



Article
scientifique

Revue de la
littérature

2019

Published
version

Open
Access

This is the published version of the publication, made available in accordance with the publisher's policy.

Time-dependent biases in observational studies of comparative effectiveness research in rheumatology. A methodological review

Iudici, Michele; Porcher, Raphaël; Riveros, Carolina; Ravaud, Philippe

How to cite

IUDICI, Michele et al. Time-dependent biases in observational studies of comparative effectiveness research in rheumatology. A methodological review. In: Annals of the Rheumatic Diseases, 2019, vol. 78, n° 4, p. 562–569. doi: 10.1136/annrheumdis-2018-214544

This publication URL: <https://archive-ouverte.unige.ch/unige:146595>

Publication DOI: [10.1136/annrheumdis-2018-214544](https://doi.org/10.1136/annrheumdis-2018-214544)

Time-dependent biases in observational studies of comparative effectiveness research in rheumatology. A methodological review

Michele Iudici,¹ Raphaël Porcher,¹ Carolina Riveros,¹ Philippe Ravaud^{1,2}

Handling editor Josef S Smolen

► Additional material is published online only. To view please visit the journal online (<http://dx.doi.org/10.1136/annrheumdis-2018-214544>).

¹Methods of Therapeutic Evaluation of Chronic Diseases (METHODS) Team, INSERM, UMR 1153, Epidemiology and Biostatistics Sorbonne Paris Cité Research Center (CRESS), Paris, France

²Cochrane France, Paris, France

Correspondence to

Dr Michele Iudici, Methods of Therapeutic Evaluation of Chronic Diseases (METHODS) team, INSERM, UMR 1153, Epidemiology and Biostatistics Sorbonne Paris Cité Research Center (CRESS), Paris 75001, France; michele_iudici@hotmail.com

Received 7 October 2018

Revised 2 January 2019

Accepted 23 January 2019

Published Online First

12 February 2019

ABSTRACT

Objective To assess to what extent time-dependent biases (ie, immortal time bias (ITB) and time-lag bias (TLB)) occur in the latest rheumatology observational studies, describe their main mechanisms and increase the awareness on this topic.

Methods We searched PubMed for observational studies on rheumatic diseases published in leading medical journals in the last 5 years. Only studies with a time-to-event analysis exploring the association of one or more interventional strategies with an outcome were included. Each study was labelled as free from bias, at risk of TLB, at risk of misclassified ITB if the period of immortal time was incorrectly attributed to an intervention group, or at risk of excluded ITB if the immortal time was discarded from the analysis.

Results We included 78 papers. Most studies were performed in Europe or North America (46% each), were not industry funded (62%) and had a safety primary outcome (59%). In total, 13 (17%) studies were considered at risk of time-dependent biases. Among the studies at risk of ITB ($n=8$; 10%), in 5 (6%), waiting time to receive treatment was wrongly attributed to the treatment exposure group, which indicated misclassified ITB. Five (6%) studies were at risk of TLB: patients on conventional synthetic disease-modifying antirheumatic drugs (DMARD; first-line drugs) were compared with patients on biologic DMARDs (second or third-line drugs) without accounting for disease duration or prior medication use.

Conclusions One in six comparative effectiveness observational studies published in leading rheumatology journals is potentially flawed by time-dependent biases.

INTRODUCTION

Time-dependent biases refer to a group of biases occurring in time-to-event analyses of observational studies. They include immortal time bias (ITB), referred to as ‘survivors treatment bias’^{1–4} or ‘guarantee-time bias’,⁵ and time-lag bias (TLB).^{6,7} Both biases tend to inflate the benefits observed with a drug in terms of higher efficacy or lower risks.^{6,7}

ITB occurs when the treatment exposure in a fixed-time model is wrongly assigned or excluded from the analysis.^{1,8} Here we describe an example using observational data to assess whether total knee arthroplasty (TKA) in patients with advanced knee osteoarthritis is associated with better survival as compared with medical therapy alone. We have data from the diagnosis of severe osteoarthritis and the concomitant prescription of chronic medical treatment. Patients

Key messages

What is already known about this subject?

► Time-dependent biases, that is, immortal time bias (ITB) and time-lag bias (TLB), are biases observed in time-to-event analyses. If present, they can distort study results by inflating the benefits of a drug in terms of higher efficacy or lower risk. These biases must be identified to avoid performing flawed analyses or wrongly interpreting the results of biased studies.

What does this study add?

► We found ITB or TLB in about one in six observational comparative effectiveness studies of rheumatology published in leading journals. A description of the main mechanisms leading to these biases in the field and a summary of the key points useful to identify and avoid them are provided.

How might this impact on clinical practice or future developments?

► A better recognition of such biases could help clinical researchers improve the quality of comparative effectiveness observational studies and clinicians critically appraise study results.

are eligible to undergo surgery. If patients are classified into ‘surgery’ or ‘medical therapy’ groups according to whether they had received surgery, the time between the diagnosis and surgery (figure 1, a) for the surgery group is immortal (‘free of events’) in the sense that they must have survived that time to be classified in this group. If this time is wrongly attributed to the surgery instead of medical therapy group (figure 1, a+b vs c), patients in the surgery group will benefit from this immortal time. This first case is reported as ‘misclassified ITB’.

A second case would be to follow-up surgery patients after they receive TKA (figure 1, b) and the medical therapy group after the diagnosis (figure 1, c). In this case, the time ‘free of events’ (immortal time) before surgery (figure 1, a) is no longer attributed to surgery patients (which is correct), but it is not even attributed to medical therapy patients (which is not correct) (figure 1, b vs c). The rate of events will be overestimated in the latter group. This is an example of ‘excluded ITB’.

