CASCADE EFFECTS OF MEDICAL TECHNOLOGY

Richard A. Deyo
University of Washington, Center for Costs and Outcomes Research, 146 North Canal Street, Suite 300, Seattle, Washington 98103-8652; e-mail: Deyo@u.washington.edu

Key Words  technology, adverse effects, quality-of-care, diagnostic tests

Abstract  Cascade effect refers to a process that proceeds in stepwise fashion from an initiating event to a seemingly inevitable conclusion. With regard to medical technology, the term refers to a chain of events initiated by an unnecessary test, an unexpected result, or patient or physician anxiety, which results in ill-advised tests or treatments that may cause avoidable adverse effects and/or morbidity. Examples include discovery of endocrine incidentalomas on head and body scans; irrelevant abnormalities on spinal imaging; tampering with random fluctuations in clinical measures; and unwanted aggressive care at the end of life. Common triggers include failing to understand the likelihood of false-positive results; errors in data interpretation; overestimating benefits or underestimating risks; and low tolerance of ambiguity. Excess capacity and perverse financial incentives may contribute to cascade effects as well. Preventing cascade effects may require better education of physicians and patients; research on the natural history of mild diagnostic abnormalities; achieving optimal capacity in health care systems; and awareness that more is not the same as better.

CASCADE EFFECTS OF MEDICAL TECHNOLOGY

Health professionals and laypersons alike tend to equate new medical technology with better-quality health care, assuming that newer is better. Much of the scientific literature on diffusion of innovations focuses on the anticipated beneficial effects of new technology and methods to ensure its rapid adoption (53). However, many new medical technologies are introduced and disseminated with only modest evaluation of efficacy, optimal indications, or impact on practice. Unfortunately, their use in routine care sometimes proves futile or even harmful. The adverse effects and consequences of new technology are often unanticipated (53). One of these is the cascade effect.

The term cascade effect, in reference to medical technology, was apparently coined by Mold & Stein in a 1986 article in the New England Journal of Medicine (44). In biology, the term cascade refers to a process that proceeds in stepwise fashion from an initiating event to a seemingly inevitable conclusion. A molecular example is the blood clotting cascade, typically initiated by a cut in the skin. The
disruption of capillary blood vessels prompts aggregation of platelets in the blood to form an initial plug, which in turn triggers a cascade of protein interactions, ultimately resulting in the formation of a firm blood clot. Mold & Stein argued that health care may be subject to similar cascade effects, in which an initiating factor is followed by a series of events with increasing momentum, so that the further events progress, the more difficult they are to stop. These events often include unnecessary tests, procedures, and risks for patients. The initiating factor is often physician anxiety, which may result, for example, in a diagnostic test, an unexpected result, and a chain of subsequent events that are ultimately to the patient’s disadvantage.

Mold & Stein offered the story of a patient admitted to the hospital for elective repair of an inguinal hernia. He had a history of coronary disease with very mild arterial narrowing on a previous cardiac catheterization. Anxious about his cardiac status, the surgeons requested a preoperative cardiology consultation. Perhaps uncertain about his own clinical judgment, the cardiologist suggested obtaining an exercise tolerance test. This was delayed for six hours while the patient waited outside the test room, during which time he became anxious, agitated, and angry, and had some mild chest discomfort. Because of the chest discomfort, the test was not done and the patient was transferred to a telemetry unit. There he became more anxious and agitated, was found to have some electrocardiogram changes, and received medications. He underwent another cardiac catheterization, which actually showed slight improvement since his previous test. At that point, the hernia repair could not be performed because of a full operating room schedule, and the primary physician was left to try to reassure the patient that he was in no danger. The procedure had to be delayed for two weeks. In this example, the chain of events seemed to be fueled by physician anxiety, and it snowballed with the addition of patient anxiety (44).

Physicians who are anxious about a patient’s problem may be tempted to do nearly anything in order to reduce their own anxiety. The first step typically appears to be a benign action, such as ordering a diagnostic test; however, the discovery of an unexpected abnormality leads to progressively riskier and costlier interventions that seem simultaneously unnecessary and unavoidable. Ober likened this situation to the story of Br’er Rabbit and the tar baby. In this classic Uncle Remus tale, Br’er Fox creates a tar baby to trap Br’er Rabbit. Br’er Rabbit greets the tar baby and takes its failure to communicate as snobbery. He hits the tar baby to teach it a lesson and gets stuck. In an effort to make the tar baby let go, he hits it with the other hand, then kicks it. With each blow he becomes more ensnared, rather than closer to a solution (45). Similarly, as physicians try to explain each new ambiguous result, they become ensnared by their own actions.

