Don’t Just Do Something, Stand There!

Intervention Bias in Medical Decision Making

Andrew J. Foy, M.D.

Case Report

The nurse paged me at 2:00 a.m. to inform me that Mr. M’s cardiac output was 3.6 L/min and down from 3.8 to 4.4 L/min over the last 8 hours. His wedge and right heart pressures were in the same range as they had been trending. I went to the cardiac intensive care unit (ICU) to assess the patient. His heart rate and blood pressure were relatively unchanged. He already had 50 cc of urine out for the hour and had been between 30 to 50cc/hr out over the last 8 hours. His skin was warm, lungs were clear, and JVP was mildly elevated. His cardiac exam was unremarkable and the patient denied being in any distress. Overall, he looked good considering he had a major anterior ST segment elevation myocardial infarction complicated by hypotension less than 48 hours earlier. I informed the nurse that we would stay the course and make no changes. I was paged again an hour later. Mr. M’s cardiac output was 3.0 L/min, and everything else was basically the same. I again told the nurse we would stay the course and politely asked her not to collect any more numbers until the morning team arrived. I could tell she was not happy. I decided that discussing the evidence for pulmonary artery catheter (PAC) monitoring would have been futile at that hour of the morning, and I was too tired anyway.

Use of the Pulmonary Artery Catheter

The modern pulmonary artery catheter (PAC) was introduced by Swan and Ganz in 1970. No significant testing was performed to establish its clinical effectiveness prior to its widespread implementation. Gore et al. found that from 1975 to 1984 the percentage of patients hospitalized with an acute myocardial infarction (MI) who received a PAC nearly tripled from 7.2% to 19.9%. This trend has reversed in recent years; however, despite the overall decline in PAC use, the practice remains prevalent in cardiac ICUs. No evidence exists to support this practice, and it may be harmful.

Three large retrospective studies, totaling nearly 40,000 patients, have looked at the use of PACs in patients with acute MI. After adjusting for risk factors, all three found that PAC use was associated with increased mortality rates except in those cases complicated by cardiogenic shock, where it was associated with no difference. Schwann et al. performed a propensity-matched analysis of 2,546 patients undergoing on-pump coronary artery bypass surgery and found that the use of PAC was associated with increased mortality and a higher risk of severe end-organ complications. The Evaluation Study of Congestive Heart Failure and Pulmonary Artery Catheterization Effectiveness (ESCAPE) trial was a prospective study that randomized patients with severe congestive heart failure (CHF) to receive therapy guided by PAC or clinical assessment alone. Investigators found that in patients with severe CHF, PAC is not superior to clinical assessment alone and leads to more complications.

The Assessment of the Clinical Effectiveness of Pulmonary Artery Catheters in Management of Patients in Intensive Care (PAC-MAN) trial found that in 1,041 patients admitted to the ICU in whom the treating team thought a PAC would aid management, PAC did not lead to an improvement in in-hospital mortality or in any of the pre-defined secondary endpoints. Furthermore, there was no difference in any of the pre-specified sub-groups such as patients in cardiogenic shock.

The Pulmonary-Artery versus Central Venous Catheter to Guide Treatment of Acute Lung Injury (FAACT) trial demonstrated that in 1,001 patients with acute lung injury (ALI) the use of PAC to guide an explicit treatment protocol of liberal versus conservative fluid replacement did not lead to an improvement in mortality.

Sandham et al. found that in 1,994 high-risk surgical patients, the use of PAC with goal-directed therapy did not improve in-hospital mortality overall or in any of the pre-specified surgical subgroups.

Shah et al. conducted a meta-analysis of 13 trials that included 5,051 patients and concluded that in critically ill patients the use of a PAC did not confer any benefits and was associated with a higher use of inotropes and intravenous vasodilators.

In conclusion, there is no evidence from the published literature that PAC monitoring improves outcomes, and it may cause harm. Despite this lack of evidence many physicians still rely on PACs and many cardiology fellows’ sleep is disturbed by the pursuit of chasing Swan numbers. Why is this so?

