Risk for Serious Gastrointestinal Complications Related to Use of Nonsteroidal Anti-inflammatory Drugs

A Meta-analysis

Sherine E. Gabriel, MD, MSc; Liisa Jaakkimainen, MSc; and Claire Bombardier, MD

■ *Objective:* To describe the relative risk for serious gastrointestinal complications due to nonaspirin nonsteroidal anti-inflammatory drug (NSAID) exposure among NSAID users as well as in selected subgroups.

■ *Design:* Overview and meta-analysis.

■ *Data Identification:* A literature search of English-language studies examining the association between NSAIDs and adverse gastrointestinal events for the period 1975 to 1990 identified using MEDLINE and communicating with three internationally recognized experts.

■ *Data Analysis:* A qualitative summary of study characteristics and a critical appraisal of study quality were done. The results of 16 primary studies were selected and combined statistically. Summary estimates were weighted by sample size and quality score.

■ *Main Results:* The overall odds ratio of the risk for adverse gastrointestinal events related to NSAID use, summarized from 16 studies (9 case-control and 7 cohort) was 2.74 (95% CI, 2.54 to 2.97). The summary odds ratios were as follows: elderly patients, (aged ≥ 60 years), 5.52 (CI, 4.63 to 6.60); patients under 65 years of age, 1.65 (CI, 1.08 to 2.53); women, 2.32 (CI, 1.91 to 2.82); and men, 2.40 (CI, 1.85 to 3.11). The summary odds ratio for NSAID users receiving concomitant corticosteroids compared with NSAID users not receiving corticosteroids was 1.83 (CI, 1.20 to 2.78). The summary odds ratio for the first gastrointestinal event was 2.39 (CI, 2.16 to 2.65). The relative risk for a subsequent or unspecified gastrointestinal event was 4.76 (CI, 4.05 to 5.59). The summary odds ratio for less than 1 month of NSAID exposure was 8.00 (CI, 6.37 to 10.06); for more than 1 month but less than 3 months of exposure, the summary odds ratio was 3.31 (CI, 2.27 to 4.82); and for more than 3 months of exposure, the summary odds ratio was 1.92 (CI, 1.19 to 3.13).

■ *Conclusions:* Users of NSAIDs are at approximately three times greater relative risk for developing serious adverse gastrointestinal events than are nonusers. Additional risk factors include age greater than 60 years, previous history of gastrointestinal events, and concomitant corticosteroid use. Another possible risk factor is the first 3 months of NSAID therapy. The risk for serious gastrointestinal events appears to be equal among men and women. These data represent summary statistics from 16 studies and cannot be considered generalizable to all NSAID users.

Annals of Internal Medicine. 1991;115:787-796.

From the Mayo Clinic and Mayo Foundation, Rochester, Minnesota; and Wellesley Hospital, Toronto, Ontario. For current author addresses, see end of text.

Nonsteroidal anti-inflammatory drugs (NSAIDs) are the most widely used agents for the treatment of musculoskeletal and arthritic syndromes (1). Use of these agents has been increasingly associated with gastrointestinal toxicity, including mild dyspepsia, as well as more serious gastrointestinal reactions such as bleeding, perforation, and other events leading to hospitalization or death. Although researchers agree that an increased risk for gastrointestinal toxicity exists with NSAID use, the size of the reported risk has varied markedly, and there is little agreement on the definition of "high risk" groups (2-19).

We reviewed the literature on NSAID-related adverse gastrointestinal events. First, we summarized study characteristics and appraised study quality. We then did a meta-analysis of all controlled trials that examined the risks for serious gastrointestinal events among NSAID users. Our primary objective was to estimate a summary odds ratio or relative risk for serious gastrointestinal complications due to nonaspirin NSAID exposure.

Methods

A comprehensive search of the English-language literature from 1975 to 1990 was conducted using MEDLINE and searching the following terms: anti-inflammatory agents, non-steroidal; gastropathy, toxicity, adverse effects, or side effects; peptic ulcer or dyspepsia; gastric erosion, gastritis, gastric ulcer, gastric mucosa, endoscopy; and human. We also searched for specific NSAIDs by name.