Mold & Stein pointed out that in clinical care, cascade effects could be triggered inappropriately by incomplete data gathering; overinterpretation of an abnormal lab result; underestimation of the risks of a test or treatment; underestimation of the possibility of false-positive results; and intolerance of ambiguity by the physician (44). Examples of some of these problems follow.
EXAMPLES OF COMMON CASCADES RESULTING FROM DIAGNOSTIC TECHNOLOGY

The problem of cascade effects may be best understood by considering some common examples (Table 1). Some of these are now reasonably well documented, though others undoubtedly remain to be described.

Endocrine Incidentalomas

Benign tumors of the adrenal glands (adenomas) are quite common and have been reported in almost 9% of patients in autopsy. Adrenal carcinomas are extremely rare, in contrast, with an incidence of 0.0004% per year (45). Unfortunately, the radiographic appearance of carcinomas and benign adenomas is generally indistinguishable. Many incidental adrenal masses are discovered in patients for whom imaging such as computed tomography is performed for vague or nonspecific abdominal complaints (underestimating the possibility of false positive results) (12). Once an incidentaloma is discovered, both patient and physician may have such a level of anxiety that additional testing becomes inevitable. This was one of the first described problems of incidental imaging findings and has spawned a growing literature and recent reviews on the management of adrenal incidentalomas (1, 2, 11, 12). A recent recommendation for follow-up for these incidentalomas includes multiple blood tests, urinary hormone assays, and repeat imaging tests within six months of discovery and annually thereafter (2). Thus, incidental findings can precipitate lifelong cascades of clinical events.

Similarly, benign adenomas of the pituitary gland are common incidental findings. As with adrenal tumors, incidental findings are identified in up to 10% of patients having MRI of the head, and autopsy studies reveal a prevalence of 1.5% to 27% for small pituitary adenomas unidentified during life. In contrast, pituitary lesions that are symptomatic as a result of either hormone excess or mass effect are rare. Their prevalence is about 20 cases per 100,000 persons, and the incidence is approximately two per 100,000 per year. Fortunately, current recommendations for follow-up of these incidentalomas include only a simple screening blood

<table>
<thead>
<tr>
<th>Initial test</th>
<th>Potential consequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Head, body scans</td>
<td>Endocrine incidentalomas</td>
</tr>
<tr>
<td>Electronic fetal monitoring</td>
<td>Unnecessary Cesarean sections</td>
</tr>
<tr>
<td>Coronary angiography in low-risk patients</td>
<td>Unnecessary invasive coronary interventions</td>
</tr>
<tr>
<td>Spinal MRI in the absence of sciatica</td>
<td>Unnecessary spine surgery</td>
</tr>
<tr>
<td>Pulmonary artery catheters</td>
<td>Tampering with cardiac physiology</td>
</tr>
<tr>
<td>Persistent testing in persons near end of life</td>
<td>Unwanted aggressive interventions</td>
</tr>
</tbody>
</table>
test for abnormal hormone secretion among patients who are truly asymptomatic (1, 11).

Electronic Fetal Monitoring

Electronic fetal monitoring for women in labor is another technology that may lead to cascade effects. These devices monitor fetal heart rate, and certain patterns are associated with a greater likelihood of fetal distress. However, the risks of the test were not well considered before its adoption into routine care. Use of such monitors requires the mother to be relatively inactive in bed and may increase anxiety levels. The combination of inactivity and anxiety may slow labor and lead to interventions to speed up labor (e.g., by artificial rupture of the membranes). When labor is accelerated, the pain of contractions increases and pain medication or epidural anesthesia may be requested by the patient. The loss of amniotic fluid may lead to higher pressures inside the baby’s skull, which could lead to more abnormal readings on the fetal heart rate monitor. Use of epidural anesthesia may lower maternal blood pressure, similarly leading to more abnormal readings. Perhaps as a consequence of such events, Cesarean section rates are 40% higher when electronic monitoring is used rather than simple auscultation of the fetal heart rate. Unfortunately, randomized trials have shown that fetal monitoring does not improve overall fetal outcomes except for a possible small reduction in neonatal seizures (56, 60). Thus, much of the intervention and cost associated with its use has little benefit and does pose potential risks.

