PHARMACOEPIDEMIOLOGY AND DRUG SAFETY 2011; **20**: 119–130
Published online 7 December 2010 in Wiley Online Library (wileyonlinelibrary.com) **DOI**: 10.1002/pds.2046

ORIGINAL REPORT

Cancer risk with tumor necrosis factor alpha (TNF) inhibitors: meta-analysis of randomized controlled trials of adalimumab, etanercept, and infliximab using patient level data

Johan Askling[1]*, Kyle Fahrbach[2], Beth Nordstrom[2], Susan Ross[2], Christopher H. Schmid[3] and Deborah Symmons[4]

[1]*Department of Medicine Solna, Karolinska Institutet, Stockholm, Sweden*
[2]*United BioSource Corporation, Lexington, MA, USA*
[3]*Institute for Clinical Research and Health Policy Studies, Tufts Medical Center, Boston, MA, USA*
[4]*ARC Epidemiology Unit, University of Manchester, UK*

ABSTRACT

Purpose Uncertain short- and long-term cancer risks with anti-TNF therapies is a concern, and led to a recent black box warning. This meta-analysis, requested by the European Medicines Agency, aimed at better assessing short-term risks by using meta-analytic techniques based on individual patient data from all corporate-sponsored randomized controlled trials (RCTs) of adalimumab, etanercept, and infliximab.

Methods All 74 RCTs of TNF inhibitors of at least 4 weeks duration were provided to independent investigators, including case narratives for events occurring between trial start until 30 days after planned end of treatment and indicating a possible cancer. Relative risks were estimated using Bayesian piecewise exponential models.

Results One hundred thirty (0.84%) of 15 418 individuals randomized to anti-TNF therapy were diagnosed with cancer, compared to 48 (0.64%) of 7486 individuals randomized to comparators. The relative risks associated with all anti-TNF were 0.99 (95%CI 0.61–1.68) for cancers excluding non-melanoma skin cancer (NMSC), and 2.02 (95%CI 1.11–3.95) for NMSC. There were indications of differences in the relative risks for the three anti-TNF drugs, but also of differences across the cancer rates in the three comparator arms for adalimumab, etanercept, and infliximab.

Conclusions Despite a reassuring overall short-term risk, we could neither refute nor verify that individual anti-TNF therapies affect the short-term clinical emergence of cancer. Despite representing the best available evidence, statistical precision, and differences in baseline cancer risk and reporting detail between trials of adilumumab, etanercept, and infliximab hampered distinction of *drug-specific* from *trial* effects, illustrating the challenges in safety-assessments using RCT meta-analyses. Long-term risk assessment requires observational studies.

KEY WORDS — TNF, rheumatoid arthritis, inflammatory bowel disease, trial, cancer, meta-analysis

Received 17 September 2009; Revised 2 August 2010; Accepted 4 August 2010

INTRODUCTION

TNF-alpha inhibitors, including adalimumab (Humira®), etanercept (Enbrel®), and infliximab (Remicade®), are indicated for the treatment of several chronic inflammatory conditions. In the most common of these, Rheumatoid Arthritis (RA), well over 20% of all patients in the US and several other countries are exposed to such therapies.

In addition to its key role in inflammation, TNF alpha exerts several effects that may be of relevance for carcinogenesis, tumor growth, and the time-point of clinical detection of incipient cancers.[1,2] Before their introduction, concerns were raised that *anti*-TNF alpha therapies would increase the risk of cancer emergence. Whereas some of these effects may require long follow-up, others (e.g., accelerated growth of incipient cancers, altered tumor phenotype, and facilitated detection) should be detectable in the short term.

* Correspondence to: J. Askling, Clinical Epidemiology Unit T2:01, Karolinska University Hospital Solna, SE-171 76 Stockholm, Sweden.
E-mail: johan.askling@ki.se

Short-term risks for cancer emergence with anti-TNF drugs have been assessed in one pooled analysis of nine RA randomized controlled trials (RCTs) of infliximab and adalimumab (suggesting a tripled risk with anti-TNF[3]), one meta-analysis of nine RA RCTs of etanercept (suggesting a near doubled risk with anti-TNF[4]), and in one pooled analysis of 18 RA trials of adalimumab, infliximab, and etanercept (suggesting a moderate excess of cancers in the treated arms[5]). The interpretation of these studies is, however, not straightforward. In all studies, the number of events was low, and only a subset of all RCTs (and only in RA) was used. Inherent differences in design and analysis preclude comparisons across these assessments, let alone across the three drugs. Since cancer is not a binary event with an easily defined date of onset, and since 'cancer' was not a predefined outcome, any event definition is inevitably post hoc and study-specific. Reliance upon the adverse events reported as 'cancer' in the trial publications did not (as acknowledged[5]) offer sufficient detail. Indeed, the multiple effects that anti-TNF therapy may have on cancer emergence beg the use of alternative definitions of 'cancer'.

Observational studies have accrued somewhat longer follow-up times, so far not observed any overall increase in cancer risk with anti-TNF treatment in class-analyses,[6–9] but typically not assessed drug-specific risks by time of follow-up. In the only such study to date, dissimilar cancer risks across the three drugs were, however, observed during the first year of follow-up.[10]

Thus, despite several attempts to assess cancer risks, the uncertainty regarding short- and long-term risks of cancer with anti-TNF therapy remains, as illustrated by the FDA black-box warning issued in June 2009.[11] With respect to short-term risks using trial data, there is a need for a comprehensive assessment that not only increases statistical power but also addresses the methodological challenges inherent in the previous analyses. It should include a robust definition and adjudication of 'cancer', and explore differences across individual drugs and conditions.

The present study followed a request by the European Medicines Evaluation Agency (EMEA) to the market authorization holders (MAH) of the anti-TNF agents to conduct a joint meta-analysis of their RCT-data. The meta-analysis was funded by the MAHs who also contributed the study data, but was prospectively planned and conducted by an independent research group working with an independent academic steering committee and a separate and independent event adjudication committee. Over 22 000 patients from 74 RCTs (13 adalimumab, 12 etanercept, and 6 infliximab trials) in RA, *plus* 43 trials of these drugs in any other condition, were included. Compared to previous analyses, our study included a considerably larger population with RA, all other investigated indications, an individual patient data meta-analysis design, classwise and drug-specific assessments, and an extensive search and blinded review of adverse events potentially representing 'cancer'. A number of cancer definitions were tested.

SUBJECTS AND METHODS

Study selection

Eligible RCTs were all placebo or standard-care controlled trials of adalimumab, etanercept, or infliximab sponsored by Abbott, Amgen/Wyeth, Centocor/Schering-Plough, or other corporate sponsors of the same products. Eligible patients were aged at least 18 years and received at least one treatment dose. Studies in which all patients had received TNF inhibitor treatment before study start were excluded, as were studies of <4 weeks duration. Each sponsor was provided with these criteria, and searched their trial databases to identify eligible trials. Trials primarily funded from sources other than corporate sponsorship and for which complete individual patient data were not accessible through any of the sponsors, (e.g., a trial of etanercept in Wegener's granulomatosis[12]) were therefore not eligible.