Conversely, TLB is observed when patients at different disease stages are compared without taking



© Author(s) (or their employer(s)) 2019. No commercial re-use. See rights and permissions. Published by BMJ.

To cite: Iudici M, Porcher R, Riveros C, et al. *Ann Rheum Dis* 2019;**78**:562–569.

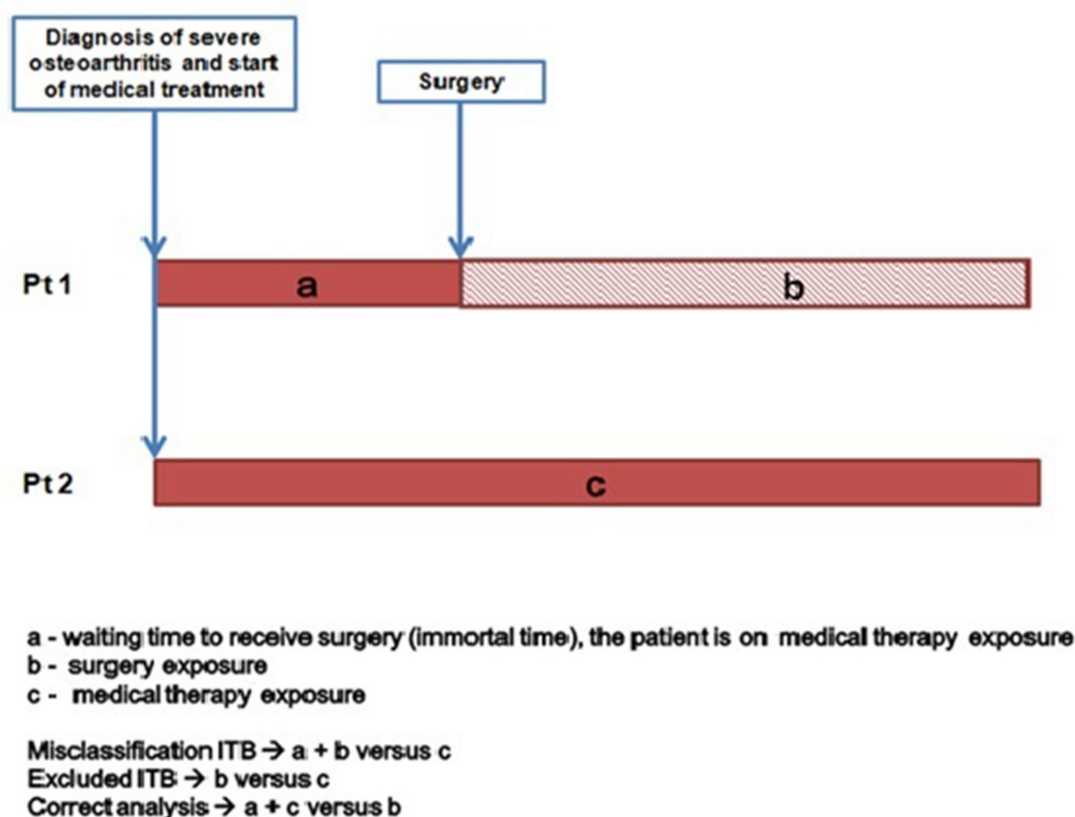


Figure 1 Illustration of mechanisms of immortal time bias (ITB). In misclassified ITB, the immortal time (box a) is wrongly assigned to the 'surgery' group (a+b vs c) and in excluded ITB, it is excluded from the analysis (b vs c).

into account disease duration, if the occurrence of the outcome of interest is potentially confounded by the disease duration itself. This bias can be observed, for example, in studies investigating patients taking first-line versus second-line or third-line drugs (figure 2).^{6,7}

In 2004, van Walraven *et al*⁹ found time-dependent biases in about 10% of cohort studies published in leading medical journals and showed that correcting the biases could have qualitatively changed study conclusions in more than half of the studies. Since then, despite an increasing number of reports highlighting