Five hundred twenty-six references were obtained. These were reviewed by one of the authors, and any citation that mentioned NSAID-related gastrointestinal events was selected (Figure 1). One hundred forty-two articles met this criterion and were entered into "Reference Manager" (20). Five additional articles were identified by communication with three investigators (Marie Griffin, MD; Michael Langman, MD; and Richard Hunt, MD) from the United States, United Kingdom, and Canada, respectively. These 5 articles were added to the data set, for a total of 147 articles.

From the 147 articles in the data set, 40 studies were selected that examined the association between NSAIDs and adverse gastrointestinal events. Specific inclusion and exclusion criteria were applied to these studies independently by two of the authors. All studies that contained a comparison group and provided an estimate of risk for serious gastrointestinal complications (defined as bleeding, perforation, or other adverse gastrointestinal events resulting in hospitalization or death) in NSAID users compared with nonusers, regardless of underlying disease, were included in the meta-analysis. A study was excluded if its primary objective was to assess effectiveness, if it involved the treatment of children (under 18 years of age), if it described fewer than ten patients, if the only NSAID studied was salicylate, or if the outcome examined was

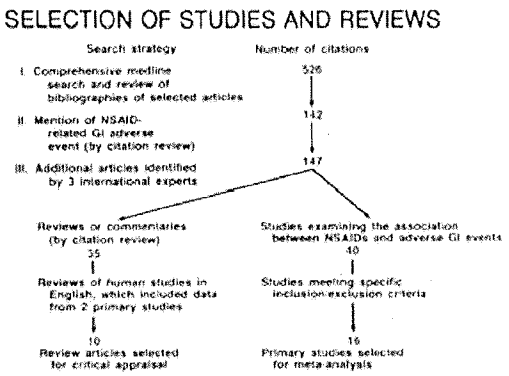

Figure 1. Selection of studies and reviews.

the identification of ulcer rather than the presence of serious gastrointestinal complications. Disagreements between the two reviewers were resolved by consensus. Sixteen studies were selected (21-36) for meta-analysis (*see* Figure 1).

Meta-analysis

The following criteria were used to evaluate the quality of the studies included in the meta-analysis: blinding, definition of outcome, case selection, control selection, matching technique, definition of exposure, and control for confounders (Appendix A). The Methods section of each study was photocopied, with care taken to exclude any mention of the authors' names, study results, or journal title. Study quality was evaluated in a blinded fashion by two of the investigators. Quality scores were assigned to each criterion according to its relative importance. A quality score of 0 indicated poor definitions and no attempt to avoid bias, and a score of 46 indicated the converse. The average score (between the two readers) among the first six categories constituted the baseline score for the study. For every 5 confounders identified in a primary study, 1 bonus point was awarded, to a maximum of 5 points for studies that identified more than 25 confounders. Thus, the maximum quality score attainable was 51. Agreement between the two readers regarding the quality score was evaluated using the kappa statistic (37).

Data from all articles were abstracted in duplicate to avoid errors. The two observers met, discussed each item, and resolved all disagreements and errors. A final copy of the completed data collection forms was then created and entered into a database (ORACLE, Oracle Corporation, Belmont, California) (38).

The results of the 16 primary studies were combined statistically using two different techniques. First, overall point estimates of the odds ratios and 95% confidence intervals (CIs) were calculated from the raw data of the 16 selected studies using the Mantel-Haenszel statistic (39). The second technique involved combining the published odds ratios and CIs directly across studies to produce an overall estimate of the odds ratio and 95% CI (40). The latter will hereafter be referred to as the "direct" method. The direct method was the primary statistical analysis technique used, and all results were calculated using this method unless otherwise stated.

The purpose of this analysis was not to estimate a common parameter, but rather to compute an average or summary statistic across the 16 selected studies. The CI for this statistic cannot, therefore, be generalized beyond the study samples. All summary estimates were weighted by sample size. The influence of the quality scores on the summary estimates was evaluated using logistic-regression analysis with quality score as a covariate.