Data extraction, exposure

For each trial, the sponsor created a patient-level dataset containing data pertaining to the

study (treated condition, TNF-inhibitor agent, comparison group therapy, study location, exclusion criteria, and required pre-trial work-up),
patient (gender, year of birth, disease duration, prior and concomitant cytotoxic or immune-suppressive therapy), and
treatment (dates and doses of study drug during the trial, duration of intended and actual treatment, date and reasons for withdrawal, and date of last patient contact).

Data extraction and definition, outcome

To identify all potential *cancer events*—irrespective of report to a regulatory agency or inclusion in the published report of the trial—each sponsor's clinical and safety database was searched using a list of over 1800 pre-defined terms. These terms included defined

DOI: 10.1002/pds

cancer types, terms for conditions potentially harboring an invasive cancer (e.g., 'gammopathy'), and terms not more specific than neoplasia-related text strings such as 'neopl'.

The study period encompassed trial initiation until 30 days after the planned end of treatment. Person-time at risk was calculated from start of treatment to the event date, or (if no event) to the latest of date of withdrawal, date of last personal contact with the patient, or 30 days after end of planned treatment.

For all events indicating a possible malignancy, sponsors submitted narratives using a pre-defined standardized format that described the history and course of the event. Narratives were subsequently blinded with respect to sponsor and treatment assignment. Based on the information in these narratives two oncologist adjudicators determined the probabilities (definite, probable, possible, unlikely) that the event was (i) a cancer including its site, (ii) prevalent at the trial start, and (iii) a recurrence of a pre-trial cancer. The earliest reported date of clinical evidence of the cancer and the date of clinical diagnosis were also abstracted. Each adjudicator reviewed all narratives. Consensus was required.

Based on the above probabilities, outcome was examined using three nested cancer definitions (see Appendix). Outcome A included all cancer events (definite or probable cancers) diagnosed during the study period, using the date of diagnosis as the event date. Outcome B excluded from A those events, in retrospect, judged definitely prevalent on the basis of a first reported date of sign or symptom pre-dating the trial start. Outcome C excluded from B those events which, for other reasons than first date of sign/symptom, were judged by the oncologists to be probably prevalent at trial start.

Because of the more 'benign' course of most non-melanoma skin cancers (NMSCs, a.k.a. keratinocyte cancer), the potential for increased detection related to the efficacy or side-effects of active therapy (e.g., rash), and the potential difficulty in assessing risks for lesions not always sent for histopathology cancer risks were assessed as all cancers including NMSC, all cancers excluding NMSC, and NMSC alone. Trials were grouped according to the treated condition: (i) primary use conditions (RA, psoriasis, psoriatic arthritis, ankylosing spondylitis, and Crohn's disease) which were the pre-defined category of main interest; (ii) RA only; (iii) investigational use/single compound approved conditions; and (iv) all conditions. Supplementary Table S2 displays drug exposure and events per condition.

Statistical analysis

We modeled hazard ratios including 95% credible intervals and Bayesian posterior probabilities using a time-to-event piecewise exponential Bayesian survival model dividing time into two intervals (0–3, >3 months) with constant baseline hazard in each interval[13] and assuming homogeneous event rates across studies (see Appendix). Models assessing *class effects* considered all anti-TNF agents as one class and included a single parameter for anti-TNF treatment versus comparator. Models assessing *drug-specific effects* employed separate parameters for each drug. These models also investigated differences in underlying risk of being diagnosed with cancer in the three sets of *trials*, i.e., whether the (control-group) cancer risk in adalimumab trials was different from the corresponding risks in etanercept and in infliximab trials, as a 'sponsor-specific control-group effect'. We also fit *separate models for each drug*. All models were adjusted for (where appropriate) age, gender, concurrent use of cytotoxic immunosuppressant therapy, condition, and RA disease duration. The proportion of missing data for all variables was less than 1%. Since the events were so rare (virtually all studies has zero events in at least one arm), stratification according to trial could not be performed.

Models with different time intervals, weakly informative priors (see Appendix), random study effects, non-proportional hazards, interactions, and time-dependent treatment were also investigated. These more complex models did not provide any better fit than the simple piecewise model. A Peto meta-analysis was run for all outcomes, and resulted in relative risks largely identical to those of the original analyses. Further sensitivity analyses (cancers events judged possible, probable, or definite cancer, excluding events with event dates during the first 90 days of the study period, excluding events with event dates during the last 30 days of the study period, RA trials, secondary use conditions, a Cox regression, respectively) were run for all outcome definitions, using cancer defined as outcome C, and all revealed a similar overall pattern to the primary analysis. SAS version 9.1 and OpenBUGS 3.0.2, run using BRugs (R version 2.6.2), were used.

RESULTS

Study characteristics

Seventy four (23 adalimumab, 28 etanercept, and 23 infliximab) trials were included. Across all studies, 22 904 patients were included, with an approximate 2.1:1 ratio between patients allocated to anti-TNF therapy and

DOI: 10.1002/pds

Table 1. Number of patients and person-years of follow-up in the meta-analysis, by drug, by treatment (Tx) and control (cont) arm, and by time since trial start

	Adalimumab		Etanercept		Infliximab		All Anti-TNF	
	Tx	Cont	Tx	Cont	Tx	Cont	Tx	Con
All trials								
N patients	4709	2646	6153	3063	4544	1769	15 406	7478
N person-years	2861	1466	4404	2073	2431	862	9696	4401
N patients at risk by time since trial start (months)								
0	4709 (100%)	2646 (100%)	6153 (100%)	3063 (100%)	4544 100%)	1769 (100%)	15 406 (100%)	7478 (100%)
1	4703 (100%)	2644 (100%)	5920 (96%)	2946 (96%)	4249 (94%)	1611 (91%)	14 872 (97%)	7201 (96%)
3	4085 (87%)	2173 (82%)	4962 (81%)	2659 (87%)	3969 (87%)	1422 (80%)	13 016 (84%)	6254 (84%)
6	2300 (49%)	1095 (41%)	2230 (36%)	1102 (36%)	2195 (48%)	745 (42%)	6725 (44%)	2942 (39%)
12	818 (20%)	384 (15%)	1352 (22%)	575 (19%)	312 (7%)	67 (4%)	2482 (16%)	1026 (14%)
18	393 (17%)	183 (6.9%)	958 (16%)	396 (13%)	0 (0%)	1 (0.06%)	1351 (8.8%)	580 (8%)
Primary use conditions								
N patients	4709	2646	4570	2248	3576	1289	12 855	6183
N person-years	2859	1465	3275	1494	1953	650	8088	3608

Primary use conditions were defined as rheumatoid arthritis, psoriasis, psoriatic arthritis, ankylosing spondylitis, and Crohn's disease.

to the comparator arms, and a 2.2:1 ratio between the corresponding person-times of follow-up (Table 1). The median duration was <6 months for each of the three anti-TNF agents. Infliximab trials had a somewhat shorter follow-up than adalimumab and etanercept trials (Table 1). In the primary use conditions the total numbers of patients treated with adalimumab and etanercept were similar, but the infliximab-treated group was 25% smaller. Moreover, whereas the ratio between treatment and comparator patient numbers was fairly similar for adalimumab (1.8:1) and etanercept (2.0:1) trials, it was higher for infliximab (2.8:1). As a result of these differences, the infliximab comparator population experience was only half the size of the adalimumab and etanercept comparators (Table 1).

