A reader’s guide to the evaluation of causation

Theresa Podrebarac, Peter Tugwell, Paul C Hébert

The Oxford Dictionary defines cause as ‘that which produces an effect’. In 1882, Koch developed his postulates as a means of determining whether an infectious organism caused a disease. He stated that the organism must be present in every case of the disease and must be isolated and grown in pure culture. Additionally, the organism when injected into a susceptible host, must cause the specific disease and then must be recovered from the animal.1 These postulates were a landmark in infectious diseases but it soon became clear that not all diseases had an infectious aetiology. In the post-antibiotic era, chronic diseases do not conform to such simple causal models. Instead, there may often be many aetiologies including remote causes along with certain genetic predispositions and interactions with the environment.

When randomised clinical trials are not possible to determine cause and effect, the evidence used may be limited to weaker study designs. In 1965, Sir Austin Bradford–Hill proposed a set of criteria used to guide decisions regarding causation.2 An updated and modified version of the criteria may be applied to evaluating aetiologies of diseases and may also be used to establish a causal relationship between a drug and an adverse effect. These guidelines are summarised in box 1.1,2

Judgments on causation are made by clinicians on a regular basis, most frequently in the context of whether a particular therapy is doing more harm than good. Increasingly, clinicians are required to estimate the risk of side-effects which may be caused by a recommended therapy. This can be rather complex since adverse events occurring during the administration of a medication, for example, are not necessarily caused by that particular drug. When two pharmacologists asked healthy university students who were not taking any medication, if they had experienced symptoms often cited as side-effects, such as nausea, rash or diarrhoea, in the past 72 hours, 81 % of the respondents answered yes.3 The emerging field of pharmaco-epidemiology has been responsible for developing strategies to assess the scientific validity of studies which attempt to provide estimates of risk of a proposed therapy. This has been vital for physicians in terms of helping patients make informed decisions about treatment. A clinical scenario will be used to illustrate how these criteria to establish causation may be applied to solve practical problems (box 2). The proposed guidelines may apply to several other situations that arise in clinical practice.

Searching for the answer

There are a number of approaches to determining the answer to difficult clinical problems. For example, you may start by asking an expert in the area to cite specific references or you could consult a reference textbook. You may proceed to the medical library or, in this instance, check with the drug information service if one is available to provide further references. With the advent of computerised searches, it is now possible to conduct an electronic search of the medical literature from your home or office to obtain directly the most current literature addressing a particular topic. In searching for the answer to our clinical problem, we tapped into most of these resources. Review articles including basic science reviews were read initially to provide an overview of the topic. These articles incorporated several expert opinions. Selected studies cited were obtained for closer scrutiny. A MEDLINE search was performed with GRATEFUL MED for English-language articles published between 1980 and 1993. Specifically, we were most interested in high-quality systematic overviews, randomised, double-blind, controlled studies and prospective cohort studies that linked nonsteroidal anti-inflammatory drugs (NSAIDs), including aspirin, with gastrointestinal complications. The MeSH (Medical Subject Heading) terms used were NSAIDs or anti-inflammatory drugs and gastrointestinal disease. It is important to use the correct index terms or MeSH headings in order to identify relevant articles. A librarian may help you find the most applicable MeSH headings for specific searches.
Since over 200 articles were found, abstracts were scanned for representative articles on well-designed studies to which the guidelines outlined in box 1 could be applied. In this search, we identified only one randomised, double-blind, placebo-controlled clinical trial, one prospective cohort study, a meta-analysis, several review articles and basic science overviews which will form the basis of this discussion.

**Guidelines for determining causation**

- was the type of study the strongest that could have been performed under the circumstances?
- what is the strength of the association? Are the relative risks or odds ratios both statistically and clinically important?
- is the temporal sequence of exposure and outcome correct?
- do the results remain consistent from study to study?
- is there a dose–response relationship?
- does the association make epidemiologic sense?
- does the association make biologic sense?
- is the association specific, that is, one cause and one effect?
- is there an analogous cause and effect relationship?