The Case for Intervention Bias in Medical Decision Making

Bias is an inclination to present or hold a partial perspective at the expense of possibly equal or more valid alternatives. Intervention bias is an unrecognized problem that affects medical decision making. I use the term to describe the bias on the part of physicians and the medical community to intervene whether it is with drugs, diagnostic tests, procedural interventions, or surgeries when not intervening would be an acceptable choice. I have established a series of conditional arguments to support the existence of intervention bias in medical decision making. If intervention bias exists, then:
1. Physicians, when presented with the option to intervene or not, when not intervening is an acceptable choice, more often choose intervention.
2. Physicians will adopt interventions without rigorous experimentation.
3. Interventions will persist after their benefit has been disproven.
4. Physicians and medical scientists acting as investigators, manuscript reviewers, and journal editors will be more likely to submit or accept manuscripts for publication that have positive findings related to intervention and to ignore or reject negative studies.

Proof exists for all of these conditions; therefore, intervention bias exists in medical decision making.

Klingman et al. collaborated with leaders of three medical societies, the American College of Cardiology (ACC), the American College of Surgery (ACS), and the American Congress of Obstetrics and Gynecology (ACOG), to design and conduct surveys of their members using hypothetical clinical scenarios to find out how they would act in each case and why. Their goal was to evaluate how often physicians performed questionable interventions for defensive reasons. They found that defensive medicine did exist in their cohort, although not to the extent they expected. Eight percent of interventions were undertaken for defensive reasons. However, in all of the scenarios, the majority of physicians chose aggressive patient management styles even though conservative management was considered medically acceptable. In most of these cases, perceived medical indications, not malpractice concerns, motivated clinical choices. For example, 60% of cardiologists would get either an exercise EKG or stress thallium study on a healthy, active 42 year-old man with no risk factors for coronary artery disease (CAD) who presented to the ER with non-cardiac chest pain (i.e., pain on rotation of the left shoulder), a normal EKG, and negative cardiac enzymes. In another vignette-based study, Ayanian and Berwick found that pediatricians displayed a propensity toward action when faced with decisions to recommend tympanostomy tube placement or to order radiography in the ambulatory setting. Their methods and results closely mirrored those of a classic study conducted by the American Child Health Association in 1934, which found that school physicians were biased toward intervention when it came to recommending tonsillectomy.

There are many examples of interventions that have been adopted prematurely. Prominent cases include vertebroplasty; anti-arrhythmic medications to suppress ventricular ectopy post-MI; routine stenting for stable CAD; revascularization for atherosclerotic renal artery stenosis (ARAS); rhythm control for atrial fibrillation (AF); the pulmonary artery catheter (PAC); the intra-aortic balloon pump (IABP); and goal-directed blood pressure, lipid, and diabetic management.

Medication reversal is the term used by Prasad and Cifu to describe the process whereby a new clinical trial, superior to its predecessors, contradicts current clinical practice. It does not mean that for every indication and purpose the therapy in question was shown not to work, but that it was contradicted for key indications.

Prasad et al. examined a large collection of high-impact literature and found that among articles making a claim regarding a medical practice, 13% were medical reversals. Ionidis has shown that 16% of highly cited articles were contradicted by future studies. Furthermore, interventions often persist after their benefit has been disproven, as in the case of the PAC. The most important examples of medical reversal being disregarded are current guidelines for goal-directed blood pressure, lipid, and diabetic management. In all these instances, data from randomized trials indicate that while treatment may be better than no treatment in certain cases, targeting a specific level is not more beneficial and may in fact be harmful. However, based on recommendations from professional guidelines, treatment to these targets is required to meet quality-of-care standards.

Finally, there is strong evidence that positive findings related to intervention are more likely to be submitted and accepted for publication than those that are negative. This is known as positive-outcome bias. In one landmark study Mahoney randomly assigned 75 journal reviewers to receive one of five similar manuscripts. All manuscripts were identical in the “Introduction” and “Methods” sections but varied in either the “Results” or “Discussion” sections. Although studies with positive and negative results had identical “Methods” sections, negative studies received significantly lower scores for publication merit.

Emerson et al. used methods similar to Mahoney and found that 210 reviewers for orthopedic journals were significantly more likely to recommend a fabricated manuscript for publication regarding the use of postoperative antibiotics if it had a positive outcome compared to an identical manuscript that found no difference. The reviewers also identified significantly more errors in the manuscript with no difference. Turner et al. evaluated 74 FDA-registered studies of 12 antidepressant agents involving 12,564 patients. A total of 37 studies viewed by the FDA as having positive results were published; one study viewed as positive was not published. Studies viewed by the FDA as having negative or questionable results were, with three exceptions, either not published (22 studies) or published in a way that, in the authors’ opinion, conveyed a positive outcome (11 studies).