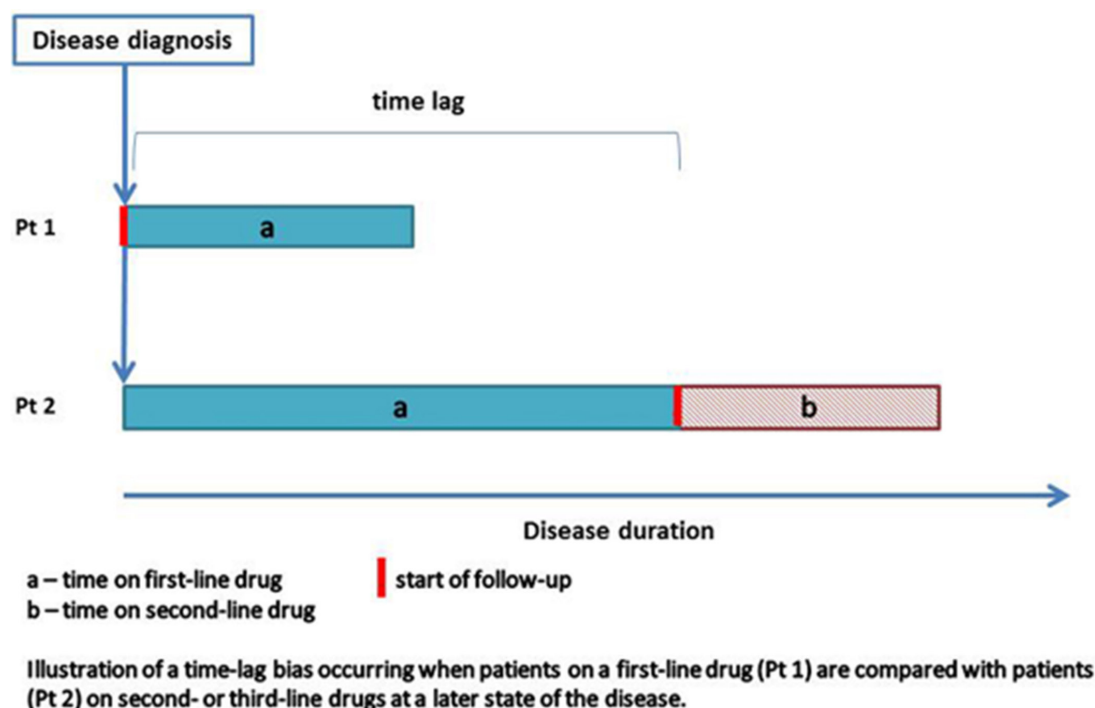


Figure 2 Illustration of the time-lag bias.

how these biases can undermine study results,^{1 6 8 10–17} no study has attempted to quantify and describe the main features of these biases in a specific medical domain.

Given the increasing number of registry and big database studies driving clinical decisions in rheumatic diseases,¹⁸ we aimed to assess to what extent time-dependent biases occur in the latest rheumatology literature and describe their main mechanisms to increase awareness on this topic.

MATERIALS AND METHODS

Search strategy

Because we considered that this methodological review lacked an outcome of direct clinical relevance, we did not record the protocol on PROSPERO. We followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines.¹⁹ We performed a search of MEDLINE via PubMed on 3 September 2017 to identify rheumatology papers published in the five journals with the highest impact factors (according to Journal Citation Reports, Clarivate Analytics) in rheumatology (*Annals of the Rheumatic Diseases*, *Arthritis & Rheumatology*, *Rheumatology* (Oxford), *Seminars in Arthritis and Rheumatism*, *Osteoarthritis and Cartilage*) and general and internal medicine (*NEJM*, *Lancet*, *JAMA*, *BMJ*, *Annals of Internal Medicine*). We used different terms referring to observational studies and survival analysis to obtain the highest sensitivity. The full search strategy is found in the online supplementary file.

Eligibility criteria

Types of studies. Observational studies exploring the effect of one or more interventional strategies on a time-to-event outcome. We excluded case-control studies, reviews, comments, editorials, letters, meta-analyses, network meta-analyses and studies with a survival analysis not comparing two exposure groups.

Types of participants. Participants of any age having a rheumatic disease.

Types of interventions. Pharmacological or non-pharmacological treatments.

Types of outcomes. We excluded studies with drug retention/discontinuation as an outcome.

Data collection

Two researchers (MI, CR) independently checked each title and abstract to exclude irrelevant articles and then independently examined the full-text articles to determine eligibility. Consensus was reached by discussion in case of disagreement. A third reviewer was available in case of unresolved disagreement. We documented the primary reason for exclusion of full-text articles.

Data extraction and management

One author (MI) extracted the data by using a standardised form, and a second author (CR) checked the extracted data. Disagreements were discussed to reach consensus. The complete list of study characteristics extracted can be found in the online supplementary file. We considered primary outcome(s) labelled 'primary' by authors or the first outcome presented in the Results section. A study was considered industry funded if the sponsor or one of the collaborators was industry.

Assessment of the risk of time-dependent biases

Each eligible full-text article was independently checked for the presence of time-dependent biases by two of the authors (MI, RP). The risk of ITB was evaluated according to the criteria proposed by Levesque *et al*⁸ consisting of the following six

questions: (1) Was treatment status determined after the start of follow-up or defined using follow-up time? (2) Was the start of follow-up different for the treated and comparator group relative to the date of diagnosis? (3) Were the treatment groups identified hierarchically? (4) Were subjects excluded on the basis of treatment identified during follow-up? (5) Was a time-fixed analysis used?

Each study was then classified as free from ITB or at risk of ITB. Studies were classified at risk of misclassified ITB if the period of immortal time was incorrectly attributed to the treated group by a time-fixed analysis, or at risk of excluded ITB if the immortal time was excluded from the analysis (figure 1).^{8 9}

A study was considered at risk of TLB if (1) two or more drugs prescribed at different disease stages were compared, (2) the study outcome was considered potentially confounded by disease duration, and (3) the analysis was not adjusted for disease duration or prior medication use. A drug was considered second line, third line, and so on according to statements by the authors or if the drug is known to be prescribed after failure of first-line treatments.^{20–25}

Studies were definitively classified at risk of ITB/TLB only when consensus was reached. We provided a description of the main features of the studies at risk of bias, the biased analysis performed and information on which direction the bias could have modified study conclusions.

Data analysis

Data are summarised as number (%) for qualitative variables and median (range) for continuous variables. Characteristics of biased and unbiased papers were compared by Fisher's exact test or Mann-Whitney test, as appropriate.