Adapted from[1,2]

**Study designs used to establish causation**

Randomised, double-blind, placebo-controlled trials provide the strongest and most reliable evidence establishing causal relationships. The random allocation of patients to either the study group or the placebo group minimises any bias and the influence of confounding variables. This study design best approximates the true biologic experiment. Individuals in both groups are followed prospectively and are then compared to determine whether a particular outcome developed with increased frequency in the study group.

The next best type of study design is the prospective cohort. Individuals treated with a given medication are simply followed over time along with patients who are not taking the medication. These groups are then evaluated to determine if a particular outcome developed in the group taking the medication versus those who were not taking it. The advantage of this type of design is that it is especially useful when it is impossible or unethical to do a randomised controlled trial. This study design can be very useful in describing the natural history of a disease and in establishing a link between exposure and eventual outcome. The disadvantage is that if the outcome is uncommon, many individuals must be followed for several years to determine an effect and this is both time-consuming and costly. Another disadvantage is that this design is vulnerable to selection or measurement biases, as well as confounding biases which can be minimised by choosing appropriate control subjects.

Most studies examining the association between NSAIDs and gastrointestinal complications have been done using either a retrospective cohort or a case-control design. In a retrospective cohort study, individuals with a certain exposure are identified retrospectively from some source such as health insurance or prescription registries. The subjects are then followed forward to the present. This study design may be limited by incomplete records and, potentially, by selection biases. Case-control studies are designed to identify subjects with a prespecified outcome of interest, in this instance, stomach ulcers or bleeding, and then to determine the influence of a certain exposure such as a medication. This design has many methodological shortcomings in determining causal associations. Frequently, multiple exposures are assessed and this type of design is most prone to selection and confounding biases. However, case-control studies are often the only option in rare diseases and can be completed in a short period of time. Both the retrospective cohort and case-control studies are very weak designs.

In addition to primary studies, integrative studies such as a meta-analysis or systematic overview are efficient methods of reviewing a topic. However, the quality of these types of studies is only as good as the primary articles from which they draw their data. Box 3 gives guidelines for evaluating the quality of such studies. The term 'meta-analysis' literally means an analysis of analysis or information synthesis. It is an approach which can be particularly useful for answering a specific question and can be done in different ways. The main objective is often to be able to combine data from various studies and in effect create one large study database with the results having more statistical power than the individual studies alone.

By using the primary studies and the meta-analysis, we will attempt to establish a causal relationship between NSAIDs and gastrointestinal complications. Each of the criteria in box 1 will be discussed in turn as they pertain to our clinical scenario and will be scored using a six-point scale according to the strength of the evidence. The results are summarised in box 4. This scoring system was used in a recent review evaluating the causal relationships between *Helicobacter pylori* infection and gastritis, duodenal ulcer, gastric cancer and nonulcer dyspepsia.

Let us now return to the clinical problem and evaluate the strength of the evidence showing a causal relationship between NSAIDs and gastrointestinal toxicity.

**Box 1**

Guidelines for determining causation

- **was the type of study the strongest that could have been performed under the circumstances?**
- **what is the strength of the association?** Are the relative risks or odds ratios both statistically and clinically important? 
- **is the temporal sequence of exposure and outcome correct?**
- **do the results remain consistent from study to study?**
- **is there a dose–response relationship?**
- **does the association make epidemiologic sense?**
- **does the association make biologic sense?**
- **is the association specific, that is, one cause and one effect?**
- **is there an analogous cause and effect relationship?**

Adapted from[1,2]

**Box 2**

Clinical problem

A 26-year-old woman comes to you with a 10-week history of an inflammatory arthritis affecting both hands, wrists and knees. She is unable to care for her two-year-old son. After completing your initial evaluation, you decide to prescribe a nonsteroidal anti-inflammatory drug (NSAID). The patient then asks if this medication can cause stomach problems. She tells you that her father recently passed away from complications due to a stomach ulcer while taking an NSAID for his arthritis. You explain that the risk of ulcers may be less in her but that you are unsure of the exact risk. You ask her to return in a week. At that time, you will provide her with more detailed information.
hospitalisation for ulcer disease (the Aspirin Myocardial Infarction Study). This was a three-year clinical trial involving 4524 subjects whose primary objective was to assess the efficacy of regular aspirin in the reduction of mortality in both men and women between the ages of 30 and 69 who had had at least one myocardial infarction. There was a marked increased risk of hospitalisation for ulcer disease in the aspirin group.