Eighty-three per cent of patients were treated for one of the primary use conditions (see supplementary Table S1 for results per condition). Forty-seven per cent were treated for RA. Mean age was 49 years (adalimumab), 53 years (etanercept), and 48 years (infliximab). The female/male distribution was 60/40 (adalimumab), 51/49 (etanercept), and 53/47 (infliximab). About one-third of all patients, similar for all three agents, were on concomitant oral steroids. Ninety-five per cent of all infliximab trial patients received concomitant methotrexate or immunosuppressive treatment. The corresponding figures for adalimumab and etanercept were 71 and 28%, respectively. The vast majority (similar for all three drugs) of patients had undergone pre-trial screening with chest X-ray and liver function tests. Around 15% of infliximab trial patients entered an open-label extension after the double-blind phase. The corresponding figures for etanercept and adalimumab were 60 and 85%, respectively.

Cancer occurrence, number of events

Presented results refer to the primary use conditions. Table 2 displays numbers and proportions of patients

Table 2. Counts (numbers and %) of cancer events including non-melanoma skin cancer by outcome definition, by drug, by treatment (Tx) and control (cont) arm for all trials

	Adalimumab		Etanercept		Infliximab		All Anti-TNF	
	Tx	Cont	Tx	Cont	Tx	Cont	Tx	Cont
Patients	4709 (100%)	2646 (100%)	6156 (100%)	3069 (100%)	4553 (100%)	1771 (100%)	15 418 (100%)	7486 (100%)
Flagged events	270 (5.73)	127 (4.80)	306 (4.97)	158 (5.15)	196 (4.30)	48 (2.71)	772 (5.01)	333 (4.45)
Adjudicated events	97 (2.06)	30 (1.13)	130 (2.11)	63 (2.05)	57 (1.25)	13 (0.73)	284 (1.84)	106 (1.42)
Outcome A[*]	41 (0.87)	15 (0.57)	57 (0.93)	25 (0.81)	32 (0.70)	8 (0.45)	130 (0.84)	48 (0.64)
Outcome B[†]	29 (0.62)	12 (0.45)	45 (0.73)	24 (0.78)	20 (0.44)	6 (0.34)	94 (0.61)	42 (0.56)
Outcome C[‡]	17 (0.36)	2 (0.08)	29 (0.47)	16 (0.52)	13 (0.28)	2 (0.11)	59 (0.38)	20 (0.27)

[*]Outcome A was defined as all cancer events (definite or probable cancers) diagnosed during the study period, using the date of diagnosis as the event date, irrespective of judgments on pre-trial prevalence.
[†]Outcome B was defined as all cancer events (definite or probable cancers) diagnosed during the study period, but excluding events in retrospect judged definitely prevalent on the basis of a first reported date of sign or symptom pre-dating the trial.
[‡]Outcome C was defined as all cancer events (definite or probable cancers) diagnosed during the study period, excluding both events with a first reported date of sign or symptom before the study period, *and* events which for other reasons were judged by the oncologists to be *probably* prevalent at trial start.

DOI: 10.1002/pds

with any potential event *flagged*, events that were *adjudicated* (i.e., those that were not clearly benign), and events that *fulfilled the definitions of Outcomes A, B, and C*. 40% of adjudicated narratives had minimal information beyond the event term (cancer): 68% from etanercept trials, 27% from adalimumab trials, and 5% from infliximab trials. The proportion of events flagged and adjudicated, but *not* fulfilling any of the outcome definitions, also differed across the three agents. It was lower in infliximab trials than in adalimumab and etanercept trials (Table 2). Depending on the outcome definition used, 0.3–0.8% of all patients developed a cancer. Supplementary Table S3 displays counts of events, by cancer site, by treatment status (anti-TNF vs. comparator).

The adjudication process judged almost 50% of all cancers *without* any reported signs or symptoms before the trial as still *probably prevalent* at trial start (i.e., if investigated or reported more in depth, chances were 50–90% that there would have been clinical signs/symptoms from the incipient cancer at the time of trial start, see Appendix). Since one of the concerns is that anti-TNF therapy might make tumors progress faster or slower and present at a more or less advanced stage than expected, exclusions based on typical tumor progression rates might under- or over-estimate the true occurrence of cancer. Judgments factoring in assumptions of typical growth speed and induction times therefore represented an unintended interpretation of the 'incident cancer' concept in the predefined analysis plan. We therefore used Outcomes A, B, and C in parallel, but added Outcome B *excluding the first 90 days of follow-up* as a less subjective alternative to Outcome C (presented in the tables only).

Cancer occurrence, rates

The event rates of the various cancer outcomes across the three anti-TNF agents were more similar than those in the comparator arms (Tables 2 and 3, supplementary Figure S1). This pattern remained when NMSCs (40% of all cancers) were excluded.

