurrence of bladder cancer. Results from previous observational studies are controversial and very heterogeneous, but, more notably, we showed how most of these studies were affected by bias and poor controlling for confounding. This should be taken into account when considering the possibility of pooling these data in meta-analyses. As the dose–response pattern seems to be a critical issue and previous studies are very heterogeneous in this regard, we suggest that a possible direction for future research would be to investigate this relationship using weighted cumulative effects models. We recommend that forthcoming studies be based only on administrative registries/databases with electronic records that have quality controls and internal quality audits because in-hospital databases have repeatedly been found to be unsuitable to adjust for confounding. We would also recommend adopting a new-user design approach and taking into account a suitable time interval for latency. Scrutinizing all possible sources of confounding, as well as adjusting multivariable models for these covariates, is also important. We found that adjustment for occupational exposures and the identification of specific bladder conditions associated with bladder cancer, such as history of bladder polyp and schistosomiasis infection, has not been carried out so far in the literature. Taking into consideration the temporal variation in exposure initiation is also crucial and can be accomplished by modelling exposure with a time-varying indicator. Prospective meta-analyses in this field should use analytical procedures to detect heterogeneity, but should also put more emphasis on the issue of time-related biases and avoid pooling the results of different studies when these contain different sources of bias and a high degree of heterogeneity. Our general conclusion for diabetologists is that the question of the association between pioglitazone use and bladder cancer has not yet been answered and prescription of this drug should be considered with care. International and regulatory agencies should promote new collaborative observational research to clarify this issue definitively.

Funding sources

None.

Competing interests

None declared.

References

1 Turner RM, Kwok CS, Chen-Turner C, Maduakor CA, Singh S, Loke YK. Thiazolidinediones and associated risk of bladder cancer: a systematic review and meta-analysis. *Br J Clin Pharmacol* 2014; **78**: 258–273.
2 Tseng C-H. A review on thiazolidinediones and bladder cancer in human studies. *J Environ Sci Heal Part C* 2014; **32**: 1–45.
3 He S, Tang Y, Zhao G, Yang X, Wang D, Zhang Y. Pioglitazone prescription increases risk of bladder cancer in patients with type 2 diabetes: an updated meta-analysis. *Tumor Biol* 2014; **35**: 2095–2102.
4 Ferwana M, Firwana B, Hasan R, Al-Mallah MH, Kim S, Montori VM *et al.* Pioglitazone and risk of bladder cancer: a meta-analysis of controlled studies. *Diabet Med* 2013; **30**: 1026–1032.
5 Bosetti C, Rosato V, Buniato D, Zambon A, La Vecchia C, Corrao G. Cancer risk for patients using thiazolidinediones for type 2 diabetes: a meta-analysis. *Oncologist* 2013; **18**: 148–156.
6 Zhu Z, Shen Z, Lu Y, Zhong S, Xu C. Increased risk of bladder cancer with pioglitazone therapy in patients with diabetes: a meta-analysis. *Diabetes Res Clin Pract* 2012; **98**: 159–163.
7 Colmers IN, Bowker SL, Majumdar SR, Johnson JA. Use of thiazolidinediones and the risk of bladder cancer among people

