Randomized controlled trials are viewed as providing gold standard evidence, and the declaratively named PRECISION trial (Prospective Evaluation of Celecoxib Integrated Safety versus Ibuprofen or Naproxen) implies the final word on the comparative cardiovascular safety of nonsteroidal anti-inflammatory drugs (NSAIDs).

The analgesic efficacy of NSAIDs derives largely from suppressing cyclooxygenase 2 (COX-2)–derived prostaglandins E₂ and I₂ (prostacyclin), whereas their gastrointestinal adverse effects result from the inhibition of cyclooxygenase 1 (COX-1)–derived gastroduodenal prostaglandins E₂ and I₂ and COX-1–derived thromboxane A₂ in platelets. This prompted the development of NSAIDs engineered to inhibit COX-2 specifically, such as rofecoxib, celecoxib, and valdecoxib. Randomized comparisons, such as VIGOR (Vioxx Gastrointestinal Outcomes Research), comparing rofecoxib and naproxen, showed fewer complicated gastrointestinal events with the COX-2 inhibitors.

The value of randomized trials was also evident when evidence, first from clinical pharmacology and then from experiments in model systems, predicted that cardiovascular events would complicate COX-2 inhibition of prostaglandin I₂ in the vasculature. Evidence consistent with this hypothesis emerged from VIGOR with more serious thromboembolic events with rofecoxib. Definitive evidence of cardiovascular hazard subsequently emerged from 10 placebo-controlled trials of structurally distinct COX-2 inhibitors. Last, overview analysis of individual data derived from ≈750 randomized trials provided a risk estimate of the magnitude of this hazard from COX-2 inhibitors with a rate ratio (RR) for serious vascular events of 1.37 (95% confidence interval, 1.14–1.66). As mechanistically predicted, the RRs for celecoxib (average daily dose, 400 mg) and rofecoxib (average daily dose, 25 mg) were superimposable. By contrast, information on older NSAIDs was fragmentary or nonexistent.

Coincident with the emergence of evidence that valdecoxib conferred a cardiovascular hazard (a finding delayed for months as the rofecoxib story unfolded), Pfizer announced in October 2004 plans to sponsor a placebo-controlled study of celecoxib in 4000 patients at high cardiovascular risk. Almost a year later, Pfizer announced that they would spend “at least $100 million” to conduct what became PRECISION, a comparison of celecoxib, ibuprofen, and naproxen in high-risk cardiovascular patients. The principal investigator predicted completion within 4 years. At the outset, the ethics and the interpretability of the trial prompted controversy. European Union countries declined to participate because of concerns about the safety of celecoxib. It was suggested that such concerns, together with the event-driven nature of the trial, would delay completion, perhaps until expiration of the patent on celecoxib. Furthermore, such high-risk patients would likely be taking low-dose aspirin, which targets platelet COX-1–derived thromboxane A₂. Both ibuprofen and naproxen interact to undermine sustained cardioprotection by aspirin; however, COX-2 is not extant in platelets, risking an intrinsic bias in favor of celecoxib.
Eleven years later, just after celecoxib, with sales >$2B annually, came off patent, the results of PRECISION were presented at the Scientific Sessions of the American Heart Association. The headline results are striking: no difference in the cardiovascular hazard from the 3 NSAIDs, no evidence supportive of an aspirin-NSAID interaction, and no evidence to suggest that naproxen is safer than the other 2 drugs. Fewer serious gastrointestinal events and renal adverse events were also noted on celecoxib-treated subjects. Given that this was a randomized trial in ≈24000 patients, has PRECISION delivered a precise outcome and should it alter practice? Unfortunately, the answer is no.

PRECISION is not a study of patients with arthritis who are at high cardiovascular risk. It mostly included patients with osteoarthritis who were at low cardiovascular risk; cardiac event rates were ≈1% per year. Yet the mechanism of cardiovascular hazard from NSAIDs is conditioned by the underlying cardiovascular risk substrate of the patient population. Furthermore, this is a noninferiority trial. Because of the low number of events accruing, the statistical upper bound was relaxed during the trial from 1.3 to 1.4 (with a power of only 80%). How likely is it that celecoxib would be found inferior? The RR for serious vascular events from celecoxib is 1.36.