Table 3. Bayesian analysis results for all cancers including non-melanoma skin cancer, trials in the primary use conditions

	Anti-TNF		Control		Hazard ratio[||] (95%CI)	Bayesian probability that hazard ratio > 1
Drug vs. control	Events[¶]/ person-years	Rate per 100 000	Events[¶]/ person-years	Rate per 100 000		
Outcome A[*]						
All anti-TNF	103/8088	1273	34/3608	942	1.30 (0.89, 1.95)	0.91
Adalimumab	41/2859	1434	15/1465	1024	1.40 (0.78, 2.61)	0.87
Etanercept	38/3275	1160	14/1494	937	1.15 (0.60, 2.29)	0.65
Infliximab	24/1953	1228	5/650	769	1.56 (0.61, 4.67)	0.81
Outcome B[†]						
All anti-TNF	76/8082	940	27/3605	749	1.21 (0.77, 1.90)	0.80
Adalimumab	29/2858	1014	12/1465	819	1.21 (0.63, 2.48)–1.35 (0.74, 2.53)	0.71–0.84
Etanercept	31/3273	947	13/1493	871	1.00 (0.51, 2.10)–1.15 (0.63, 2.19)	0.48–0.67
Infliximab	16/1951	820	2/647	309	2.31 (0.90, 6.67)–2.64 (0.73, 12.78)	0.91–0.96
Outcome B[†] excluding first 90 days of follow-up						
All anti-TNF	56/5082	1101	18/2172	829	1.27 (0.75, 2.20)	0.81
Adalimumab	23/1748	1315	7/848	825	1.62 (0.72, 4.10)–1.70 (0.83, 3.73)	0.87–0.93
Etanercept	23/2226	1033	10/972	1029	0.87 (0.41, 1.94)–1.10 (0.55, 2.33)	0.37–0.60
Infliximab	10/1108	902	1/351	285	2.33 (0.82, 7.21)–3.50 (0.66, 34.43)	0.88–0.94
Outcome C[‡]						
All anti-TNF	44/8071	545	11/3598	306	1.75 (0.90, 3.63)	0.95
Adalimumab	17/2853	595	2/1460	137	2.90 (1.21, 7.71)–3.99 (1.18, 17.90)	0.99–0.99
Etanercept	18/3269	550	8/1491	537	0.87 (0.36, 2.19)–1.12 (0.52, 2.56)	0.37–0.61
Infliximab	9/1949	461	1/647	155	2.31 (0.79, 7.24)–3.18 (0.61, 22.87)	0.87–0.93

[*]Outcome A was defined as all cancer events (definite or probable cancers) diagnosed during the study period, using the date of diagnosis as the event date, irrespective of judgments on pre-trial prevalence.

[†]Outcome B was defined as all cancer events (definite or probable cancers) diagnosed during the study period, but excluding events in retrospect judged definitely prevalent on the basis of a first reported date of sign or symptom pre-dating the trial.

[‡]Outcome C was defined as all cancer events (definite or probable cancers) diagnosed during the study period, excluding both events with a first reported date of sign or symptom before the study period, *and* events which for other reasons were judged by the oncologists to be *probably* prevalent at trial start.

[¶]Numbers of events do not add up to total because of patients with more than one event (of different types).

[||]All-drug analysis uses non-informative prior. Drug-specific Hazard Ratios and the 95%CI reflect the range of Hazard Ratios found under three different Priors: (1) a mean HR of 2.0 and a 95% range of roughly [0.04, 109], (2) a mean HR of 0.0 and a 95% range of roughly [0.02, 55], and (3) a mean HR of 2.0 and a 95% range of roughly [0.50, 8.2].

Covariates in the all anti-TNF models were age, gender, concurrent CIT use, and disease condition, and the individual drug models included only age as a covariate.

DOI: 10.1002/pds

Cancer risks, hazard ratios, all sites

The Bayesian analyses by drug and outcome definition are shown in Table 3. These hazard ratios are largely similar to those obtained from ratios of the crude incidence rates (or incidence proportions) in the anti-TNF and comparator arms. Classwise analysis of all-site cancer risk across all three anti-TNF agents (Outcome A including NMSC) gave an overall hazard ratio of 1.30 (95%CI 0.89–1.95, Table 3) corresponding to a posterior probability of an increased cancer risk with anti-TNF vs. comparator of 0.91. This outcome showed little indication of any difference, irrespective of treatment assignment, across trials of the three drugs (using etanercept trials as reference: adalimumab trials HR = 0.96, infliximab trials HR = 0.93), nor across the *three anti-TNF treatments* when modeled separately (Table 3). Classwise analyses of outcome B including NMSC suggested an overall hazard ratio with anti-TNF of 1.21 (95%CI 0.77–1.90), but with indications of a difference in control group risk for infliximab (using etanercept trials as reference, adalimumab trials RR = 0.80, infliximab trials RR = 0.37), and some indication of an increased risk with infliximab treatment assignment (Table 3), though based on only two events in the infliximab comparator.

Cancer risks, hazard ratios, all sites minus NMSC

When NMSCs were excluded (Table 4), neither outcome A nor B suggested any increased overall risk with anti-TNF as a class (RR = 0.99 [95%CI 0.61–1.68] and RR = 0.90 [95%CI 0.51–1.56], respectively). Again, there were indications of a potential difference in control group cancer risk between the three sets of trials (using etanercept trials as reference, adalimumab trials RR = 0.79, infliximab trials RR = 0.17) and between *the three anti-TNF treatments* (Table 4) though numbers now dropped to single, or no, cases with cancer in the infliximab comparator. As a result of low numbers, the infliximab analysis of Outcome B did not converge using a non-informative prior (Table 4).

Table 4. Bayesian Analysis Results for all site cancers *excluding* non-melanoma skin cancers, trials in the primary use conditions

	Anti-TNF		Control		Hazard Ratio[||] (95%CI)	Bayesian probability that hazard ratio > 1
Drug vs. control	Events[¶]/ person-years	Rate per 100 000	Events[¶]/ person-years	Rate per 100 000		
Outcome A[*]						
All anti-TNF	52/8111	641	22/3614	609	0.99 (0.61, 1.68)	0.47
Adalimumab	18/2869	627	10/1468	681	0.92 (0.42, 2.14)–1.09 (0.56, 2.23)	0.40–0.59
Etanercept	23/3285	700	11/1497	735	0.88 (0.41, 1.84)–1.07 (0.56, 2.19)	0.36–0.57
Infliximab	11/1957	562	1/650	154	2.47 (0.88, 7.32)–3.71 (0.72, 27.69)	0.92–0.96
Outcome B[†]						
All anti-TNF	40/8107	493	19/3614	526	0.90 (0.51, 1.56)	0.35
Adalimumab	12/2869	418	8/1468	545	0.77 (0.31, 1.94)–1.02 (0.46, 2.33)	0.28–0.51
Etanercept	20/3283	609	11/1496	735	0.77 (0.37, 1.72)–0.96 (0.48, 1.93)	0.25–0.45
Infliximab	8/1955	409	0/649	0	3.02 (0.96, 9.99)–8.99 (0.92, 178.04)	0.96–0.97
Outcome B[†], excluding first 90 days of follow-up						
All anti-TNF	29/5103	568	13/2179	597	0.92 (0.48, 1.83)	0.40
Adalimumab	10/1757	569	4/851	470	1.23 (0.41, 4.28)–1.53 (0.63, 3.97)	0.63–0.81
Etanercept	15/2235	671	9/975	923	0.70 (0.29, 1.73)–0.96 (0.45, 2.14)	0.22–0.45
Infliximab	4/1111	360	0/353	0	2.51 (0.75, 9.34)–6.73 (0.50, 140.61)	0.89–0.93
Outcome C[‡]						
All anti-TNF	28/8104	346	9/3608	249	1.40 (0.66, 3.08)	0.81
Adalimumab	9/2867	314	1/1464	68	2.69 (0.91, 8.21)–4.46 (0.86, 38.74)	0.95–0.96
Etanercept	12/3281	366	8/1495	535	0.63 (0.24, 1.70)–0.90 (0.41, 2.03)	0.18–0.40
Infliximab	7/1956	358	0/649	0	2.89 (0.89, 10.18)–10.18 (1.01, 194.81)	0.95–0.98

[*]Outcome A was defined as all cancer events (definite or probable cancers) diagnosed during the study period, using the date of diagnosis as the event date, irrespective of judgments on pre-trial prevalence.