Overall odds ratios for all studies included in the meta-analysis as well as odds ratios for various subgroups were calculated. The overall odds ratios referred to the odds ratios combined from the main research questions of each of the studies. Summary odds ratios for various subgroups were calculated from those studies which provided data on these subgroups. The method of Breslow and Day was used to test for homogeneity of the Mantel-Haenszel estimates (41). Tests of homogeneity were also performed for the direct method according to the method of Greenland (40).

Results

We selected 16 studies (9 case-control and 7 cohort) that specifically examined the risks for clinically defined, NSAID-related, adverse gastrointestinal events (21-36). The reported relative risks varied from 1.0 (34) (indicating no increased risk for gastrointestinal events) to 13.7 (29) (indicating a risk for NSAID users 13.7 times greater than that for nonusers). Two potential sources of variability were identified: differences in study characteristics and differences in study quality.

Study Characteristics

Study characteristics are shown in Appendix B. For both the case-control and cohort studies, serious gastrointestinal events were defined among hospital-based cases. Among the case-control studies, the ascertainment of gastrointestinal outcome was not done in a uniform manner. Gastrointestinal events were assessed based on the results of endoscopy, roentgenography, or surgery (27-29, 33, 35) or on a clinical diagnosis of hematemesis or melena (26, 30-32). Some case-control studies used community controls (31, 33, 35); others compared cases with hospital controls (28-30, 32) or used both types of controls (26, 27). Most studies matched controls directly with cases (26-28, 30, 31, 33). Two case-control studies used a nested case-control

Table 1. Study Quality Scores

Study (reference)	Baseline (range, 0-46)*	Bonus (range, 0-5)†	Total (range, 0-51)
Griffin et al. (33)	25.5	4.00	29.5
Levy et al. (32)	24.5	5.00	29.5
McIntosh et al. (35)	22.5	5.00	27.5
Somerville et al. (26)	22.5	4.00	26.5
Bartle et al. (27)	23.0	3.00	26.0
Henry et al. (30)	20.5	2.00	22.5
Jick et al. (31)‡	20.0	1.00	21.0
Carson et al. (24)‡	15.5	5.00	20.5
Guess et al. (25)‡	16.0	3.00	19.0
Bloom (22)‡	14.5	4.00	18.5
Beard et al. (23)‡	14.5	4.00	18.5
Beardon et al. (21)‡	13.5	2.00	15.5
Armstrong and Blower (29)	14.5	0.00	14.5
Collier and Pain (28)	13.5	1.00	14.5
Jick et al. (34)‡	10.0	2.00	12.0
Alexander et al. (36)	9.50	1.00	10.5

* Baseline scores were assigned based on an evaluation of the following design items: explicit definitions of exposure, outcome, case and control status as well as the use of blinding and matching.

† Bonus points were assigned based on the number of confounders, which were accounted for in the analysis. *See* text for method of bonus-point assignment.

‡ Cohort studies.