RESULTS

Among the 1029 records retrieved, 78 articles were selected (figure 3). Each study included a median of 6393 (range 49–341 749) patients. The studies were mostly published in rheumatology journals (n=71; 91%), were not industry funded (n=48; 62%) and had safety primary outcomes (n=46; 59%) (table 1). Overall, 13 (17%) studies were considered at risk of time-dependent biases: 8 (10%) were at risk of ITB and 5 (6%) were at risk of TLB. Factors found to be associated with the risk of time-dependent biases were the continent of origin of the corresponding

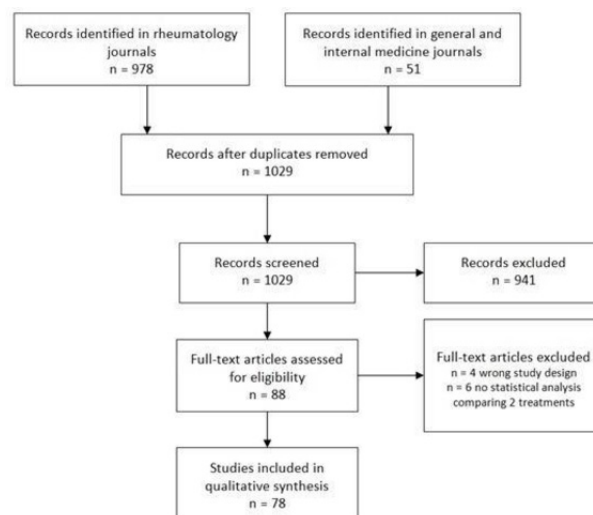


Figure 3 Flow chart of study selection.

Table 1 Main characteristics of the included studies at risk or not of immortal time bias (ITB) or time-lag bias (TLB)

	Overall sample	At risk of ITB or TLB	Not at risk of ITB or TLB	P value*
Studies, n	78 (100)	13 (17)	65 (83)	
Continent of corresponding author				0.002
North America	36 (46)	2 (15)	34 (52)	
Europe	36 (46)	7 (54)	29 (45)	
Asia	6 (8)	4 (31)	2 (3)	
Industry funded	30 (38)	3 (23)	27 (41)	0.349
Journal				
<i>Annals of Rheumatic Disease</i>	32 (41)	—	—	
<i>Arthritis & Rheumatology</i>	20 (26)	—	—	
<i>Rheumatology</i>	10 (13)	—	—	
<i>Seminars in Arthritis and Rheumatology</i>	7 (9)	—	—	
<i>BMJ</i>	5 (7)	—	—	
<i>Osteoarthritis and Cartilage</i>	2 (2)	—	—	
<i>JAMA</i>	1 (1)	—	—	
<i>Lancet</i>	1 (1)	—	—	
Journal impact factor, median (IQR)	12.8 (6.9–12.8)	4.8 (4.8–12.8)	12.8 (6.9–12.8)	0.033
Data sources				
Registries/cohorts	57 (73)	11 (85)	46 (71)	
Administrative databases	21 (27)	2 (15)	19 (29)	
Disease†				0.0002
Inflammatory arthritides	51 (63)	6 (46)	45 (67)	
Osteoarthritis	12 (16)	0 (0)	12 (18)	
Connective tissue diseases	9 (11)	4 (31)	5 (8)	
Vasculitides	3 (4)	3 (23)	0 (0)	
Gout/hyperuricaemia	3 (4)	0 (0)	3 (4)	
Other	2 (2)	0 (0)	2 (3)	
Intervention investigated‡				0.317
bDMARDs	41 (50)	6 (43)	35 (51)	
csDMARDs	13 (16)	5 (36)	8 (12)	
Analgesics/NSAIDs/corticosteroids	6 (7)	1 (7)	5 (7)	
Drugs for gout	3 (4)	0 (0)	3 (4)	
Other	19 (23)	2 (14)	17 (26)	
Primary outcome				0.958
Effectiveness	18 (23)	3 (23)	15 (23)	
Safety	46 (59)	7 (54)	39 (60)	
Survival	5 (7)	1 (8)	4 (6)	
Safety±effectiveness±survival	9 (11)	2 (15)	7 (11)	
Total sample size, median (range)	6393 (49–341 749)	602 (49–48 782)	6806 (75–341 749)	0.131
Statistical analysis†				
Kaplan-Meier curves	32 (30)	6 (31)	26 (30)	
Time-fixed variable Cox regression	51 (48)	12 (63)	39 (45)	
Time-dependent variable Cox regression	16 (15)	0 (0)	16 (18)	
MSM	4 (4)	0 (0)	4 (5)	
Other	3 (3)	1 (6)	2 (2)	
Adjustment for baseline confounders				
Propensity score	32 (41)	3 (23)	29 (45)	
Cox regression	34 (44)	7 (54)	27 (41)	
None	5 (6)	3 (23)	2 (3)	
Other	7 (9)	0 (0)	7 (11)	

Data are expressed as number (%), if not otherwise specified.

*Comparison between studies at risk or not of TLB/ITB.

†More than one choice for each study.

bDMARD, biologic disease-modifying antirheumatic drug; csDMARD, conventional synthetic disease-modifying antirheumatic drug; MSM, marginal structural models; NSAID, non-steroidal anti-inflammatory drug.

author ($p=0.002$), the impact factor of the journal ($p=0.033$) and the disease investigated ($p=0.0002$) (table 1).

Studies at risk of ITB

Eight (10%) studies^{26–33} were considered at risk of misclassified ITB and none were at risk of excluded ITB. Most of the studies at risk of ITB were published in journals with a median

(IQR) impact factor of 4.8 (4.6–9.8), investigated treatment for systemic connective tissue diseases or vasculitis ($n=7$) and were non-industry funded ($n=7$). The statistical tools more commonly used were time-fixed Cox regression and Kaplan-Meier curves (table 2).