[†]Outcome B was defined as all cancer events (definite or probable cancers) diagnosed during the study period, but excluding events in retrospect judged definitely prevalent on the basis of a first reported date of sign or symptom pre-dating the trial.

[‡]Outcome C was defined as all cancer events (definite or probable cancers) diagnosed during the study period, excluding both events with a first reported date of sign or symptom before the study period, *and* events which for other reasons were judged by the oncologists to be *probably* prevalent at trial start.

[¶]Numbers of events do not add up to total because of patients with more than one event (of different types).

[||]All-drug analysis uses non-informative prior. Drug-specific Hazard Ratios and the 95%CI reflect the range of Hazard Ratios found under three different Priors: (1) a mean HR of 2.0 and a 95% range of roughly [0.04, 109], (2) a mean HR of 0.0 and a 95% range of roughly [0.02, 55], and (3) a mean HR of 2.0 and a 95% range of roughly [0.50, 8.2].

Covariates in the all anti-TNF models were age, gender, concurrent CIT use, and disease condition, and the individual drug models included only age as a covariate.

DOI: 10.1002/pds

Cancer risks, specific sites

When NMSC were assessed separately (Table 5) Outcomes A and B both showed an approximate doubling in risk for adalimumab and etanercept treatment assignment but less for infliximab treatment. Again, however, there were indications of heterogeneity in control group risk across the three sets of trials (using etanercept trials as reference, adalimumab trials RR = 1.46, infliximab trials RR = 3.37).

Because of low anticipated numbers, site-specific risks for other cancer types were not pre-planned. There were 12 lymphomas (9 vs. 3 for Outcome A, 7 vs. 3 for Outcome B, and 6 vs. 1 for Outcome C). The crude proportions would suggest a (classwise) relative risk largely similar to that of all-site cancers (data not shown). Supplementary Table S3 displays the crude counts of cancers (Outcome A) by treatment status across all indications.

Analyses using three alternative *informative* priors resulted in a range of similar hazard ratios and/or of convergence of models *that had not* converged using the initial non-informative prior probability (Tables 3–5). Use of alternative statistical methods (Peto meta-analysis) revealed largely similar results as the Bayesian approach for Outcomes A and B (data not shown). Crude hazard ratios by specific conditions treated were based on only single cases (supplementary Table S2). Analyses of each outcome excluding any study that used a dose of anti-TNF outside of the range indicated in the product label revealed a generally similar pattern of results to the primary analyses, except that the hazard ratio for cancers excluding NMSC for infliximab (defined per Outcome A) was attenuated (RR = 1.43 [95%CI 0.11–33.72]).

DISCUSSION

Overall, the occurrence of cancer diagnoses (other than NMSC) in these trials was not higher with anti-TNF treatment than with comparator treatment. By contrast, for NMSC, the overall analysis suggested a statistically significant doubling in risk with anti-TNF therapy.

Table 5. Bayesian analysis results for non-melanoma skin cancer only, trials in primary use conditions

	Anti-TNF		Control		Hazard ratio[||] (95%CI)	Bayesian probability that hazard ratio > 1
Drug vs. Control	Events[¶]/ person-years	Rate per 100 000	Events[¶]/ person-years	Rate per 100 000		
Outcome A[*]						
All anti-TNF	53/8094	655	12/3613	332	2.02 (1.11, 3.95)	0.99
Adalimumab	23/2863	803	5/1468	341	2.59 (1.01, 7.67)	0.98
Etanercept	16/3278	488	3/1495	201	2.40 (0.71, 9.97)	0.91
Infliximab	14/1954	716	4/650	615	1.19 (0.39, 4.61)	0.59
Outcome B[†]						
All anti-TNF	38/8092	470	8/3611	222	2.18 (1.03, 4.92)	0.98
Adalimumab	17/2862	594	4/1468	272	2.11 (0.79, 6.13)–2.29 (0.81, 8.29)	0.93–0.97
Etanercept	12/3277	366	2/1495	134	2.14 (0.82, 5.86)–2.84 (0.73, 15.07)	0.91–0.94
Infliximab	9/1953	461	2/648	309	1.50 (0.39, 8.07)–1.78 (0.64, 5.44)	0.68–0.86
Outcome B[†], excluding first 90 days of follow-up						
All anti-TNF	28/5090	550	5/2176	230	2.34 (0.96, 7.38)	0.97
Adalimumab	13/1751	742	3/851	352	2.12 (0.85, 5.70)–2.27 (0.72, 8.38)	0.89–0.94
Etanercept	9/2230	403	1/974	103	2.37 (0.82, 7.18)–3.73 (0.61, 350.02)	0.89–0.94
Infliximab	6/1109	541	1/351	285	1.88 (0.33, 17.17)–2.21 (0.37, 17.31)	0.71–0.87
Outcome C[‡]						
All anti-TNF	18/8086	223	2/3609	55	4.96 (1.21, 41.06)	0.99
Adalimumab	8/2859	280	1/1467	68	2.59 (0.91, 7.85)–4.74 (0.77, 50.40)	0.94–0.96
Etanercept	7/3276	214	0/1494	0	3.16 (0.99, 10.53)–9.74 (1.00, 212.30)	0.97–0.98
Infliximab	3/1951	154	1/648	154	1.19 (0.17, 12.74)–1.72 (0.50, 6.11)	0.54–0.81

[*]Outcome A was defined as all cancer events (definite or probable cancers) diagnosed during the study period, using the date of diagnosis as the event date, irrespective of judgments on pre-trial prevalence.

[†]Outcome B was defined as all cancer events (definite or probable cancers) diagnosed during the study period, but excluding events in retrospect judged definitely prevalent on the basis of a first reported date of sign or symptom pre-dating the trial.

[‡]Outcome C was defined as all cancer events (definite or probable cancers) diagnosed during the study period, excluding both events with a first reported date of sign or symptom before the study period, *and* events which for other reasons were judged by the oncologists to be *probably* prevalent at trial start.

[¶]Numbers of events do not add up to total because of patients with more than one event (of different types).

