Increased risk of lymphoma among inflammatory bowel disease patients treated with azathioprine and 6-mercaptopurine

A Kandieli, A G Fraser, B I Korelitz, C Brensinger, J D Lewis

**Background:** Inflammatory bowel disease (IBD) is commonly treated with immunomodulators such as azathioprine and 6-mercaptopurine (6-MP). Studies examining lymphoma risk in IBD patients treated with these medications have been underpowered and have yielded conflicting conclusions.

**Aims:** The purpose of this meta-analysis was to provide a more precise estimate of the relative risk of lymphoma among IBD patients treated with azathioprine or 6-MP.

**Methods:** Studies were included if they were English language, full article, cohort studies specifically designed to evaluate cancer as an adverse outcome of treatment with azathioprine or 6-MP. Pooled standardised incidence ratios were calculated to estimate the relative risk of lymphoma associated with therapy. Heterogeneity was assessed using Poisson regression. Sensitivity analyses examined the influence of individual studies on risk estimate and heterogeneity statistics.

**Results:** Six studies were identified that met our inclusion criteria. When the data were combined across all studies, the pooled relative risk was 4.18 (95% confidence interval 2.07–7.51; 11 observed cases, 2.63 expected). Sensitivity analyses showed that exclusion of any one study had a relatively small effect on the pooled relative risk estimate (range 3.49–5.21) but excluding either the study with the highest or lowest estimated relative risk eliminated the statistically significant heterogeneity.

**Conclusions:** Our data suggest an approximate fourfold increased risk of lymphoma in IBD patients treated with azathioprine/6-MP. The increased risk of lymphoma could be a result of the medications, the severity of the underlying disease, or a combination of the two.
6-MP, mean/median dose, mean median duration of follow up, expected number of patients with lymphoma, observed number of patients with lymphoma, and type of lymphoma.

**Statistical analyses**

In order to calculate a pooled estimate of the relative risk of lymphoma, it was necessary to have all studies analysed in the same manner. All but two studies\(^{16,17}\) calculated standardised incidence ratios (SIRs) using indirect standardisation.\(^{18}\) In this method observed numbers of cancers in each study are compared with the expected number of cases of lymphoma based on age and sex specific rates for the general population that gave rise to the study cohort; 95% confidence intervals (CI) are calculated assuming that the incidence of lymphoma follows a Poisson distribution. If the 95% CI excludes 1.0, this is consistent with a statistically significant increase or decrease in the risk of lymphoma with a type I error of less than 0.05.

For studies by Fraser and colleagues\(^{16}\) and Korelitz and colleagues,\(^{17}\) we reanalysed the primary data to calculate the SIRs. For the study by Fraser and colleagues,\(^{16}\) we used the primary data from the original study and compared these data with lymphoma rates for England and Wales.\(^{19}\) Thus the SIR was calculated as the ratio of observed lymphomas divided by the expected number of lymphomas on the age and sex distribution of the study population. Confidence intervals were estimated assuming that the incidence of lymphoma follows a Poisson distribution.

For the study by Korelitz and colleagues,\(^{17}\) the data set that was used in the original publication was no longer available. As a result, we used more up to date data obtained from the SEER cancer registry.\(^{20}\) Confidence intervals were again estimated assuming that the incidence of lymphoma follows a Poisson distribution.

The study by Farrell and colleagues\(^{15}\) included two cases of lymphoma in patients treated exclusively with methotrexate and/or ciclosporin. For this study, we included only those lymphoma patients treated with azathioprine or 6-MP, and excluded patients treated exclusively with methotrexate and/or ciclosporin. We reduced the expected number of lymphomas in proportion to the fraction of the total cohort of patients treated with immunomodulators other than azathioprine or 6-MP (expected 0.06 in original article, reduced to 0.053 in our analysis).

In the study by Körleitz et al.,\(^{17}\) we reanalysed the primary data to calculate the SIRs for the Lewis, Fraser, and Korelitz studies on the pooled SIR estimate, we performed a series of sensitivity analyses in which the pooled SIR and the deviance statistic were recalculated after excluding each study.

Secondary analyses were performed examining the relative risk of non-Hodgkin lymphoma. For these analyses, we recalculated the SIRs for the Lewis, Fraser, and Korelitz studies using only the expected rates of non-Hodgkin lymphoma. Patients who developed Hodgkin's disease were censored at the time of lymphoma diagnosis.

![Figure 1](http://gut.bmj.com/10.1136/gut.2004.049460)
Three studies provided sufficient data to compare lymphoma risk in azathioprine or 6-MP treated IBD patients to that in IBD patients not treated with these medications.15 16 The results of these studies were pooled using Mantel-Haenszel methods with weights proportional to the inverse variance, as implemented in the STATA “ir” routine (STATA version 8.2, College Station, Texas, USA). Two of the studies15 16 allowed for calculation of age adjusted relative risk estimates. Results of these studies were pooled using the random effects summary estimate of the data, as implemented in the STATA “meta” routine (STATA version 8.2).