In studies with a hierarchical treatment exposure model (ie, treated vs untreated), misclassified ITB occurred when (1)

Table 2 Main features of studies at risk of 'misclassified' immortal time bias

Author, year	Disease	Patients (n)	Outcome	Drug exposure	Statistical analysis	HR or RR for primary outcome(s)*	Description of the bias	Details	Potential effect of the bias on study results
Azar <i>et al</i> ²⁷ 2014	GPA	89	Disease relapse	RTX+GC followed or not by maintenance therapy (AZA or MMF or MTX)	KM, time-fixed Cox	0.43 (0.22–0.84) for patients receiving maintenance therapy	Waiting period to receive therapy	The time from remission achievement to maintenance therapy initiation potentially attributed to no treatment group	Hazard of relapse underestimated for patients receiving maintenance therapy, overestimated for those not receiving it
Ma <i>et al</i> ²⁰ 2016	SLE	23 084	Death	TCM/non-TCM	KM, time-fixed Cox	0.73 (0.68–0.78) for TCM users	Waiting period to receive therapy	TCM users or non-users classified within 3 years from diagnosis	Hazard of death underestimated for TCM-treated patients, overestimated for non-TCM-treated patients
Chen <i>et al</i> ²⁸ 2015	SLE	8628	Development of diabetes	No HCQ, HCQ cumulative dose <129 or ≥129 g	KM, time-fixed Cox	0.26 (0.18–0.37) for development of DI in HCQ≥129 vs <129 g	Treatment defined as a cumulative dosage to be reached	Two groups of patients were identified according to the cumulative dose of HCQ received (<129 vs ≥129 g)	Hazard of DI underestimated for HCQ≥129 g, overestimated for patients who received <129 g
Kurita <i>et al</i> ²⁹ 2015	PM/DM	49	Time to relapse or death for respiratory cause or serious adverse events	GC±immunosuppressor±tacrolimus	KM, time-fixed Cox	0.32 (0.14–0.75) for tacrolimus patients for RFS; 0.25 (0.10–0.66) for tacrolimus for disease-free survival	Waiting period to receive therapy	The tacrolimus group was defined as composed by patients treated with tacrolimus within 28 days from the start of immunosuppressive treatment	Hazard of relapse or serious adverse events underestimated for tacrolimus group, overestimated for non-tacrolimus group
Alba <i>et al</i> ¹⁶ 2014	GCA	106	Disease relapse	ACEI, ARB, no treatment	KM, time-fixed Cox	0.32 (0.12–0.81) for ARB versus nothing	Waiting period to receive therapy	Patients were labelled as ACEI or ARB users if started the drug during the first year after diagnosis. If not, they were considered non-users	Hazard of relapse underestimated for ACEI and ARB group, overestimated for non-treatment group
Roubille <i>et al</i> ³¹ 2017	RA	602	New-onset safety events including all-cause mortality, cardiovascular events, severe infection and fracture	Ever or never GC use	Time-fixed Cox	0.88 (0.62–1.27) for GC use	Treatment status defined as at least one prescription dispensed	The patients were classified into 2 groups by whether or not they received GC at least once over the 7 years of follow-up	Hazard of safety outcome events underestimated for GC users, overestimated for never GC users
van Daalen <i>et al</i> ²² 2017	AAV	323	Malignancy	CYC, RTX, CYC+RTX, no treatment	Poisson regression	4.61 (1.16–39.98) for CYC versus RTX; 3.05 (1.40–7.35) for CYC versus both CYC and RTX; 1.48 (0.60–4.36) for CYC versus none	Treatment defined as a cumulative dosage/number of treatments to be reached	Patients were classified according to their CYC and/or RTX exposure into the following categories: CYC treated, RTX treated, CYC and RTX treated, not treated with CYC or RTX	Hazard of malignancy underestimated for CYC+RTX group
Yap <i>et al</i> ²³ 2013	SLE	65	Renal survival	MMF as induction for all. MMF or AZA or CNi as maintenance	KM, time-fixed Cox, mixed-model analysis	Non-significant difference in log rank test among the 3 groups	Waiting period to receive therapy	In 23 patients, MMF was replaced with AZA after 16.5 (11.9 months (MMF-AZA group), and in 11 patients, MMF was replaced with CNi after 23.9 (20.4 months (MMF-CNi group)	Hazard of loss of renal function underestimated for MMF-AZA and MMF-CNi groups, overestimated for MMF-MMF group

* All results refer to adjusted analysis, when available.

† The authors did not provide a satisfactory definition of the group CYC+RTX: they did not specify whether this group included patients who received CYC+RTX at the same time (condition unusual but possible in clinical practice) or whether patients had received these two drugs over the follow-up but not necessarily at the same time. In the first case, the study would not be at risk of ITB.

AAV, ANCA-associated vasculitis; ACEI, ACE inhibitors; ARB, angiotensin II receptor blockers; AZA, azathioprine; CNi, calcineurin inhibitors; CYC, cyclophosphamide; DI, diabetes; GC, glucocorticoids; GCA, giant cell arteritis; GPA, granulomatosis with polyangiitis; HCQ, hydroxychloroquine; KM, Kaplan-Meier curves; MMF, mycophenolate mofetil; MTX, methotrexate; PM/DM, polymyositis/dermatomyositis; RA, rheumatoid arthritis; RFS, relapse-free survival; RTX, rituximab; SLE, systemic lupus erythematosus; TCM, traditional Chinese medicine.

waiting time to receive treatment was accounted for in the treatment group (n=5 studies)^{26 27 29 30 33}; (2) treatment status was defined as cumulative dosage reached/number of treatments received (n=2 studies)^{28 32}; and (3) treatment status was defined as at least one prescription dispensed (n=1 study).³¹ In all these cases, the hazard of the outcome was likely underestimated in the exposure group and overestimated in the comparator, which led to an exaggerated beneficial effect or a lowered risk of harms for the treated group.⁴ A description of mechanisms leading to bias and its potential effects on study results is found in [table 2](#).