[||]All-drug analysis uses non-informative prior. Drug-specific Hazard Ratios and the 95%CI reflect the range of Hazard Ratios found under three different Priors: (1) a mean HR of 2.0 and a 95% range of roughly [0.04, 109], (2) a mean HR of 0.0 and a 95% range of roughly [0.02, 55], and (3) a mean HR of 2.0 and a 95% range of roughly [0.50, 8.2].

Covariates in the all anti-TNF models were age, gender, concurrent CIT use, and disease condition, and the individual drug models included only age as a covariate.

DOI: 10.1002/pds

The overall relative risk (Outcome A, all sites including NMSC, relative risk = 1.30, 95%CI 0.89–1.95) of our meta-analysis, which included at least 50% more RA trials than either of the previous analyses, *plus* 43 trials in other conditions, and a robust individual patient-data design and blinded adjudication of events, is in the same range as that presented in the global pooled data safety analyses by Leombruno *et al.* of 18 trials of any of the three anti-TNF agents in RA, based on data in the reported publications.[5] In that study, the overall relative risk for cancer (excluding NMSC) in patients receiving recommended doses of the anti-TNF was 1.31 (95%CI 0.69, 2.48). Similar to our study, the previous pooled analysis of malignancies in nine RA RCTs of adalimumab and infliximab by Bongartz *et al.* (reported relative risk = 3.3), reported very low rates of cancer in the comparator groups (based on three cancers). However, in contrast to our meta-analysis that was based on events and number of person-years of follow-up until 30 days post planned end of treatment, their pooled analysis was based on number of patients randomized and events throughout the trials. A subsequent meta-analysis of nine etanercept trials of RA by the same authors found a hazard ratio of 1.8 derived from 26 cancers in anti-TNF arms and seven cancers in comparator arms.[4] The Bongartz[3] and the Leombruno[5] analyses both found increased risks with higher doses of anti-TNF therapy. The data provided in our study did, however, not allow for detailed assessment of risks by dose, although analyses restricted to the labeled dose yielded largely similar results. A combined efficacy/safety meta-analysis on biologics (other agents than adalimumab, etanercept, and infliximab were included) in Crohn's disease trials found no significant increase in cancer risk with anti-TNF therapy, though again, numbers were small.[14]

Observational studies[7–9,15] have not, so far, suggested any marked increase in the risk of all-site cancer following anti-TNF therapy. These studies have typically assessed hazard ratios for the class of anti-TNF agents rather than for individual drugs, for RA only, and not by time of follow-up. Hence, short-term risks have, paradoxically, not been specifically addressed. In the one study that assessed drug-specific risks by time since treatment start, dissimilar cancer risks were observed during the first year, but not thereafter.[10]

Several studies of immunocompromised patients, including patients with RA, have reported increased risks of NMSC. Observational studies of anti-TNF-agents in RA have further reported increased NMSC risks with anti-TNF therapy.[7,9,16] In our analyses, there were consistent indications of increased risks for NMSC for both etanercept and adalimumab using Outcomes A and B, but not for infliximab.

Although lymphomas were of particular interest, there were too few cases to permit statistical modeling of hazard ratios. The lymphoma rate among the comparators in the RA trials was 83 per 100 000 person-years, which is more than double the US general population rate[17] but only slightly lower than the rate in anti-TNF arms of the same RA trials (111 cases per 100 000 person-years). The incidence rate in the Swedish Biologics Register of RA was in the same order of magnitude (92 per 100 000 in biologics-naive RA, and 96 per 100 000 person-years following anti-TNF therapy).[6]

For each of the outcome definitions, and irrespective of whether NMSC was included, our results indicated a pattern of (i) relatively similar cancer incidences in the anti-TNF arms across the three agents, (ii) relatively similar cancer incidences in the anti-TNF arms and the comparator arms for etanercept, but (iii) lower cancer incidences in the comparator arms of infliximab trials. The adalimumab results fell between those of etanercept and infliximab. In assessing the results of drug-specific analyses, and before concluding the absence of a signal for non-cutaenous cancers for etanercept, and the presence of an increased risk for non-cutaenous cancers with infliximab, a series of issues need be discussed: The analyses that indicated differences in drug-specific hazard ratios typically also indicated differences in baseline cancer risk across the three sets of sponsors' trials, irrespective of treatment assignment (anti-TNF or comparator). These two types of differences were correlated, as more variable baseline differences in cancer rate occurred with more variable hazard ratios. A similar correlation was true for statistical precision, with analyses suggesting the largest difference in hazard ratio across anti-TNF agents coming from studies with the least precision due to too few events. It is thus difficult to determine whether apparent differences in cancer risk among the three drugs were related to (i) differences related to the three anti-TNF agents; (ii) differences in the comparator arms; (iii) study-specific differences in reporting of cancer adverse events; or (iv) chance.

Regarding drug-specific differences, all anti-TNF agents are antagonists of TNF-alpha, but whereas adalimumab and infliximab are monoclonal antibodies, etanercept is a receptor fusion protein. Because of their different biological properties, a different effect on cancer emergence cannot be excluded. Other differences (half-life, fully humanized vs. not, infusion vs. self injection) may also be of relevance.

DOI: 10.1002/pds

Regarding comparator differences, it is unclear whether the comparator rates in etanercept (735 per 100 000) or adalimumab (681/100 000) RCTs for Outcome A (NMSC excluded) were unusually high, or the infliximab comparator rate (154/100 000) unusually low. As benchmark, the incidence of all cancers excluding NMSC in the general US population is 740 cases per 100 000 person-years,[17] and 900–1300 per 100 000 in observational studies of cancer in patients with RA.[7,8,18,19] Although randomization should make the rates internally comparable within each RCT, any effect modification of risks with anti-TNF by baseline rate of cancer would limit comparability across different sets of trials.

Regarding study-related differences, the frequency of events requiring adjudication differed across the three sets of trials, as did the amount of information available to the adjudicators. Although the adjudication was performed blinded to trial/treatment/control status, the adjudication of many events was based upon minimal information (as in the etanercept studies). If the level of detail provided in the narratives varied systematically by treatment assignment within trials of any single anti-TNF agent, this could have affected the numbers of events and hazard ratios across the three drugs.

Regarding limited statistical precision, the small number of events in some groups means that small changes in absolute event numbers may translate into large changes in hazard ratio. For instance, using the crude data from Tables 3 and 4, one or two additional cases in the infliximab comparator would essentially remove the suggestion of increased risks with infliximab treatment.

Although the randomized design of the RCTs made within-study groups comparable at the start of each trial, it does not guarantee ongoing comparability for the duration of each study. The very efficacy of anti-TNF therapy may alter the probability of having a cancer detected, or the timing of its diagnosis. Given the relatively short timeframes of these RCTs, even small changes in the time taken to diagnose cancer (lead time bias) may have considerable effects on the hazard ratio estimates. Screening of potential subjects before, and surveillance during, clinical trials may have eliminated patients at highest risk of developing cancer in the short-term, or, triggered investigations ultimately leading to a cancer diagnosis during the trial.