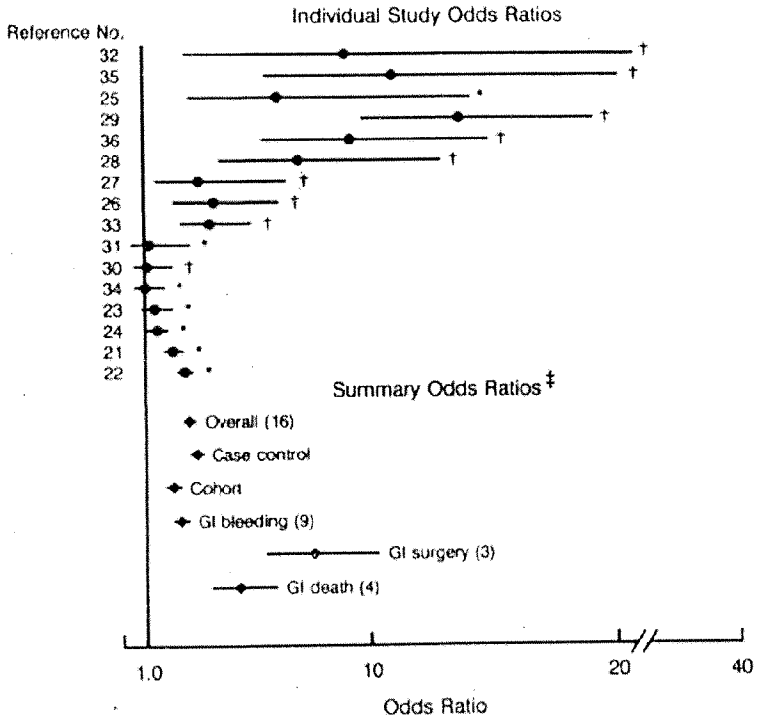

Figure 2. Individual study and summary odds ratios. Individual study odds ratios are arranged in order of increasing sample size (top to bottom). Individual study odds ratios were provided in the original studies (21-28, 30-32, 34, 36) or calculated from data provided in original studies (29, 33, 35). (●, Individual study odds ratio; ◆, summary odds ratio; the 95% confidence intervals are indicated by the extended lines; * cohort study; † case-control study; ‡ odds ratios summarized by "direct" technique [40]; numbers in parentheses are the number of studies combined.)

design (31, 33). Determinations of NSAID exposure were made by an unblinded review of clinical notes (28-30), a structured questionnaire with interviewers who were blinded (26, 27, 32, 35), or an extraction of prescription data from pharmacy computer files (31, 33). In all cohort studies, the assessment of NSAID exposure was based on prescription files. Estimates of the duration of NSAID exposure varied from 30 days (24, 25) to 90 days (22, 23, 31, 34). One cohort study (25) examined deaths from gastrointestinal causes, whereas the remainder looked at hospitalizations caused by gastrointestinal complications. Samples examined in the cohort studies included the Group Health Cooperative in Puget Sound; the Pennsylvania Medicaid group; the residents of Saskatchewan, Canada; and the residents of the Tayside Region, Scotland. The Puget Sound Group Health Cooperative represents a younger, employed population, the Medicaid group is elderly, and the Tayside and Saskatchewan groups represent residents of geographically diverse districts.

Study Quality

Table 1 shows the study quality scores. Methodologic assessment of the 16 studies showed acceptable agreement between two observers for the six study quality categories evaluated (mean kappa, 0.70; minimum, 0.56; maximum, 0.83). The mean kappa for the quality category of blinding was 0.67 (minimum, 0.0; maximum, 1.0); for case selection, 0.75 (minimum, 0.66; maximum, 0.90); for control selection, 0.68 (minimum, 0.4; maximum, 1.0); for definition of exposure, 0.74 (minimum, 0.59; maximum, 0.96); for matching technique, 0.83 (minimum, 0.66; maximum, 1.0); and for definition of outcome, 0.56 (minimum, 0.0; maximum, 1.0). Disagreements regarding control of confounders were reexamined and resolved by consensus. The six studies with the highest quality scores were case-control studies (Table 1). These studies gave more explicit definitions of cases, controls, and exposure and used blinding more frequently. The study quality score was not found to be a significant covariate in the regression model ($P > 0.2$).

Summary Odds Ratios

Published odds ratios and summary odds ratios from the primary studies are shown in Figure 2. The overall odds ratio of the risk for adverse gastrointestinal events related to NSAID use (summarized from 16 case-control and cohort studies) is 2.74 (CI, 2.54 to 2.97). The summary odds ratio (combined from 8 studies) for elderly persons is 5.52 (CI, 4.63 to 6.60). In the cohort studies, the term "elderly" refers to persons 65 years of age or older. In the case-control studies, "elderly" refers to persons 60 years of age or older. The summary odds ratio for nonelderly persons, combined from 3 studies, is 1.65 (CI, 1.08, 2.53). These data show a greater than threefold increase in relative risk for serious gastrointestinal events among elderly NSAID users when compared with nonelderly users.