**Ethics approval**

This study was approved by the Institutional Review Board of the University of Pennsylvania.

**RESULTS**

Six studies met our inclusion criteria (table 1).

### Relative risk of lymphoma

When the data were combined across all of the included studies, the total number of observed cases of lymphoma was 11, and the total number of expected cases was 2.63, resulting in an SIR of 4.18 (95% CI 2.07–7.51). However, as can be seen in fig 1, there was significant heterogeneity among the studies (deviance statistic p = 0.03).

Sensitivity analyses demonstrated that the significant heterogeneity in pooled analyses was explained by the extreme difference in results of the studies of Connell and colleagues14 and Farrell and colleagues.15 When either of these studies was excluded from the analysis, the test for heterogeneity was no longer statistically significant (table 2). Importantly, in these analyses, exclusion of any one study had a relatively small effect on the pooled relative risk estimate (range 3.49–5.21). Furthermore, excluding any one study did not eliminate the statistical significance of the pooled estimate of the relative risk of lymphoma.

Because the study of Kielen12 included patients without IBD, we performed additional sensitivity analyses excluding this study and the two previously identified outliers. When both the Kinlen study12 and the Connell and colleagues study14 were excluded from the analysis, the SIR was 4.61 (95% CI 2.09–8.79); when the Kinlen study12 and the Farrell and colleagues study15 were excluded from the analysis, the SIR was 2.90 (95% CI 1.15–6.00).

Three studies provided sufficient data to directly compare azathioprine or 6-MP treated IBD patients with IBD patients who had not received this therapy.15 16 The pooled analysis yielded a combined relative risk of 2.92 (95% CI 1.05–8.13, test for heterogeneity p = 0.18). If we excluded the Farrell study,16 the combined relative risk estimate was 2.03 (95% CI 0.66–6.29). Using age adjusted estimates of the relative risk from the studies of Fraser and colleagues16 and Lewis and colleagues,1 the pooled analysis yielded a slightly higher relative risk estimate of 3.11 (95% CI 0.66–14.62).

To help place the significance of our findings into a clinical context, we estimated the number of person years of follow up after therapy with azathioprine or 6-MP needed to result in one additional case of lymphoma. Based on the results from our meta-analysis, we performed this calculation for a range of relative risk estimates (low of 2.0 to high of 6.0) (table 3). Based on these calculations, assuming a relative risk of lymphoma of 4.0, the number of patients needed to be treated to cause one additional lymphoma per year ranged from approximately 4357 persons aged 20–29 years to 355 persons aged 70–79 years.

In secondary analyses, we focused specifically on the relative risk of non-Hodgkin lymphoma. There was significant heterogeneity (p = 0.01) among the studies when we examined only the risk of non-Hodgkin lymphoma, with two studies estimating very high relative risks of lymphoma (Farrell and colleagues15 SIR 37.4 and Kielen12 SIR 12.5) while two studies (Lewis and colleagues1 and Connell and colleagues14) observed no cases of non-Hodgkin lymphoma, thus resulting in SIR estimates of zero. Combining all studies, there were nine cases of non-Hodgkin lymphoma compared with an expected 2.30 (SIR 3.92; 95% CI 1.78–7.47). The increased relative risk of lymphoma and the test for heterogeneity remained statistically significant when any one study was excluded. The lowest estimate was observed when the study by Farrell and colleagues15 was excluded (SIR 3.12; 95% CI 1.24–6.46). Excluding the study by Lewis and colleagues1 resulted in the highest SIR (SIR 5.07; 95% CI 2.30–9.66). When we dropped the studies by both Kielen12 and Farrell and colleagues,15 the resulting estimated SIR was 2.40 (95% CI 0.76–5.64). Because of the heterogeneity, these secondary pooled analyses need to be viewed with caution.

### Case descriptions

Overall, there were 11 cases of lymphoma, of which two were Hodgkin’s (table 4). Patients had received a median of 14 months of therapy prior to the diagnosis of lymphoma (range 6–94 months). Of the nine non-Hodgkin lymphomas, four originated in the bowel, one originated in the central nervous system, two originated elsewhere, and the site was not reported for the remaining two lymphomas.