Testing and Treatment for Coronary Disease Among Low-Risk Patients

Twenty years ago, Graboys considered the likely consequences of screening for occult heart disease among prospective runners over age 35. Such screening by means of exercise stress testing was a common recommendation at the time. Approximately 10% would be expected to have positive results. Because asymptomatic persons have a low probability of disease and because exercise electrocardiography has a substantial false-positive rate, only one fourth of patients with positive tests would actually have multivessel coronary artery narrowing. Assuming low rates of cardiac catheterization complications and operative mortality, as well as low rates of perioperative heart attacks, Graboys estimated the following consequences of the routine screening strategy. If 20 million persons were screened, 2 million would have positive exercise studies. If all of these underwent coronary catheterization, approximately 2000 would die. Of the half million persons who would then undergo bypass surgery, 10,000 would die, and there would be 40,000 new heart attacks. A conservative estimate of the cost in 1979 was $13 billion (24). This analysis describes a typical cascade effect that may result from an ill-advised strategy of routine testing in otherwise healthy adults.

A more recent study examined the use of cardiac catheterization among Medicare patients following heart attacks. Experience was compared between New
York, with a low rate of testing, and Texas, where the rates of post–heart attack catheterization were 50% higher. Patients in the two states appeared to have approximately equal severity of coronary artery disease, but Texans with mild disease were much more likely to undergo angiography and to receive revascularization procedures. Paradoxically, after two years of follow-up, the Texas patients had lower exercise tolerance, more angina, and higher overall mortality (26).

A randomized trial compared the use of coronary angioplasty versus medical therapy for patients with angina. Most patients had relatively mild coronary artery disease. After two years, angioplasty had reduced symptoms in the group with severe angina, but the risk of a heart attack or death was almost twice that in the medical therapy group (52). There are now four randomized trials that are remarkably consistent in showing that routine cardiac catheterization and revascularization does not reduce the likelihood of repeat heart attacks or death compared with a more conservative approach following acute coronary syndromes (unstable angina and acute heart attacks) (38). In a study of patients with relatively small heart attacks, an aggressive strategy was actually associated with higher mortality both during hospitalization and at one-year follow-up (9). An editorialist noted that there is a strong relationship between the availability of angiography in a geographic area and the likelihood that the more aggressive diagnostic and treatment strategy would be chosen. Factors encouraging the use of aggressive therapy include the abundance of facilities for cardiac catheterization and revascularization in the United States, large numbers of physicians trained to perform the procedures, and generous insurance reimbursement (38).

Spinal MRI

Spinal MRI exemplifies the problem of discovering more and more abnormalities with most having no clinical relevance (15). There are now many studies of asymptomatic patients demonstrating that herniated discs, degenerative discs, and bulging discs are frequent incidental findings (8, 32, 33). Clinically, such findings lead to overdiagnosis, anxiety on the part of patients, and conviction about the presence of disease. Some authors have suggested that clinically irrelevant findings on MRI may lead to unnecessary back surgery with its occasional complications (8). An Institute of Medicine study concluded that lumbar spine surgery is overused and misused in the United States (46), and the wide use of imaging studies may be a driver of this excess use.

In a study of geographic variations in back surgery rates, Keller et al. found that where surgery rates were highest in the state of Maine, patients had the mildest preoperative disease and the worst postoperative outcomes. Conversely, where surgery rates were the lowest, patients had the most severe preoperative disease and the best postoperative outcomes (36). These findings support the notion that some back surgery is unnecessary and may lead to poor patient outcomes. We may speculate that some of the unnecessary surgery is driven by irrelevant imaging results.
Other Unnecessary Surgery

Back surgery may be just one example of a larger problem of unnecessary surgical intervention. International comparisons suggest that the United States has a much higher rate of many forms of surgery than other developed countries. In an early study of surgical second opinion programs, 17.6% of recommendations for surgery were not confirmed. Other studies have used explicit criteria to identify indications for surgery and have demonstrated high rates of unnecessary surgery for procedures such as coronary bypass operations, hysterectomy, pacemaker insertion, and tonsillectomy. Extrapolating from the second opinion study, a Congressional subcommittee concluded that perhaps 2.4 million unnecessary operations are performed annually in the United States, at a cost of $3.9 billion and 11,900 deaths. The number of operative complications would be substantially higher (39).