Odds ratios were subdivided by gastrointestinal out-

Table 2. Comparison of Summary Odds Ratios and Confidence Intervals Obtained by Two Methods

Category	Number of Studies Combined	Summary Odds Ratio	95% CI
Overall	12*/16†	2.86*/2.74†	2.62 to 3.12*; 2.54 to 2.97†
Patient ≥ 60 years of age	6/8	6.24/5.52	5.21 to 7.48; 4.63 to 6.60
Patient < 60 years of age	2/3	3.07/1.65	1.62 to 5.82; 1.08 to 2.53
Gastrointestinal bleeding	7/9	2.71/2.39	2.26 to 3.24; 2.11 to 2.70
Gastrointestinal surgery	3/3	7.04/7.75	5.34 to 9.29; 5.83 to 10.31
Gastrointestinal cause of death	3/4	4.22/4.79	3.24 to 5.50; 3.64 to 6.22
Unspecified adverse gastrointestinal event	2/3	2.68/1.79	2.42 to 2.98; 1.70 to 1.90

* Mantel-Haenszel technique for case-control studies only.
† Direct technique method of Greenland (reference 40).

come. The odds ratio for gastrointestinal bleeding, combined from nine studies, was 2.39 (CI, 2.11 to 2.70). The odds ratio for gastrointestinal surgery, combined from three studies, was 7.75 (CI, 5.83 to 10.31). The summary odds ratio for gastrointestinal death, combined from four studies, was 4.79 (CI, 3.64 to 6.22). Thus, the relative risk for surgical or fatal outcomes among NSAID users is 2- or 3-fold higher than the relative risk for gastrointestinal bleeding.

The summary odds ratio for women was 2.32 (CI, 1.91 to 2.82), whereas the summary odds ratio for men was 2.40 (CI, 1.85 to 3.11). The summary odds ratio for women compared with men was 1.15 (CI, 0.89 to 1.50). These data do not support gender as an independent risk factor. The risk for first compared with subsequent gastrointestinal event was also examined. The summary odds ratio for the first gastrointestinal event, combined from six studies, was 2.39 (CI, 2.16 to 2.65). The relative risk for subsequent or unspecified gastrointestinal event, combined from the remaining 10 studies, was 4.76 (CI, 4.05 to 5.59). These data suggest that patients with a history of gastrointestinal events may have an increased relative risk for further events. The use of concomitant corticosteroids was also examined. The summary odds ratio for NSAID users receiving concomitant corticosteroids compared with NSAID users not receiving corticosteroids was 1.83 (CI, 1.20 to 2.78). This finding suggests an approximately twofold increase in the relative risk among NSAID users who are receiving corticosteroids compared with NSAID users not receiving corticosteroids.

Summary odds ratios were also obtained using the Mantel-Haenszel statistic. A comparison of the results obtained by the two statistical techniques showed that the direct method enabled the use of data from more studies, resulting in narrower CIs. Summary odds ratios by both methods were similar in most categories (Table 2).

Summary odds ratios calculated according to individual NSAID used and duration of NSAID exposure were as follows: piroxicam, 11.12 (CI, 6.19 to 20.23); indomethacin, 4.69 (CI, 2.97 to 7.41); aspirin, 3.38 (CI, 2.26 to 5.01); naproxen, 2.84 (CI, 1.68 to 4.82); and ibuprofen, 2.27 (CI, 1.85 to 2.80). There is substantial overlap in the CIs among NSAIDs. The duration of NSAID consumption may be related to the size of the odds ratio (Figure 3). The summary odds ratio for less than 1 month of NSAID exposure was 8.00 (CI, 6.37 to 10.06); for longer than 1 month but less than 3 months, 3.31 (CI, 2.27 to 4.82); and for longer than 3 months, 1.92 (CI, 1.19 to 3.13). The highest odds ratios were obtained from studies in which the duration of NSAID consumption was less than 1 month.