---

**Table 2** Sensitivity analysis of lymphoma risk when individual studies were excluded from the analysis.

<table>
<thead>
<tr>
<th>Study</th>
<th>SIR (95% CI)</th>
<th>p value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Connell</td>
<td>5.21 (2.59–9.35)</td>
<td>0.09</td>
</tr>
<tr>
<td>Kielen</td>
<td>3.64 (1.65–6.94)</td>
<td>0.03</td>
</tr>
<tr>
<td>Farrell</td>
<td>3.49 (1.58–6.66)</td>
<td>0.12</td>
</tr>
<tr>
<td>Lewis</td>
<td>5.02 (2.39–9.27)</td>
<td>0.03</td>
</tr>
<tr>
<td>Fraser</td>
<td>4.03 (1.72–7.98)</td>
<td>0.01</td>
</tr>
<tr>
<td>Korelitz</td>
<td>3.91 (1.69–7.84)</td>
<td>0.01</td>
</tr>
</tbody>
</table>

SIR, standardised incidence ratio (expected/observed).

*p < 0.05 implies there is significant heterogeneity among the studies included in the analysis.

---

**Table 3** Number needed to treat to cause one additional lymphoma per year.

<table>
<thead>
<tr>
<th>Age (y)</th>
<th>Lymphoma incidence*</th>
<th>NNH if relative risk of lymphoma = 2</th>
<th>NNH if relative risk of lymphoma = 4</th>
<th>NNH if relative risk of lymphoma = 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>20–29</td>
<td>7.65</td>
<td>3107</td>
<td>4357</td>
<td>2614</td>
</tr>
<tr>
<td>30–39</td>
<td>10.70</td>
<td>9346</td>
<td>3115</td>
<td>1869</td>
</tr>
<tr>
<td>40–49</td>
<td>16.60</td>
<td>6024</td>
<td>2008</td>
<td>1205</td>
</tr>
<tr>
<td>50–59</td>
<td>29.60</td>
<td>3378</td>
<td>1126</td>
<td>676</td>
</tr>
<tr>
<td>60–69</td>
<td>56.45</td>
<td>1771</td>
<td>591</td>
<td>354</td>
</tr>
<tr>
<td>70–79</td>
<td>93.90</td>
<td>1065</td>
<td>355</td>
<td>213</td>
</tr>
</tbody>
</table>

*Incidence rates per 100 000 person years from 1996–2000 (SEER data17).

NNH, number needed to harm one additional patient per year.

---

Accessed as 10.1136/gut.2004.049460 on 11 July 2005. Downloaded from http://gut.bmj.com on June 8, 2022 by guest. Protected by copyright.
DISCUSSION
Most previous research suggests that the risk of lymphoma among patients with IBD is similar to that in the general population. In contrast, this meta-analysis demonstrates that there is an approximate fourfold increased risk of lymphoma in the subgroup of IBD patients treated with azathioprine and/or 6-MP relative to the lymphoma rate expected in the general population. These findings are consistent with prior research demonstrating a modest increased risk of lymphoma among patients receiving these medications for rheumatoid arthritis. The increased risk of lymphoma among IBD patients is markedly lower than that observed after organ transplantation, a condition in which much greater levels of immunosuppression are achieved.

Another way to place these results into a clinical context is to ask how great of a risk of lymphoma is necessary for the risks of therapy to outweigh the benefits. Lewis et al previously conducted a decision analysis using a Markov model to assess the efficacy of alternative management strategies for maintaining remission in patients with Crohn’s disease. That study showed that azathioprine must result in a greater than 9.8-fold increased risk of lymphoma for therapy with alternative medications to be the preferred treatment strategy, assuming that Crohn’s disease does not itself lead to a baseline increased risk of lymphoma. Using this model, we can then assume that a fourfold increased risk of lymphoma does not preclude the use of azathioprine and/or 6-MP in the treatment of IBD, especially in young patients. An important objective of meta-analysis is to look for evidence of heterogeneity among studies and to determine whether differences in study design explain the heterogeneity. We observed significant heterogeneity among studies in both our primary and secondary analyses. The greatest difference was between the results of studies by Connell and colleagues and Farrell and colleagues. Both studies were set in single centres and had similar durations of therapy, durations of follow up, and dose. The most obvious difference was sample size. The study by Farrell and colleagues was the smallest and had the lowest expected number of lymphomas. Because the expected number of lymphomas was approximately 0.05, every case of lymphoma increased the relative risk estimate by approximately 20. Because of the small sample size, the study by Farrell and colleagues produced a less precise estimate of the relative risk than the other studies. Thus while we agree with Farrell’s conclusions that IBD patients treated with azathioprine/6-MP appear to be at increased risk of lymphoma, the magnitude of that increased risk appears to be far lower than that estimated by Farrell and colleagues.

As with all meta-analyses, we designed specific inclusion and exclusion criteria prior to initiating our study. Our decision to focus exclusively on studies that included cancer as an outcome of interest was implemented to avoid misclassification bias. Studies designed to look at other outcomes may examine cancer as a secondary outcome, or more commonly as an adverse event. Data collection for these secondary outcomes or adverse events may be less complete. If so, inclusion of such studies may bias the results toward the null hypothesis.