with type 2 diabetes: a meta-analysis. *Can Med Assoc J* 2012; **184**: E675–E683.
8 Filipova E, Uzunova K, Kalinov K, Vekov T. Pioglitazone and the risk of bladder cancer: A meta-analysis. *Diabetes Ther* 2017; **8**: 705–26.
9 Adil M, Khan RA, Ghosh P, Venkata SK, Kandhare AD, Sharma M. Pioglitazone and risk of bladder cancer in type 2 diabetes mellitus patients: A systematic literature review and meta-analysis of observational studies using real-world data. *Clin Epidemiol Glob Heal* 2018; **6**: 61–68.
10 Li Z, Sun M, Wang F, Shi J, Wang K. Association between pioglitazone use and the risk of bladder cancer among subjects with diabetes mellitus: A dose-response meta-analysis. *Int J Clin Pharmacol Ther* 2017; **55**: 210–219.
11 Suissa S. Immortal time bias in pharmacoepidemiology. *Am J Epidemiol* 2007; **167**: 492–499.
12 Suissa S, Dell'Aniello S, Vahey S, Renoux C. Time-window bias in case-control studies: statins and lung cancer. *Epidemiology* 2011; **22**: 228–231.
13 Suissa S, Azoulay L. Metformin and the risk of cancer. *Diabetes Care* 2012; **35**: 2665–2673.
14 Moher D, Liberati A, Tetzlaff J, Altman DG; PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Ann Intern Med* 2009;**151**:264–269.
15 Sterne J, Hernan M, Reeves B, Savovic J, Berkman N, … ROBINS-I: a tool for assessing risk of bias in non-randomized studies of intervention. *BMJ* 2016;**355**:i4919.
16 Higgins JPT, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Stat Med* 2002; **21**: 1539–1558.
17 Viechtbauer W. Conducting meta-analyses in R with the metafor package. *J Stat Softw* 2010; **36**: 1–48.
18 Orsini N, Li R, Wolk A, Khudyakov P, Spiegelman D. Meta-analysis for linear and nonlinear dose-response relations: examples, an evaluation of approximations, and software. *Am J Epidemiol* 2012; **175**: 66–73.
19 Harrell F. R package "rms." Nashville: 2018.
20 Greenland S, Longnecker MP. Methods for trend estimation from summarized dose-response data, with applications to meta-analysis. *Am J Epidemiol* 1992; **135**: 1301–1309.
21 Crippa A, Orsini N. Multivariate Dose-Response Meta-Analysis: The dosresmeta R Package. *J Stat Softw* 2016; **72**: 1–15.
22 Wei L, MacDonald TM, Mackenzie IS. Pioglitazone and bladder cancer: a propensity score matched cohort study. *Br J Clin Pharmacol* 2013; **75**: 254–259.
23 Vallarino C, Perez A, Fusco G, Liang H, Bron M, Manne S *et al*. Comparing pioglitazone to insulin with respect to cancer, cardiovascular and bone fracture endpoints, using propensity score weights. *Clin Drug Investig* 2013; **33**: 621–631.
24 Tuccori M, Filion KB, Yin H, Oriana HY, Platt RW, Azoulay L. Pioglitazone use and risk of bladder cancer: population based cohort study. *Br Med J* 2016; **352**: i1541.
25 Tseng C-H. Pioglitazone and bladder cancer. *Diabetes Care* 2012; **35**: 278–280.
26 Neumann A, Weill A, Ricordeau P, Fagot JP, Alla F, Allemand H. Pioglitazone and risk of bladder cancer among diabetic patients in France: a population-based cohort study. *Diabetologia* 2012; **55**: 1953–1962.
27 Mamtani R, Haynes K, Bilker WB, Vaughn DJ, Strom BL, Glanz K *et al*. Association between longer therapy with thiazolidinediones and risk of bladder cancer: a cohort study. *J Natl Cancer Inst* 2012; **104**: 1411–1421.
28 Lewis JD, Habel L, Quesenberry C, Mamtani R, Peng T, Bilker WB *et al*. Proteinuria testing among patients with diabetes mellitus is associated with bladder cancer diagnosis: potential for unmeasured confounding in studies of pioglitazone and bladder cancer. *Pharmacoepidemiol Drug Saf* 2014; **23**: 636–645.
29 Lewis JD, Habel LA, Quesenberry CP, Strom BL, Peng T, Hedderson MM *et al*. Pioglitazone use and risk of bladder cancer and other common cancers in persons with diabetes. *J Am Med Assoc* 2015; **314**: 265–277.
30 Lewis JD, Ferrara A, Peng T, Hedderson M, Bilker WB, Quesenberry CP *et al*. Risk of bladder cancer among diabetic patients treated with pioglitazone. *Diabetes Care* 2011; **34**: 916–922.