A second series of concerns relate to whether pharmacoequivalent levels of drug exposure were attained. Clearly, less drug exposure means less efficacy, but also fewer cardiovascular and gastrointestinal adverse effects. Practitioners in PRECISION could increase the dose of ibuprofen and naproxen to attain efficacy. However, restraint on dose escalation of celecoxib, a regulatory response in many countries to the cardiovascular signal detected in 2 previous randomized controlled trials, limited the average daily dose to just >209 mg. In the overview of randomized controlled trials, this dose was indistinguishable from placebo; its RR for serious vascular events was 0.95, in contrast to RRs of 1.29 and 2.96 at total daily doses of 400 mg and 800 mg. There were no primary analgesic efficacy end points in PRECISION. However, visual analog scale reporting of efficacy, reports of arthralgia and osteoarthritis, and the number of patients with insufficient clinical response are all significantly worse on celecoxib, consistent with the possibility of reduced comparative efficacy. Indeed, the higher rates of hypertension and renal effects of the other 2 drugs are also consistent with lower drug exposure in the celecoxib group. It is interesting to note that Pfizer sponsored a concurrent comparison of switching to prescribed celecoxib or continuing on conventional NSAID therapy in a European population at low cardiovascular risk; here, interpretation of the trial was undermined by asymmetrical withdrawal from celecoxib (average ≈170 mg mean daily dose in a limited data set) because of the lack of efficacy.

A third major constraint to the interpretation of PRECISION is that, of ≈8000 patients randomly assigned to each treatment, ≈5000 had stopped taking their assigned therapy by the end of the study. Approximately 30% were lost to follow-up, and, of those who ceased taking their allocated treatment, a fraction recommenced taking some NSAIDs. All these observations intersect with the comments above to question the validity of the conclusions around noninferiority.

This trial was not designed to address differences in the likelihood of an NSAID interaction with low-dose aspirin as a source of bias in favor of celecoxib. The patients were not randomly assigned as to aspirin use, and there was no objective measurement of aspirin action. We do not know whether aspirin was taken as prescribed (in ≈45%) at outset and whether it was discontinued or started during the study, either by prescription or by patient access to this over-the-counter drug. Thus, it is unknown who took aspirin throughout the study and whether, if they did, cardiovascular events might have ensued in the ibuprofen and naproxen groups because of an interaction undermining the antiplatelet effects of the drug.

Similarly, the trial was not powered or designed to address the reported comparative cardiovascular safety of high-dose naproxen. Naproxen pharmacokinetics are highly variable, and an ill-defined proportion of patients have an extended half-life. Naproxen would be expected to confer cardioprotection comparable to the irreversible platelet inhibitor aspirin only in those individuals who take high doses or have a long naproxen half-life and are not already taking aspirin. As with the aspirin interaction, the absence of evidence is not evidence of absence.

In summary, there are so many problems with the interpretation of PRECISION that it fails to inform clinical practice. Thus, despite the enrollment of >24 000 patients and more than a decade of study, we are no closer to being able to advise the millions of patients with chronic arthritic pain regarding relative efficacy and safety of the treatments available to them. Such a disappointment indicates a need to move away from such blunt instruments as poorly designed trials to deep phenotyping studies that identify factors that predispose to benefit and risk at an individual level, thereby bring more precision to the use of NSAIDs.

**DISCLOSURES**

Dr FitzGerald has served as a consultant for Pfizer.

**AFFILIATION**

From Institute for Translational Medicine and Therapeutics, Perelman School of Medicine, University of Pennsylvania, Philadelphia.
FOOTNOTES
Circulation is available at http://circ.ahajournals.org.

REFERENCES


Imprecision: Limitations to Interpretation of a Large Randomized Clinical Trial
Garret A. FitzGerald

Circulation. 2017;135:113-115; originally published online November 13, 2016;
doi: 10.1161/CIRCULATIONAHA.116.026324
Circulation is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2016 American Heart Association, Inc. All rights reserved.
Print ISSN: 0009-7322. Online ISSN: 1524-4539

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://circ.ahajournals.org/content/135/2/113

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in Circulation can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the Permissions and Rights Question and Answer document.

Reprints: Information about reprints can be found online at:
http://www.lww.com/reprints

Subscriptions: Information about subscribing to Circulation is online at:
http://circ.ahajournals.org//subscriptions/