In an additional 12 studies, the authors investigated differences between a cohort of new users and a comparator cohort of prevalent user patients. Although (1) the start of follow-up was different between the two groups (start of treatment for treated patients, entry in the cohort for comparator), (2) treated patients had received the comparator drug before entry in the study, and (3) person-time on the comparator drug for treated patients was excluded from the analysis, we did not classify these studies at risk of excluded ITB because the exclusion of person-time occurred in both treatment and comparison cohorts. Actually, patients who initiated the comparator drug and experienced an event of interest before being included in the cohort never entered the study, which led to underestimating the event rate also in this group. We checked with numerical simulations that no bias was induced by this design (data not shown).

Studies at risk of TLB

We considered 5 (6%) studies at risk of TLB: these studies compared the risk of cancer^{34–37} or tuberculosis³⁸ development in patients treated with conventional synthetic disease-modifying antirheumatic drugs (DMARD) versus biologic DMARDs. In four studies,^{34 35 37 38} the authors did not adjust for disease duration or past medication exposure, and in one study³⁶ the two groups were propensity score matched but only for cumulative steroid dose. [Table 3](#) summarises the main features of these studies.

DISCUSSION

We found ITB or TLB in about one in six observational studies of rheumatology published in leading journals. ITB was always due to ‘misclassification’ rather than ‘exclusion’ of the immortal time. TLB was observed when patients receiving conventional synthetic DMARDs were compared with those on biologic

therapy without accounting for disease duration or previous drug intake.

Some examples taken from the papers included in the present review can help better identify ITB in published literature. A misclassification of the immortal time should be searched in the following cases. (1) *When there is a waiting time between the start of follow-up (ie, time of diagnosis, entry in the cohort) and treatment initiation, and such time is incorrectly attributed to the exposed group.* For example, in one of the studies classified at risk of ITB,²⁹ patients were defined as tacrolimus ‘treated’ if they had received tacrolimus as maintenance therapy within 28 days from the first immunosuppressor used to induce remission. Because follow-up started after the achievement of remission for both treated and untreated patients, the waiting time to receive tacrolimus, by definition ‘immortal’ for tacrolimus-treated patients, was wrongly attributed to this group, thereby potentially conferring to tacrolimus a spurious protective advantage on relapse prevention. (2) *When the exposure is handled as ‘ever’ or ‘never’ drug intake over follow-up.* An example is one study³⁰ investigating the impact of traditional Chinese medicine on survival of patients with systemic lupus. The authors classified patients as traditional Chinese medicine ‘users’ or ‘never users’ if they had ever (or never) received such treatment within 3 years from study entry. In line with the previous example, the wrong attribution of the immortal time (from the start of follow-up to the introduction of treatment for ‘ever-treated’ patients) likely led to underestimating the rate of deaths in ‘treated’ patients and overestimating it in ‘untreated’ patients. (3) *When a given duration of drug use or a given cumulative dose is required for a participant to be classified as exposed.* In one study,²⁸ patients were classified according to their cumulative dose of hydroxychloroquine (<129 or ≥129 g) over the follow-up. Receiving more hydroxychloroquine was associated with a 74% reduction in risk of diabetes developing. Again, this apparent longer survival free from diabetes is at least in part an artefact of the wrong attribution of the immortal time to patients with higher hydroxychloroquine cumulative intake.

Moreover, even if we did not find any case of excluded ITB, attention should be paid to avoiding discarding the immortal time from the analysis ([figure 1](#), b vs c).

Several methodological aspects should be considered when performing comparative effectiveness observational studies. In an ideal scenario, a treatment should be compared with another

Table 3 Main features of the studies at risk of time-lag bias

Author, year	Disease	Patients (n)	Outcome	Drug exposure	Statistical analysis	Control for baseline confounders	HR or RR for primary outcome(s)*
Cordtz <i>et al</i> ³⁴ 2016	IA	13 905	Virus-associated cancers	Ever or never bDMARD user	Cox regression	Cox regression	0.9 (0.7–1.2) for ever versus never bDMARD users
Silva-Fernandez <i>et al</i> ³⁷ 2016	RA	425	Malignancy recurrence	csDMARDs, TNFi, RTX	Cox regression	Cox regression	0.55 (0.35–0.86) for TNFi versus csDMARDs; 0.43 (0.10–1.80) for RTX versus csDMARDs
Kim <i>et al</i> ³⁶ 2016	RA	44 534	High-grade cervical dysplasia and cervical cancer	csDMARDs, bDMARDs	Cox regression	PS matching	1.32 (0.86–2.01) for bDMARDs versus csDMARDs
Arkema <i>et al</i> ³⁸ 2015	RA	48 782	Tuberculosis infection	Ever or never bDMARD user	Cox regression	Cox regression	4.4 (2.3–8.5) for ever versus never bDMARDs
Dreyer <i>et al</i> ³⁵ 2013	IA	9696	Malignancy	Ever or never TNFi user	Cox regression	Cox regression	1.02 (0.80–1.30) for ever versus never TNFi users

*All results refer to adjusted analysis.