To conclude, with respect to short-term cancer risks with anti-TNF therapies, the overall occurrence of cancers other than NMSC was not markedly higher in individuals randomized to anti-TNF therapy than to the comparator arms. We noted, however, indications of drug-specific increases in risk, the nature and cause of which could not be fully established. Different levels of statistical precision and potential differences in study conduct and reporting practices across the three sets of trials precluded this comprehensive and exhaustive meta-analysis from a conclusive determination of whether individual anti-TNF therapies affect the short-term risks of being diagnosed with cancer. For NMSC, there was a more consistent pattern across trials suggesting an increase with anti-TNF inhibitors, which together with previous findings also suggesting an increased risk with anti-TNF therapies, would call for increased clinical vigilance for this cancer type.

From a methodological point of view, our study highlights the challenges in re-analyzing clinical trials for outcomes not originally planned, and comparing data across sponsors (but, conversely, the potential for pre-planned safety-meta analyses of emerging trial data, using pre-defined safety endpoints). In particular, it shows the substantive complexity behind seemingly straightforward events like 'cancer', which are uncommon, insidious, not identified as pre-specified outcomes, and for which no standardized reporting or definition of 'onset' exists. Although this individual patient data meta-analysis represents the best available evidence to date regarding short-term risks, analyses of long-term risks as well as further explorations of the risk for site-specific cancers, and in the different populations for which anti-TNF agents are used, require larger study populations and longer follow-up. The ongoing long-term follow-up via registries is important in this respect.

CONFLICT OF INTEREST

JA has performed scientific analyses using data from the Swedish Biologics Register that is run by the

KEY POINTS

- Overall, no increased short-term risk for cancer other than non-melanoma skin cancer (NMSC) was observed with anti-TNF therapy as a class, although indications of increased risks with individual anti-TNF drugs could neither be verified nor refuted.
- Anti-TNF therapies may increase the short-term risk of NMSC diagnosis.
- Meta-analysis of uncommon safety outcomes based on data from randomized controlled trials is a powerful tool but poses a series of methodological challenges that require due attention and action.

DOI: 10.1002/pds

Swedish Society for Rheumatology. For the maintenance of this register, the Swedish Society for Rheumatology has received funding, independent of the conduct of these scientific analyses, from Schering-Plough, BMS, Wyeth, Abbott Laboratories, and Roche. JA reports financial disclosures of <2000 USD in speaker's honorarium for presentations of results from safety studies based on data from the Swedish Biologics Register at scientific meetings which in turn had received funding or were organized by sponsors of this meta-analysis.

CS has consulted for Pfizer. KF and BN are employed by UBC, which has provided various consulting services to all of the sponsors of this project. SR has been both an employee of and a consultant to MetaWorks, later acquired by UBC, which has provided various consulting services to all of the sponsors of this project. DS has performed scientific analyses using data from the British Society for Rheumatology Biologics Register that is funded by the British Society for Rheumatology. The BSR receive funding independent of the conduct of the scientific analyses from Abbott Laboratories, Amgen, Biovitrum, Schering-Plough, Wyeth Pharmaceuticals, and Roche.

ACKNOWLEDGEMENTS

This study was supported by a contract between three pharmaceutical company sponsors (Abbott Laboratories, Centocor Inc., and Wyeth Pharmaceuticals) and United BioSource Corporation (UBC). JA, CS, and DS received compensation from UBC for their work as consultants and voting members of the project Steering Committee (SC). SR was initially a UBC employee and subsequently a consultant, serving as a non-voting member of the SC, and KF, and BN are employees of UBC. Drs Kevin Knopf and Babis Andreadis served as the cancer adjudication committee.

The corporate sponsors were allowed to comment on the analysis plan and the paper, but all final decisions on the analysis and paper were the responsibility of JA, CS, and DS. CS served as the independent statistician, had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article:

Supplementary Figure 1. Cumulative incidence of outcome A, including non-melanoma skin cancer as a function of time since start of the study period, by drug and treatment status.

Supplementary Figure 2. Cumulative incidence of outcome A, excluding non-melanoma skin cancer as a function of time since start of the study period, by drug and treatment status.

Supplementary Table 1. Number of Cancer Events (Outcome A, including Non-Melanoma Skin Cancer) number of person-years of follow-up, by condition, by drug, and by treatment and control status.

Supplementary Table 2. Study characteristics for all included trials.

Supplementary Table 3. Cancer events (outcome A) during follow-up in all included trials.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

REFERENCES

1. Balkwill F. TNF-alpha in promotion and progression of cancer. *Cancer Metastasis Rev* 2006; **25**(3): 409–416.
2. van Horssen R, Ten Hagen TL, Eggermont AM. TNF-alpha in cancer treatment: molecular insights, antitumor effects, and clinical utility. *Oncologist* 2006; **11**(4): 397–408.
3. Bongartz T, Sutton AJ, Sweeting MJ, Buchan I, Matteson EL, Montori V. Anti-TNF antibody therapy in rheumatoid arthritis and the risk of serious infections and malignancies: systematic review and meta-analysis of rare harmful effects in randomized controlled trials. *Jama* 2006; **295**(19): 2275–2285.
4. Bongartz T, Warren FC, Mines D, Matteson EL, Abrams KR, Sutton AJ. Etanercept therapy in rheumatoid arthritis and the risk of malignancies. A systematic review and individual patient data meta-analysis of Randomized Controlled Trials. *Ann Rheum Dis* 2009; **68**(7): 1177–1183.
5. Leombruno JP, Einarson TR, Keystone EC. The safety of anti-Tumor Necrosis Factor treatments in rheumatoid arthritis: meta and exposure adjusted pooled analyses of serious adverse events. *Ann Rheum Dis* 2009; **68**(7): 1136–1145.
6. Askling J, Baecklund E, Granath F, *et al*. Anti-tumour necrosis factor therapy in rheumatoid arthritis and risk of malignant lymphomas: relative risks and time trends in the Swedish Biologics Register. *Ann Rheum Dis* 2009; **68**(5): 648–653.
7. Askling J, Fored CM, Brandt L, *et al*. Risks of solid cancers in patients with rheumatoid arthritis and after treatment with tumour necrosis factor antagonists. *Ann Rheum Dis* 2005; **64**: 1421–1426.
8. Setoguchi S, Solomon DH, Weinblatt ME, *et al*. Tumor necrosis factor alpha antagonist use and cancer in patients with rheumatoid arthritis. *Arthritis Rheum* 2006; **54**(9): 2757–2764.
9. Wolfe F, Michaud K. Biologic treatment of rheumatoid arthritis and the risk of malignancy: analyses from a large US observational study. *Arthritis Rheum* 2007; **56**(9): 2886–2895.
10. Askling J, van Vollenhoven RF, Granath F, *et al*. Cancer risk in patients with rheumatoid arthritis treated with anti-tumor necrosis factor alpha therapies: does the risk change with the time since start of treatment? *Arthritis Rheum* 2009; **60**(11): 3180–3189.
11. Available from: http://www.fda.gov/Drugs/DrugSafety/PostmarketDrugSafetyInformationforPatientsandProviders/DrugSafetyInformationforHeathcareProfessionals/ucm174474.htm
12. Wegener's Granulomatosis Etanercept Trial (WGET) Research Group. Etanercept plus standard therapy for Wegener's granulomatosis. *N Engl J Med* 2005; **352**(4): 351–361.
13. Berry S, Berry D, Natarajan K, *et al*. Bayesian survival analysis with nonproportional hazards: metaanalysis of combination pravastatin-aspirin. *JASA* 2004; **99**: 36–44.
14. Peyrin-Biroulet L, Deltenre P, de Suray N, *et al*. Efficacy and safety of tumor necrosis factor antagonists in Crohn's disease: meta-analysis of placebo-controlled trials. *Clin Gastroenterol Hepatol* 2008; **6**(6): 644–653.
15. Geborek P, Bladström A, Turesson C, *et al*. Tumour necrosis factor blockers do not increase overall tumour risk in patients with rheumatoid arthritis, but may be