31 Lee M-Y, Hsiao P-J, Yang Y-H, Lin K-D, Shin S-J. The association of pioglitazone and urinary tract disease in type 2 diabetic Taiwanese: bladder cancer and chronic kidney disease. *PLoS One* 2014; **9**: e85479.
32 Korhonen P, Heintjes EM, Williams R, Hoti F, Christopher S, Majak M *et al*. Pioglitazone use and risk of bladder cancer in patients with type 2 diabetes: retrospective cohort study using datasets from four European countries. *BMJ* 2016; **354**: i3903.
33 Jin S-M, Song SO, Jung CH, Chang J-S, Suh S, Kang SM *et al*. Risk of bladder cancer among patients with diabetes treated with a 15 mg pioglitazone dose in Korea: a multi-center retrospective cohort study. *J Korean Med Sci* 2014; **29**: 238–242.
34 Garry E, Buse B, Lund J, Pate V, Stürmer T. Comparative safety of pioglitazone verus clinically meaningful treatment alternatives concerning the risk of bladder cancer in older US adults with type 2 diabetes. *Diabetes Obes Metab* 2018; **20**: 129–140.
35 Fujimoto K, Hamamoto Y, Honjo S, Kawasaki Y, Mori K, Tatsuoka H *et al*. Possible link of pioglitazone with bladder cancer in Japanese patients with type 2 diabetes. *Diabetes Res Clin Pract* 2013; **99**: e21–e23.
36 Levin D, Bell S, Sund R, Hartikainen SA, Tuomilehto J, Pukkala E *et al*. Pioglitazone and bladder cancer risk: a multipopulation pooled, cumulative exposure analysis. *Diabetologia* 2015; **58**: 493–504.
37 Levin D, Bell S, Sund R, Hartikainen SA, Tuomilehto J, Pukkala E *et al*. Comparative safety of pioglitazone verus clinically meaningful treatment alternatives concerning the risk of bladder cancer in older US adults with type 2 diabetes. *Diabetologia* 2012; **20**: 129–140.
38 Azoulay L, Yin H, Filion KB, Assayag J, Majdan A, Pollak M *et al*. The use of pioglitazone and the risk of bladder cancer in people with type 2 diabetes: nested case-control study. *Br Med J* 2012; **344**: e3645.
39 Chang C-H, Lin J-W, Wu L-C, Lai M-S, Chuang L-M, Arnold Chan K. Association of thiazolidinediones with liver cancer and colorectal cancer in type 2 diabetes mellitus. *Hepatology* 2012; **55**: 1462–1472.
40 Hsiao F-Y, Hsieh P-H, Huang W-F, Tsai Y-W, Gau C-S. Risk of bladder cancer in diabetic patients treated with rosiglitazone or pioglitazone: a nested case–control study. *Drug Saf* 2013; **36**: 643–649.
41 Song SO, Kim KJ, Lee B-W, Kang ES, Cha BS, Lee HC. The risk of bladder cancer in Korean diabetic subjects treated with pioglitazone. *Diabetes Metab J* 2012; **36**: 371–378.
42 Kuo H-W, Tiao M-M, Ho S-C, Yang C-Y. Pioglitazone use and the risk of bladder cancer. *Kaohsiung J Med Sci* 2014; **30**: 94–97.
43 Higgins JPT, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ Br Med J* 2003; **327**: 557.
44 Blais L, Ernst P, Suissa S. Confounding by indication and channeling over time: the risks of β2-agonists. *Am J Epidemiol* 1996; **144**: 1161–1169.
45 McIntosh B, Cameron C, Singh SR, Yu C, Ahuja T, Welton NJ *et al*. Second-line therapy in patients with type 2 diabetes inadequately controlled with metformin monotherapy: a systematic review and mixed-treatment comparison meta-analysis. *Open Med* 2011; **5**: e35–e48.

46 Dormandy JA, Charbonnel B, Eckland DJA, Erdmann E, Massi-Benedetti M, Moules IK *et al*. Secondary prevention of macrovascular events in patients with type 2 diabetes in the PROactive Study (PROspective pioglitAzone Clinical Trial In macroVascular Events): a randomised controlled trial. *Lancet* 2005; **366**: 1279–1289.
47 Abrahamowicz M, Beauchamp M-E, Sylvestre M-P. Comparison of alternative models for linking drug exposure with adverse effects. *Stat Med* 2012; **31**: 1014–1030.
48 Robins JM, Hernan MA, Brumback B. Marginal structural models and causal inference in epidemiology. *Epidemiology* 2000; **11**: 550–560.
49 Greenland S, Robins JM, Pearl J. Confounding and collapsibility in causal inference. *Stat Sci* 1999; **14**: 29–46.
50 Ray WA. Evaluating medication effects outside of clinical trials: new-user designs. *Am J Epidemiol* 2003; **158**: 915–920.
51 Rothman KJ. Induction and latent periods. *Am J Epidemiol* 1981; **114**: 253–259.