The ARAMIS (Arthritis, Rheumatism and Aging Medical Information System) database is the largest prospectively followed cohort of some 2700 rheumatoid arthritis patients who have been followed for 9525 patient-years. It was the only prospective cohort study identified. These data identified 116 patients who had had 128 episodes of gastrointestinal disease requiring hospitalisation and all but nine of these patients were taking NSAIDs at the time of these events. The major disadvantage of this database is that it lacks a control group with no NSAID exposure. Without a proper control group, it is not possible to calculate a relative risk. Even with concurrent controls in a cohort design, there may be other unknown factors acting as confounders which make it difficult to assess the pre-existing risk for gastrointestinal disease.

The meta-analysis from 1991 found in the search was comprehensive and included 16 primary studies, nine of which were case-control studies and seven retrospective cohort studies. The investigators did not identify any randomised, controlled trials. The results showed a significantly increased risk of gastrointestinal adverse events across the studies but identified flaws in the study designs which could have underestimated the risk of these events. For example, by identifying exposure to NSAIDs through health insurance and prescription registries, there is the potential for significant bias and certain assumptions are made. Data gathered through these registries are often used for other purposes and may not be of sufficient quality.

Consider the following: to simply identify the fact that a prescription was written does not give sufficient information on whether the medication was taken, how it was taken or the duration of therapy. This information can only be obtained through structured interviews or laboratory tests which are time-consuming and costly. Some of the retrospective cohort studies assumed 100% compliance which may overestimate the actual exposure and hence bias towards a lower relative risk. Case-control studies are at even greater risk of significant bias. For example, patients who come in complaining of abdominal complaints are more likely to be asked about their NSAID use while patients taking NSAIDs are more likely to be investigated for ulcers.

In summary, one randomised, double-blind, placebo-controlled trial comparing aspirin to placebo demonstrated significantly increased gastrointestinal toxicity. The best data on NSAID come from the ARAMIS database, a prospective cohort study and the meta-analysis which included 16 primary studies. These studies, although flawed, have shown increased adverse gastrointestinal events on NSAIDs.

What is the strength of the association? Are the relative risks or odds ratios both statistically and clinically important?

In other words, what are the odds of developing the specific complication from being exposed to the therapy versus not being exposed. In general, the higher the odds the greater the strength of the association. If the treatment is harmless, the relative risk or odds ratio should be one. Statistical significance is often expressed using p-values or 95% confidence intervals to indicate the probability that the results are due to chance alone. Clinical significance is usually expressed as a relative risk when dealing with a cohort study. An odds ratio, an estimate to risk, comparable to the relative risk, is primarily used when analysing a case-control study. A relative risk is defined as the number of patients considered at risk who develop the disease and are exposed to the therapy (in this case NSAIDs) compared to similar patients who are not exposed. A range of relative risks and odds ratios have been reported in the various studies examining NSAIDs and gastrointestinal toxicity. The differing estimates of risk are probably a reflection of the various study designs and the different outcome measures that were used. Because of all the potential shortcomings affecting cohort and case-control studies, many epidemiologists would like to see relative risks of greater than two to demonstrate true significance.

The Aspirin Myocardial Infarction Study reported a relative risk of 10.7 times higher for developing a duodenal ulcer for the aspirin group versus placebo and 9.1 times higher for developing a gastric ulcer. The results were adjusted for age and smoking. These results are very impressive considering that the relative risk for developing lung cancer in smokers, compared to nonsmokers, ranges from 4 to 16.