associated with an increased risk of lymphomas. *Ann Rheum Dis* 2005; **64**(5): 699–703.
16. Chakravarty EF, Michaud K, Wolfe F. Skin cancer, rheumatoid arthritis, and tumor necrosis factor inhibitors. *J Rheumatol* 2005; **32**(11): 2130–2135.
17. Age-adjusted SEER incidence and U.S. death rates and 5-year relative survival rates by primary cancer site, sex and time period. . Available from: http://seer.cancer.gov/csr/1975_2005/results_merged/topic_survival.pdf
18. Askling J, Fored CM, Baecklund E, *et al*. Haematopoietic malignancies in rheumatoid arthritis: lymphoma risk and characteristics after exposure to tumour necrosis factor antagonists. *Ann Rheum Dis* 2005; **64**(10): 1414–1420.
19. Mellemkjaer L, Linet MS, Gridley G, *et al*. Rheumatoid arthritis and cancer risk. *Eur J Cancer* 1996; **32A**(10): 1753–1757.

METHODS APPENDIX

Definitions used for adjudication of cancer events

Definite cancer events were those with a histology listed in the narrative, implying a pathologic examination of a specimen. Probable cancers were those with no histological diagnosis, but with other strong evidence (e.g., multiple liver or lung metastases on X-ray or scan, where the likelihood of cancer was judged to be between 50 and 90%). Possible events included those with a clinical diagnosis of a malignancy but no supporting test findings, and diagnoses that were not clearly malignant but that had supplementary information suggesting that the event was a cause for concern and/or not obviously benign. Unlikely events were those with less than 10% chance of being cancer.

Definitely prevalent events had documented earliest evidence (sign or symptom) of cancer appearing before the start of treatment. Probably prevalent events (chance of existing at trial start 50–90%) had no documentation of any sign or symptom related to the cancer before trial start, but the adjudicators still believed that clinical signs would have appeared before trial start. A possibly or unlikely prevalent event had less than 50% chance of having started before the trial.

A recurrent cancer indicated a recurrence during the trial of a cancer, with the same type and at the same or closely adjacent location to a cancer that was diagnosed prior to trial entry. No recurrent cancers were found in the primary disease conditions; all events reported in the primary analysis were thus incident events.

Interpretation of Bayesian risk estimates

The Bayesian approach uses the language of probability to describe uncertainty. Bayesian results are, therefore, appropriately interpreted in terms of probability (not likelihood or confidence). Bayesian 95% credible intervals, for example, describe a range within which the true model parameter lies with 95% probability. Sometimes these are called Bayesian confidence intervals. When non-informative priors are used, the Bayesian credible interval will be similar to a frequentist ('common') confidence interval. The posterior distribution provides a probability with which a model parameter (such as a treatment effect expressed as a hazard) exceeds a specified numerical value such as 1.0. Such probabilities are one-tailed by definition as they derive directly from a probability distribution.

Bayesian computations

The Bayesian models were estimated by Markov chain Monte Carlo employing the Gelman–Rubin convergence diagnostic and evaluating model fit with the deviance information criterion.[14] We used non-informative prior distributions except with sparse data when treatment effects were given weakly informative priors. Generally, inferences varied only slightly.

Bayesian analyses based on running three Markov chain Monte Carlo (MCMC) simulations at dispersed starting values were used to model event risk. Convergence of the algorithm was assessed by looking at time series plots of the chains, their within-chain autocorrelation and the Gelman–Rubin convergence diagnostic using a convergence limit of between to within-chain variance set to 1.1. Chains were run long enough to achieve convergence and also to yield at least 40 000 iterations for conducting inferences. The relative performance of the Bayesian survival models was evaluated using the deviance information criteria (DIC). The proportional hazard assumption was evaluated using log–log survival plots.

Selection of prior probabilities

Prior distributions on model parameters were initially chosen to be vague, including normal distributions with mean 0 and SD of 1000 for the log-treatment hazard and all other predictors (e.g., age), and a uniform distribution over the interval [0, 100] for the standard deviation of the log baseline hazard. Because several control arms contained fewer than three events, the data did not provide sufficient information to accurately estimate their posterior distributions with no prior information. Therefore, we set up three weakly informative priors for the treatment effects in these cases. All three assumed normal distributions and none implied any substantive effect.

Prior 1 assumed a non-zero treatment effect, with a mean log(hazard) of 0.7 and a standard deviation of 2. This roughly translates to an estimated mean hazard ratio of 2 with a 95% chance for a hazard ratio between 0.04 and 110.

Prior 2 assumed a zero treatment effect, with a mean log(hazard) of 0 and a standard deviation of 2. This roughly translates to an estimated mean hazard ratio of 1 with a 95% chance for a hazard ratio between 0.02 and 55.

Prior 3 assumed a non-zero treatment effect that was more informative than that of Prior 1, with a mean log(hazard) of 0.7 and a standard deviation of 0.7. This roughly translates to an estimated mean hazard ratio of 2 with a 95% chance for a hazard ratio between 0.5 and 8.2.

All of these prior distributions therefore imposed weak restrictions on the size of the treatment effect. Because this is a safety study, priors 1 and 3 conservatively assumed a prior increased risk in treatment compared with control.

DOI: 10.1002/pds