Data were also subdivided by gastrointestinal event and age (Table 3). The relative risk for gastrointestinal surgery for nonelderly individuals, combined from three studies, was 0.44 (CI, 0.29 to 0.66), whereas the risk for gastrointestinal surgery among elderly persons, combined from three studies, was 10.42 (CI, 7.40 to 14.66). These data suggest a tenfold increase in relative risk for gastrointestinal surgery among elderly users when compared with younger users.

Estimates of the prevalence of serious gastrointestinal events among NSAID users were summarized from four cohort studies (7, 23, 25, 34). The summary, 1-year prevalence among NSAID users was 1 per 1000; the prevalence among elderly users (≥ 65 years of age) was 3.2 per 1000; and the prevalence among younger users (< 65 years of age) was 0.39 per 1000.

Sources of Heterogeneity

Tests for homogeneity were statistically significant ($P < 0.05$) for all analyses, indicating that the differ-

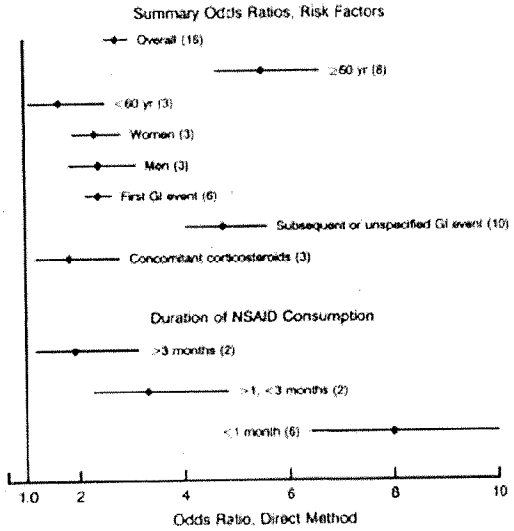

Figure 3. Summary odds ratios and risk factors. (◆, Summary odds ratio; the 95% confidence intervals are indicated by the extended line; numbers in parentheses are the number of studies combined.)

ences among the results of individual studies are greater than can be expected on the basis of chance alone.

We did two different types of analyses to identify sources of heterogeneity. Heterogeneity across studies is composed of intrastudy heterogeneity and inter-study heterogeneity. In an effort to describe intra-study heterogeneity, tests of homogeneity were conducted for several subgroups across studies. These subgroups were subdivided according to gastrointestinal outcome, age, age and gastrointestinal outcome, and use of individual NSAIDs. Each of these subgroups accounted for a portion of the variability, thus reducing the test statistic for homogeneity. There was, however, no subgroup identified that accounted for most of the observed heterogeneity. In an effort to describe interstudy heterogeneity, we did a multivariate regression analysis using the log of the study odds ratio as the dependent variable and study design, duration of NSAID use, gastrointestinal outcome, and average age as the independent variables. The regression was weighted using the individual study variances. The four independent variables accounted for approximately half of the interstudy variability.

Discussion

Two research designs have been used to study the risk for gastrointestinal events related to NSAID therapy: retrospective cohort and case-control studies. Most of the cohort studies used secondary analysis of health insurance registries in which data were collected primarily for billing purposes. The computerized case definition for gastrointestinal events is subject to substantial misclassification (42-44). Misclassification rates of up to 29% were noted in studies using retrospective chart review to confirm computerized diagnoses (23, 31, 34), resulting in contamination of the case group by controls and of the control group by cases and thus reducing the relative-risk estimate. Similarly, the information on NSAID exposure obtained from these registries may not have been of optimal quality. The duration of NSAID exposure is often unknown and assumptions are made from prescription registries regarding the average duration of NSAID use. Some studies estimated an average prescription duration of 90 days with full patient compliance (23, 31, 34). Such an assumption may overestimate the duration of NSAID exposure, biasing the results toward a falsely low relative risk. The frequency of NSAID use in a study sample determines the power of that study to detect a statistically significant relative risk (45). Nonsteroidal anti-inflammatory drug use among patients with prepaid health plans may be lower than that of the general population, further underestimating the relative risk. These factors contribute to the lower relative risks reported by the cohort studies when compared with the case-control studies.