The meta-analysis also reported a statistically significant increase in the risk of
developing gastrointestinal complications while on NSAIDs. The relative risks ranged from 1.05 to 13.7 with an overall odds ratio of 2.74. The data suggested that the risk was greatest during the first three months of therapy with NSAIDs and that concomitant use of steroids and advanced age also increased the risk. From the ARAMIS data, gastrointestinal pathology accounted for over 70,000 hospitalisations and over 7,000 deaths annually in the US. Patients with disability, using steroids and with prior NSAID side-effects were at increased risk. The chance of hospitalisation or death for a patient with rheumatoid arthritis due to gastrointestinal adverse events averaged between 1.3 to 1.6% over one year with higher risk patients having a risk of 4–5% per year. In summary, the strength of the association demonstrated in our selected studies was convincing both clinically and statistically across the different study designs.

Is the temporal sequence of exposure and outcome correct?
This criterion is usually self-evident in randomised, double-blind controlled trials and in prospective cohort studies. The data from the Aspirin Myocardial Infarction Study is convincing and the temporal sequence is correct. However, in case-control studies the outcome of interest is present at the time that risk factors and exposures are sought. It is often more difficult to ascertain whether the outcome, namely gastrointestinal ulceration, may have been silently present prior to the exposure to the drug. For example, some patients may take aspirin or other NSAIDs to relieve the gastrointestinal symptoms that were not recognised as being due to an ulcer. The gold standard in terms of the temporal sequence for identifying developing ulcers is through endoscopy and none of our studies included pre-exposure endoscopy which would have strengthened the hypothesis for causation.

Is the association consistent from study to study?
Consistency from study to study can provide some of the strongest evidence for causation. The best evidence for consistency is often derived from meta-analysis. Studies from several different countries also report an increased risk of gastrointestinal complications associated with NSAID usage. No studies have shown a protective effect of NSAIDs, and in the meta-analysis, 11 of the 16 studies showed an increased risk. A case-control study found during our search, compared different NSAIDs to each other with respect to the risk of perforation and haemorrhage. These data suggest that some classes of NSAIDs were more potent in terms of causing toxicity. Nevertheless, all NSAIDs have consistently shown the potential to cause adverse gastrointestinal events.

Is there a dose-response relationship?
If the frequency of an adverse event increases with increasing dose, this can be supportive evidence for causation although little is known about toxic thresholds and susceptible hosts. There is some evidence that acute and chronic gastrointestinal ulceration may be dose-dependent. In one of the case-control studies of elderly patients who had been hospitalised for confirmed peptic ulcer disease, the relative risk increased from 2.8 for taking 0.5 of a standard dose of an NSAID per day to 8 if more than three standard doses were taken. However, there are other data that show taking low-dose aspirin, 325 mg every other day, also increases risk.

Does the association make epidemiologic sense?
There are several approaches one can use to interpret this criterion. Firstly, have other plausible explanations for our observation been ruled out? For example, what is the role of H pylori in patients who are also on NSAIDs? This remains largely unknown. Is the association reversible, that is, if the exposure is withdrawn, does the gastrointestinal event rate also diminish? Has there been a corresponding increase in the number of NSAID prescriptions written and the number of hospitalisations for gastric ulcer complications? Data from the US have shown that hospital discharges for uncomplicated duodenal ulcer from 1970 to 85 have decreased while the rates of complications, ie, haemorrhage or perforation, have remained stable. However, since 1978 hospitalisation for gastric ulcers with haemorrhage have nearly doubled. In Canada in 1989, an estimated 10.6 million prescriptions were written for NSAIDs, an increase of 6% per year. Studies from several other countries report increased rates of gastric ulcer complications, trends which appear related to NSAID usage. The repeated demonstration of an association lends much credibility to our hypothesis that NSAIDs do indeed cause gastrointestinal ulceration and further complications.
Does the association make biologic sense? Several experimental studies have examined the multifactorial effects of NSAIDs on the gastric mucosa.\(^3\) NSAIDs produce mucosal effects through local irritation and through the systemic effect of prostaglandin inhibition. By depleting endogenous prostaglandins in rabbits through immunisation, upper gastrointestinal tract ulceration occurred that resembled the stress ulcers and erosions often seen with NSAIDs.\(^4\) NSAIDs affect several components of the mucosa which are necessary to preserve the acid barrier. NSAID use may lead to deep epithelial injury by interfering with the regulation of blood flow.\(^5\) Animal ulcer models show that NSAIDs inhibit mucosal proliferation at the edge of an ulcer which can be reversed with prostaglandin E\(_1\).\(^6\) As we are able to recreate the same ulcerations in animal models as in human subjects through the administration of NSAIDs, it stands to reason that this association is biologically sensible and this greatly strengthens our hypothesis.