It is possible that relevant studies could have been missed during the literature searches. Although we used multiple different combinations in our Medline search to increase our sensitivity and carefully reviewed reference lists from all studies, we limited our search to English language studies. Meta-analyses may be biased by inability to identify unpublished studies. Our data support the possibility that publication bias may have existed. We observed lower relative risk estimates among large studies and higher estimates among smaller studies. This suggests that some smaller studies showing an increased risk of lymphoma may have been published while small studies demonstrating no increased risk remain unpublished. Importantly, small studies contribute relatively little weight in the analyses. For example, in our sensitivity analysis, where we exclude the study by Farrell and colleagues, the estimated relative risk only decreased from 4.18 to 3.49.

Azathioprine/6-MP tend to be used in more severe cases of IBD. It is possible that the increased risk of lymphoma with immunomodulator therapy seen in these studies is confounded by the severity of IBD (that is, those who are at the greatest risk for lymphoma are the same population that are most likely to receive immunomodulator therapy). Arguing against confounding by indication are the results of a previous cohort study addressing the risk of lymphoma among patients with IBD. In the study by Lewis and colleagues, IBD patients with lymphoma were no more likely than IBD patients without lymphoma to have used 5-ASA medications, steroids, or to have undergone surgery in the previous two years. Thus, in that study, severity of IBD did not appear to be strongly associated with lymphoma risk. None the less, a conservative interpretation of our data is that IBD patients who receive immunomodulator medications are at higher risk of lymphoma than the general population, and that this increased risk could be due to the medication, disease activity, or both.

Meta-analyses are limited by the quality of the published data. Several of the studies included in our analyses had potential limitations. For example, the study by Farrell and colleagues had a small sample size, and as such relatively imprecise relative risk estimates. Likewise, we elected to include the study by Kinlen although only one third of the non-rheumatoid arthritis patients analysed had IBD. However, when the Farrell and Kinlen studies were excluded from the analysis, we obtained similar results.
All of the studies included in our meta-analysis compared the incidence of lymphoma among IBD patients treated with immunomodulator therapy to that in the general population. Justification for this approach comes from multiple population-based studies that support a lack of association between IBD and lymphoma.11–17 Thus the incidence of lymphoma in the general population should approximate that observed in IBD populations. Further evidence to support this comes from our pooled analysis of the studies with a second control IBD population. This evidence further supports this conclusion. Further evidence to support this comes from our pooled analysis of the studies with a second control IBD population.

It is unclear how the risk of lymphoma changes when therapy is discontinued, and whether the risk is dose related. The studies included in our meta-analysis all continued to follow patients after azathioprine/6-MP were discontinued. If immunosuppression is the primary mechanism leading to lymphoma, it is plausible that the risk of lymphoma would return to normal after the medications are discontinued. To the extent that this is true, individual studies and our meta-analysis may have underestimated the risk of lymphoma during therapy with azathioprine/6-MP. Regardless, our estimates should remain valid for the long term follow up of patients treated with these medications.

In conclusion, our data suggest an approximate fourfold increased risk of lymphoma in IBD patients treated with azathioprine/6-MP. This is consistent with relative risk estimates observed in previous studies among patients with rheumatoid arthritis. Our data suggest that one additional lymphoma will occur every 300 to 4500 years after therapy with azathioprine or 6-MP, depending on the age of the patient. Because these data were obtained from observational studies, it is not possible to fully exclude the possibility that the increased risk of lymphoma is associated with the severity of the disease, rather than being caused by the medications. Regardless, given the magnitude of the association, even if the increased risk is entirely attributable to the medications, it is unlikely to outweigh the potential benefits of these medications for most patients. In fact, to the extent that the observed increased risk of lymphoma is in part or entirely attributable to the severity of the underlying IBD, the benefit to risk ratio for these medications would be even greater.

ACKNOWLEDGEMENT
This study was supported in part by a grant from the National Institutes of Health (grant K08-DK02589 to JDL). The funding agency had no input into the design, interpretation of the results, or presentation of the data.

Conflict of interest: declared (the declaration can be viewed on the Gut website at http://www.gutjnl.com/supplemental)

REFERENCES

Authors’ affiliations
A Kandiel, Department of Gastroenterology and Hepatology, The Cleveland Clinic, Cleveland, Ohio, USA
A G Fraser, Department of Medicine, University of Auckland, Auckland, New Zealand
B I Korelitz, Department of Medicine, Lenox Hill Hospital, New York, USA
C Brensinger, Department of Biostatistics and Epidemiology, and Center for Clinical Epidemiology and Biostatistics, University of Pennsylvania, Philadelphia, PA
J D Lewis, Department of Medicine, Department of Biostatistics and Epidemiology, and Center for Clinical Epidemiology and Biostatistics, University of Pennsylvania, Philadelphia, PA, USA

www.gutjnl.com