Causes of Intervention Bias

Why does intervention bias exist in medicine? It is likely the manifestation of two well recognized forms of bias, self-interest bias and confirmation bias. Theories of political and economic science view self-interest as the ultimate goal of many aspects of human behavior. It also appears that self-interest plays a strong role in attitude judgment and persuasion.

Through a series of experiments Darke and Chaiken showed that self-interest biases attitude judgment in a directional manner. Because intervention is often in the self-interest of
physicians, clinical researchers, and the health care industry from a financial perspective it could bias them to more easily accept arguments in favor of intervention and less inclined to accept those that go against it. Prasad and Cifu state that “financial incentives are strongly aligned to promote new technologies…. Conflicts of interest among trialists, industry-sponsored studies, and industry-sponsored economic analyses all encourage wrongful optimism, facilitating approval.”

Besides being in their financial self-interest, the act of intervention could serve physicians’ self-interest in another way. According to Ayanian and Berwick, “clinical satisfaction may be greater for doctors when they recommend an intervention, giving a sense of greater activism in their patients’ care.” For many physicians, the ability to intervene is tied directly to job satisfaction and personal fulfillment. Those who subspecialize often do so based on their affinity for the nature of the interventions involved. Therefore, their ability to render judgments about the appropriateness of intervention would be affected by self-interest bias.

Confirmation bias is the tendency of people to favor information that confirms their beliefs or hypotheses. It is harmful to objective evaluation, which is required as part of the scientific method. One explanation offered for medical reversal by Prasad and Cifu was “an unjustified confidence [hubris] in basic science models and surrogate outcomes” that can be explained by confirmation bias. Confirmation bias may also explain why individual physicians favor interventions based on personal experience and belief even after their benefits have been disproven. An experimenter’s confirmation bias could affect which data are reported. Data that conflict with the experimenter’s expectations may be more readily discarded as unreliable, producing the so-called “file drawer effect.” The finding by Turner et al. that 60% of negative trials registered with the FDA were not reported confirms this. Confirmation bias can also explain why data that conflicts with reviewers’ expectations would more likely be dismissed, as Mahoney’s and Emerson’s studies suggest.

Fear of malpractice lawsuits is also a likely contributor to intervention bias. According to Kowey, “defensive medicine is pervasive and takes many forms.” “It extends from ordering too many tests all the way to performing unnecessary surgical procedures.” One mail survey of physicians in six high-risk specialties in Pennsylvania found that nearly all (93%) reported practicing defensive medicine and “assurance behavior,” such as ordering tests and performing diagnostic procedures, was very common (92%). A national survey administered by the American Medical Association (AMA) found that an overwhelming majority of respondents (91%) reported believing that physicians order more tests and procedures than needed to protect themselves from malpractice suits. These survey results suggest that fear of malpractice lawsuits or the charge of malpractice may be the overwhelming cause of intervention bias; however, results from Klingman’s landmark study contradict this.

In the study by Klingman et al. only 8% of interventions were undertaken for defensive reasons, but in all of the scenarios the majority of physicians chose aggressive patient management styles even though conservative management was considered medically acceptable by the expert panels. In most of these cases, perceived medical indications, not malpractice concerns, motivated clinical choices. These results do indeed confirm the existence of intervention bias but indicate that self-interest and confirmation bias probably play a more important role than malpractice-related concerns. Also, fear of malpractice would play a minor role if any in influencing professional guidelines that have been responsible for codifying overtreatment in certain cases like goal-directed blood pressure and glucose management.

Moral hazard due to third-party payment for health care services also likely contributes to intervention bias. Traditional health insurance reimburses as a function of expenditure or use. Because insurance drives the marginal price of medical care at the point of use to near zero, consumers—or physicians acting as their agents—demand care until the marginal price of additional care is nearly zero. Studies have found that a fully insured population spends about 40% to 50% more than a population with a large deductible and that their status is not measurably improved by the additional services. Oboler et al. found that the majority of patients expect more care than is prudent to deliver. Sixty-six percent of respondents believed that in addition to regular care, an annual physical examination is necessary. Many tests, including Papanicolaou smear (75%), mammography (71%), cholesterol measurement (65%), prostate-specific antigen test (65%), urinalysis (40%), blood glucose measurement (41%), fecal occult blood testing (39%), and chest radiography (36%), were desired. Interest in these tests decreased substantially when the charges were known.