Is the association specific, that is, one cause and one effect? This criterion is not essential in terms of showing causation. Even in infectious diseases where causation is more clearcut, a single infecting organism can cause a spectrum of disease. This also brings into play certain host factors which may influence the expression of disease and the fact that multiple factors may interact to produce the final effect. For example, group A streptococci can cause a sore throat or develop into full blown acute rheumatic fever. The term NSAID gastropathy has been used but it does not define a single entity. A recent post mortem study found that nonspecific small intestinal ulceration was increased in NSAID users and that gastric ulcers were more prevalent in short-term users.\(^7\) Ulcers have been well documented to occur due to \(H\) pylori infection especially duodenal ulcers. These ulcers are indistinguishable from NSAID-induced injury.\(^4\) Hence, there does not appear to be only one cause for gastric ulceration nor does it appear that NSAIDs confine their injury to the gastric mucosa.

Is there an analogous cause and effect relationship? In other words, is there good evidence which has already been well established for a similar exposure causing the outcome of interest, namely gastrointestinal ulceration? This again is often a weak link in establishing causation. Only if we were to separate aspirin from other NSAIDS could we state that what had been previously illustrated with respect to increased gastrointestinal toxicity and aspirin is analogous to the effect seen with NSAIDs.

How to apply these principles to your patient

Simple rigorous criteria should be considered prior to claiming a causal link between an exposure, eg, a medication and an outcome, such as an adverse reaction. Randomised, double-blind, placebo-controlled trials serve as the model design. These trials and very often weaker designs are identified and evaluated to determine the cumulative strength of the evidence to establish causation.

The final step is to now individualise the results. Is your particular patient similar to the average patient in the studies? If not, are they similar to a well-described subgroup of patients? These questions reflect the generalisability or external validity of the results. This can be determined by examining the eligibility criteria of the studies. In this instance, our patient would best fit with the ARAMIS database. If your patient is dissimilar to those described, you need to then determine whether the causal link is biologically plausible and therefore applicable to your patient.

In addition to this rigorous exercise, you can further individualise the case for causation by applying the following pharmaco-epidemiological principles. You may discontinue the suspect medication to see if the adverse effect disappears and then rechallenge the patient to see if it returns. This is done quite routinely in everyday practice. With some medications, a specific antagonist can be administered to reverse the adverse effect or drug levels may be used to explain certain toxicities.\(^2\) This may be further evaluated in a more formal fashion using an \(n = 1\) trial which randomly allocates the administration of a placebo with the active drug in a single patient to see if the adverse effect persists.\(^8\) These principles are not valid for evaluating potentially life threatening events such as gastrointestinal bleeding or anaphylaxis. A careful search for any other causes of the adverse event is always indicated. These guidelines are summarised in box 5.

Let us return to our clinical case of the 26-year-old female with inflammatory polyarthritis. You explain that there is strong evidence that NSAIDs do indeed increase the risk of ulcers and bleeding. The risk is estimated to be approximately three times greater in users when compared to nonusers of NSAIDs. Her risk of...
developing gastrointestinal bleeding is estimated to be between two in 1000 to
four in 100.\textsuperscript{13} However, the risk in young adults is probably less than in older
individuals who often have other diseases. Given all of the information,
the patient’s fears were allayed. The medication was proving quite effective
and despite the increased risk of an ulcer, she decided to continue the NSAID
while taking the appropriate precautions. She appreciates your thorough explanation.

2 Sackett DL, Haynes RB, Guyatt GH, Tugwell P. Clinical epidemiology: a basic science for clinical medicine, 2nd edn. Boston: Little, Brown and
4 Sanders JO, van Zanten V, Sherman PM. Helicobacter pylori infection as a cause of gastritis, duodenal ulcer, gastric cancer and nonulcer