ARB, angiotensin II receptor blocker; bDMARD, biologic disease-modifying antirheumatic drug; csDMARD, conventional synthetic disease-modifying antirheumatic drug; IA, inflammatory arthritides; PS, propensity score; RA, rheumatoid arthritis; RTX, rituximab; TNFi, tumour necrosis factor inhibitor.

having the same indications and that might be used interchangeably.³⁹ However, if we want to compare drugs given at different disease stages (ie, patients on first-line vs second-line drugs), some points should be considered to avoid time-related biases, in particular TLB. First, different disease durations can be associated with the outcome of interest. Second, if the exposure to a first-line treatment is associated with the development of the outcome, even if after a long period, the attribution of the event to the first-line or second-line drug becomes challenging. In this setting, statistical analyses taking into account latency time windows and disease duration are needed.⁷

This study has some limitations. First, the choice to select high-impact journals could have underestimated the presence of time-related biases in the literature. Second, despite we tried to maximise the sensitivity of our search, we could have missed studies not providing in the title or abstract the keywords used. Moreover, we could identify which direction the bias could have affected study results but not quantify its impact on point estimates. Indeed, the potential impact of ITB on treatment effect estimates depends on the respective amounts of misclassified/excluded and correctly classified person-time, as well as number of events. Methods to quantify the magnitude of ITB require those data being detailed.^{40,41} Unfortunately, none of the articles at risk of ITB presented all that information. We could anticipate that the choice of a 3-year waiting time to define treatment exposure in the study from Ma *et al*³⁰ could have impacted more the estimation of the treatment effect than the choice to classify patients as receiving or not maintenance therapy within 28 days from the achievement of remission in the study from Kurita *et al*.²⁹

In conclusion, time-related biases are common in the rheumatology literature and can be avoided by using appropriate study designs and statistical analyses such as time-dependent Cox regression or landmark analysis.⁴² Attention should be paid to avoid a hierarchical treatment exposure based, for example, on (1) 'ever' or 'never' intake of the drug, (2) treated designation if the drug has been received for at least a given period, and (3) treated designation if a given cumulative dosage intake of the drug has been reached. Moreover, the comparison of groups of patients at different disease stages should consider differences of disease duration, if they are believed to potentially confound the study outcome. A better recognition of such biases could help improve study designs and the interpretation of published observational studies.

Acknowledgements The authors thank Laura Smales for proofreading the manuscript.

Contributors MI, RP and PR contributed to the conception and design of the work. MI and CR contributed to the acquisition of data. MI, RP, CR and PR contributed to the analysis and interpretation of data. MI, RP, CR and PR helped draft the paper and critically revised it for important intellectual content. All the authors gave final approval of the version submitted.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement All the data collected are available upon request.