The problem of moral hazard is likely compounded by the use of patient satisfaction surveys, which are being widely used as health care quality metrics. Fenton et al. found that in a national survey of 51,946 adults conducted between 2000 and 2007, higher patient satisfaction was associated with greater inpatient use, higher overall health care and prescription drug expenditures, and increased mortality. Studies have shown that physicians often give in to whatever patients want, whether it is medically necessary or not. Wilson et al. found that patients’ perceived need for radiological studies was significantly associated with use of those services for outpatients with respiratory problems and low back pain. In another study, 36% of physicians told researchers they would yield to a patient who asks for a clinically unwarranted magnetic resonance imaging exam.

Consequences of Intervention Bias

Intervention bias has several important consequences. For one, informed decision making relies on the validity of unbiased, balanced, and objective data from published studies, independent of the reported outcome. This is corrupted by intervention bias,
rendering clinical recommendations flawed towards specific intervention strategies. Next, intervention bias can lead medical professionals to violate the principle of “primum non nocere” or “first do no harm.” This can occur whenever interventions are undertaken without rigorous experimentation and then persist after their benefits have been disproven. Goal-directed blood pressure management calls for certain high-risk groups to achieve a blood pressure level of less than 130/80 mm Hg. The most definitive blood pressure targeting trial was the Action to Control Cardiovascular Risks in Diabetes (ACCORD) study. At 4.7 years, there was no difference in the primary end point of nonfatal MI, nonfatal stroke, or cardiovascular death, despite achieving a significant difference in mean systolic blood pressure after the first year (119.3 vs. 133.5 mm Hg). There was a significant increase in serious adverse events in the intensive-therapy group (3.3% vs. 1.3%, P < .001). Therefore, in regard to intensive blood pressure management based on results from the ACCORD trial, the number of patients needed to treat to provide a therapeutic benefit is theoretically infinite, but only 50 patients need to be treated to harm one.

Intervention bias also contributes to the disproportionate increase in medical spending in the U.S. Since 1970, U.S. health care spending per capita has been more than double the real growth in GDP per capita (4.3% vs. 2.0%). Over that same period, countries belonging to the Organization for Economic Cooperation and Development (OECD) averaged an annual growth rate of 3.8% in health care spending per capita compared to only a 2.1% annual growth in GDP per capita. Eight of 20 countries had higher average annual growth rates in health care spending per capita than the U.S. A portion of this growth is due to the adoption and overutilization of technologies that are either futile or confer only minor clinical benefits. Between 1987 and 2000, spending on heart disease increased by more than $26 billion, 69% of which was attributable to increased cost per treated case. During that time, the rate of stenting increased 128% despite repeated negative trials involving the use of this modality for its most often cited indication.

The Role of Medical Judgment

In today’s political climate, with pressures to limit treatment with the goal of cost containment, the opposite problem of nonintervention bias could present a danger in the future. While guarding against overuse of treatments shown to be of limited value in rigorous trials, the physician caring for an individual patient must bear in mind that subgroups of patients may benefit, and that his patient may be substantially different from the majority of study patients. Additionally, insistence on proof of value in a rigorous trial may deprive his patients of potentially lifesaving treatments that cannot by their nature be tested in a randomized controlled trial, e.g. devices for assisting in cardiopulmonary resuscitation, or that simply have not yet been tested. Proper use of evidence-based medicine requires appreciation of its limitations. The physician’s knowledge of his individual patient, his understanding of basic physiologic and pharmacologic principles, and his expert clinical intuition must always be respected.

In every case, the physician’s concern is the response of one patient, and therapy must be tailored to that patient to optimize outcome for that patient, not necessarily for the endpoints chosen by researchers.

Conclusion

Intervention bias often corrupts the informed decision making process and leads physicians to adopt interventions prematurely and to continue using them after their benefits have been disproven. This subjects patients to unnecessary risk and also increases nonproductive costs. Sources include self-interest bias, confirmation bias, publication bias, the moral hazard of third-party payment, and fears of professional liability.

Recognition is the first step toward overcoming bias. To guard against it, we should always remain skeptical, insist that trials report clinically meaningful endpoints, and be unafraid to protest the widespread adoption of “practice guidelines” or “best practices” that ignore intervention bias or undermine the ability of physicians to use their best judgment in caring for individual patients.

Andrew Foy, M.D., is a fellow in cardiology, Penn State Hershey Medical Center, 500 University Dr., P.O. Box 850, Hershey, PA 17033. Telephone (302) 750-1108; email: Andrew.foy@gmail.com.

REFERENCES

33 White C. Health care spending growth: How different is the United States from the rest of the OECD? Health Aff 2007;26:154-161.