In two case-control studies, different techniques were used to determine NSAID exposure among case patients and controls (28, 29). Physicians hospitalizing patients with gastrointestinal bleeding are more likely to inquire about NSAID use than are physicians questioning controls or their relatives. Such differences in the determination of NSAID exposure bias the results toward a falsely large relative risk. The use of a structured interview administered by an investigator who is blinded to the status (case patient or control) of the patient results in more valid estimates of relative risk (17, 26). Well-designed, nested, case-control studies minimize the selection bias, inherent in hospital-based case-control studies (33).

Although there have been many studies examining the gastrointestinal risks of NSAID use, important methodologic limitations and differences in study characteristics contribute to the conflicting results. Retrospective cohort studies probably underestimate the relative risk, whereas some case-control studies probably overestimate it. The aggregation of the results from observational studies is controversial (46). The strongest studies are those that defined cases, controls, outcome, and exposure accurately and reproducibly (26, 27, 32, 33, 35), as reflected by the quality-assessment scores in this meta-analysis (Table 1).

We conducted a structured overview of all previous reviews of NSAID-related adverse gastrointestinal events. The quality of the 10 reviews selected (3, 6-12, 18, 19) was assessed according to several criteria: the comprehensiveness of the literature search, the minimization of bias in the selection of primary studies, the assessment of the quality of the primary studies, the appropriateness of the techniques used in data synthesis, and the validity of the conclusions made by the authors as supported by the data. Most of the published reviews on this topic cite only a portion of the available literature, do not provide a critical assessment of the quality of the studies cited, and fail to combine the results of these studies statistically. Only 1 of the 10 reviews used a clearly defined, comprehensive search strategy (6). Inclusion criteria were stated for 2 of the 10 reviews (6, 8). A quality assessment of the studies was done in only 1 review (6). Appropriate, explicitly stated methods of data synthesis were given in only 2 reviews (6, 19).

Table 3. Subgroup Odds Ratios Combined from Case Control and Cohort Studies Using the "Direct" Method*

Variable	Number of Studies Combined	Summary Odds Ratio	95% CI
Gastrointestinal event by age*			
< 60 years			
Gastrointestinal bleeding	1	1.03	0.60 to 1.76
Gastrointestinal surgery	3	0.44	0.29 to 0.66
≥ 60 years			
Unspecified gastrointestinal adverse event	3	1.78	1.69 to 1.87
Gastrointestinal bleeding	9	2.38	2.10 to 2.69
Gastrointestinal surgery	3	10.42	7.40 to 14.66
Gastrointestinal cause of death	4	4.40	3.35 to 5.79

* Gastrointestinal events occurring in hospitalized patients.

DOCKET ALARM

Explore Litigation Insights

Docket Alarm provides insights to develop a more informed litigation strategy and the peace of mind of knowing you're on top of things.

Real-Time Litigation Alerts

Keep your litigation team up-to-date with **real-time alerts** and advanced team management tools built for the enterprise, all while greatly reducing PACER spend.

Our comprehensive service means we can handle Federal, State, and Administrative courts across the country.

Advanced Docket Research

With over 230 million records, Docket Alarm's cloud-native docket research platform finds what other services can't. Coverage includes Federal, State, plus PTAB, TTAB, ITC and NLRB decisions, all in one place.

Identify arguments that have been successful in the past with full text, pinpoint searching. Link to case law cited within any court document via Fastcase.

Analytics At Your Fingertips

Learn what happened the last time a particular judge, opposing counsel or company faced cases similar to yours.

Advanced out-of-the-box PTAB and TTAB analytics are always at your fingertips.

API

Docket Alarm offers a powerful API (application programming interface) to developers that want to integrate case filings into their apps.

LAW FIRMS
Build custom dashboards for your attorneys and clients with live data direct from the court.

Automate many repetitive legal tasks like conflict checks, document management, and marketing.

FINANCIAL INSTITUTIONS
Litigation and bankruptcy checks for companies and debtors.

E-DISCOVERY AND LEGAL VENDORS
Sync your system to PACER to automate legal marketing.