REFERENCES

- 1 Suissa S. Immortal time bias in observational studies of drug effects. *Pharmacoepidemiol Drug Saf* 2007;16:241–9.
- 2 Austin PC, Mamdani MM, van Walraven C, *et al*. Quantifying the impact of survivor treatment bias in observational studies. *J Eval Clin Pract* 2006;12:601–12.
- 3 Glesby MJ, Hoover DR. Survivor treatment selection bias in observational studies: examples from the AIDS literature. *Ann Intern Med* 1996;124:999–1005.
- 4 Wolkewitz M, Allignol A, Harbarth S, *et al*. Time-dependent study entries and exposures in cohort studies can easily be sources of different and avoidable types of bias. *J Clin Epidemiol* 2012;65:1171–80.
- 5 Giobbie-Hurder A, Gelber RD, Regan MM. Challenges of guarantee-time bias. *JCO* 2013;31:2963–9.
- 6 Suissa S. Lower risk of death with SGLT2 inhibitors in observational studies: real or bias? *Diabetes Care* 2018;41:6–10.
- 7 Suissa S, Azoulay L. Metformin and the risk of cancer: time-related biases in observational studies. *Diabetes Care* 2012;35:2665–73.
- 8 Lévesque LE, Hanley JA, Kezouh A, *et al*. Problem of immortal time bias in cohort studies: example using statins for preventing progression of diabetes. *BMJ* 2010;340.
- 9 van Walraven C, Davis D, Forster AJ, *et al*. Time-dependent bias was common in survival analyses published in leading clinical journals. *J Clin Epidemiol* 2004;57:672–82.
- 10 Wallis CJD, Saskin R, Narod SA, *et al*. Estimating the effect of immortal-time bias in urological research: a case example of testosterone-replacement therapy. *BJU Int* 2017;120:584–90.
- 11 Jones M, Fowler R. Immortal time bias in observational studies of time-to-event outcomes. *J Crit Care* 2016;36:195–9.
- 12 Suissa S. Reduced mortality with sodium-glucose cotransporter-2 inhibitors in observational studies: avoiding immortal time bias. *Circulation* 2018;137:1432–4.
- 13 Suissa S, Klil-Drori AJ. Statins in multiple myeloma: survival benefit or immortal time bias? *JCO* 2017;35:1377–8.
- 14 Matok I, Azoulay L, Yin H, *et al*. Immortal time bias in observational studies of drug effects in pregnancy. *Birth Defects Res A Clin Mol Teratol* 2014;100:658–62.
- 15 Crowson CS. Immortal time bias: Comment on the article by Gwinnutt *et al*. *Arthritis Rheumatol* 2017;69.
- 16 Tong SYC, Cheng AC, Denholm JT. Immortal time bias in assessing evidence-based care processes for *Staphylococcus aureus* Bacteremia. *JAMA Intern Med* 2018;178:295–6.
- 17 Targownik LE, Suissa S, Understating SS. Understanding and avoiding immortal-time bias in gastrointestinal observational research. *Am J Gastroenterol* 2015;110:1647–50.
- 18 Desai RJ, Solomon DH. Use of large healthcare databases for rheumatology clinical research. *Curr Opin Rheumatol* 2017;29:138–43.
- 19 Moher D, Liberati A, Tetzlaff J, *et al*. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Ann Intern Med* 2009;151:264–9.
- 20 Kloppenburg M, Kroon FPB, Blanco FJ, *et al*. 2018 update of the EULAR recommendations for the management of hand osteoarthritis. *Ann Rheum Dis* 2019;78:16–24.
- 21 Kowal-Bielecka O, Fransen J, Avouac J, *et al*. Update of EULAR recommendations for the treatment of systemic sclerosis. *Ann Rheum Dis* 2017;76:1327–39.
- 22 Ramiro S, Smolen JS, Landewé R, *et al*. Pharmacological treatment of psoriatic arthritis: a systematic literature review for the 2015 update of the EULAR recommendations for the management of psoriatic arthritis. *Ann Rheum Dis* 2016;75:490–8.
- 23 Smolen JS, Landewé R, Bijlsma J, *et al*. EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs: 2016 update. *Ann Rheum Dis* 2017;76:960–77.
- 24 van der Heijde D, Ramiro S, Landewé R, *et al*. 2016 update of the ASAS-EULAR management recommendations for axial spondyloarthritis. *Ann Rheum Dis* 2017;76:978–91.
- 25 Yates M, Watts RA, Bajema IM, *et al*. EULAR/ERA-EDTA recommendations for the management of ANCA-associated vasculitis. *Ann Rheum Dis* 2016;75:1583–94.
- 26 Alba MA, García-Martínez A, Prieto-González S, *et al*. Treatment with angiotensin II receptor blockers is associated with prolonged relapse-free survival, lower relapse rate, and corticosteroid-sparing effect in patients with giant cell arteritis. *Semin Arthritis Rheum* 2014;43:772–7.
- 27 Azar L, Springer J, Langford CA, *et al*. Rituximab with or without a conventional maintenance agent in the treatment of relapsing granulomatosis with polyangiitis (Wegener's): a retrospective single-center study. *Arthritis Rheumatol* 2014;66:2862–70.
- 28 Chen YM, Lin CH, Lan TH, *et al*. Hydroxychloroquine reduces risk of incident diabetes mellitus in lupus patients in a dose-dependent manner: a population-based cohort study. *Rheumatology* 2015;54:1244–9.
- 29 Kurita T, Yasuda S, Oba K, *et al*. The efficacy of tacrolimus in patients with interstitial lung diseases complicated with polymyositis or dermatomyositis. *Rheumatology* 2015;54:39–44.
- 30 Ma YC, Lin CC, Li CI, YC M, CI L, *et al*. Traditional Chinese medicine therapy improves the survival of systemic lupus erythematosus patients. *Semin Arthritis Rheum* 2016;45:596–603.
- 31 Roubille C, Rincheval N, Dougados M, *et al*. Seven-year tolerability profile of glucocorticoids use in early rheumatoid arthritis: data from the ESPOIR cohort. *Ann Rheum Dis* 2017;76:1797–802.
- 32 van Daalen EE, Rizzo R, Kronbichler A, *et al*. Effect of rituximab on malignancy risk in patients with ANCA-associated vasculitis. *Ann Rheum Dis* 2017;76:1064–9.

- 33 Yap DY, Ma MK, Mok MM, *et al.* Long-term data on corticosteroids and mycophenolate mofetil treatment in lupus nephritis. *Rheumatology* 2013;52:480–6.
- 34 Cordtz R, Mellemkjær L, Grintborg B, *et al.* Risk of virus-associated cancer in female arthritis patients treated with biological DMARDs-a cohort study. *Rheumatology* 2016;55:1017–22.
- 35 Dreyer L, Mellemkjær L, Andersen AR, *et al.* Incidences of overall and site specific cancers in TNF α inhibitor treated patients with rheumatoid arthritis and other arthritides - a follow-up study from the DANBIO Registry. *Ann Rheum Dis* 2013;72:79–82.
- 36 Kim SC, Schneeweiss S, Liu J, *et al.* Biologic disease-modifying antirheumatic drugs and risk of high-grade cervical dysplasia and cervical cancer in rheumatoid arthritis: a cohort study. *Arthritis Rheumatol* 2016;68:2106–13.
- 37 Silva-Fernández L, Lunt M, Kearsley-Fleet L, *et al.* The incidence of cancer in patients with rheumatoid arthritis and a prior malignancy who receive TNF inhibitors or rituximab: results from the british Society for rheumatology biologics Register- Rheumatoid arthritis. *Rheumatology* 2016;55:2033–9.
- 38 Arkema EV, Jonsson J, Baecklund E, *et al.* Are patients with rheumatoid arthritis still at an increased risk of tuberculosis and what is the role of biological treatments? *Ann Rheum Dis* 2015;74:1212–7.
- 39 Schneeweiss S. Developments in post-marketing comparative effectiveness research. *Clin Pharmacol Ther* 2007;82:143–56.
- 40 Suissa S. Immortal time bias in pharmaco-epidemiology. *Am J Epidemiol* 2008;167:492–9.
- 41 Liang C, Seeger JD, Dore DD. Implications of immortal person-time when outcomes are nonfatal. *Ann Epidemiol* 2016;26:212–7.
- 42 Cho IS, Chae YR, Kim JH, *et al.* Statistical methods for elimination of guarantee-time bias in cohort studies: a simulation study. *BMC Med Res Methodol* 2017;17:017–405.