Adverse outcomes due to poor surgical technique have long been a target of quality review efforts, but Roos and colleagues proposed similar attention to adverse outcomes produced simply by high rates of interventions. In a review of coronary artery surgery rates, they focused on communities with unusually high surgery rates and high mortality rates. At least as many deaths could have been prevented by reducing surgical rates to the U.S. average as by improving the technical quality of surgery (54).

For many common operations, indications remain controversial, and diagnostic test results may often drive ill-advised surgery. Leape notes that uncertainty stemming from a lack of consensus about surgical indications may lead surgeons to recommend operations for patients who wish to have them but from which they will not benefit. From a patient perspective, the combination of risk and the potential for dramatic cure gives surgery an aura of excitement lacking in other forms of therapy (39).

Pulmonary Artery Catheterization

Pulmonary artery catheters are inserted into peripheral veins, floated through the right side of the heart, and into the pulmonary artery. In this position, the catheter allows measurement of pressures of the central veins and the pulmonary artery, mixed venous blood gases, cardiac output, and pulmonary capillary wedge pressure (reflecting pressures on the left side of the heart). These devices are used in critically ill patients to assess the need for intravenous fluids or for treatments to improve cardiac output. Approximately 1.5 million catheters are sold in the United States annually, and these are most frequently used in cardiac surgery, cardiac catheterization laboratories, coronary care units, medical intensive care, and high risk surgery and trauma (4).

When the catheter was introduced in 1970, there were no prospective trials to evaluate its clinical impact. In 1996, an observational study suggested that the use of pulmonary artery catheters may not only be unhelpful in patient care, but may on balance do more harm than good (27). Negative consequences of the pulmonary catheter can result from serious complications (e.g., cardiac arrhythmias,
thromboembolism, and sepsis), but also from operational problems, errors in data interpretation, or exaggerated or inappropriate treatment responses to catheter data. The observational studies have led to randomized trials that are currently under way for some indications, but definitive data are not yet available.

Although it remains uncertain whether pulmonary artery catheterization is, on average, truly harmful, the potential for it to cause harm is based partly on the possible cascade effects of measuring numerous physiologic parameters, then tampering to try to optimize these parameters. This device may also illustrate the problem of judging a technology’s benefits from surrogate outcomes such as cardiac output. Such physiologic measures may be optimized by the technology, even though the end results (patient survival) are unaffected or adversely affected.

Unwanted Aggressive Care at the End of Life

This may be one of the most familiar cascade effects that has adverse consequences for patients. Many are familiar with the elderly parent who wants to die with dignity but gets drawn into a series of well-intentioned medical interventions, only to die in an intensive care unit with intravascular catheters, artificial ventilation, multiple medications to support circulation or treat infection, and perhaps even renal dialysis to support failing kidneys. It may only be in retrospect that the family and physicians recognize that the patient did not want interventions but was drawn into them by frequent diagnostic testing, monitoring of blood parameters, and efforts to identify reversible disease.

Physicians may fail to appreciate that because older patients have many competing risks for death, the absolute effect of a new diagnosis on life expectancy may be very small (64). Thus, the potential gain in survival even from perfect therapy may be small. Furthermore, risks of any therapy generally increase with age, so the high burden of competing risks and high rates of treatment complications may reduce the net benefit of many treatments.

COMMON TRIGGERS FOR THE CASCADE EFFECT

Because of the near inevitability of certain cascades once they are initiated, the best chance of preventing unnecessary adverse consequences may be to prevent the triggering event. A wide variety of likely triggers have been identified, relating to both psychological and cognitive factors, as well as cultural attitudes and perverse incentives.

Shotgun Testing

A nearly ubiquitous feature of modern medical practice is the panel of laboratory tests ordered as a cluster. For many biochemical tests of blood or urine, the normal range is simply defined as two standard deviations from the mean of a healthy population. By definition, therefore, about 5% of results on each test from normal
TABLE 2  Probability that a healthy person will have abnormal results in a biochemical profile

<table>
<thead>
<tr>
<th>Number of tests</th>
<th>Probability of at least one abnormal test, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>6</td>
<td>26</td>
</tr>
<tr>
<td>12</td>
<td>46</td>
</tr>
<tr>
<td>20</td>
<td>64</td>
</tr>
<tr>
<td>100</td>
<td>99.4</td>
</tr>
</tbody>
</table>

a Adapted from references 10 and 66.
b Assuming that each test in the battery is independent of the others.

persons will be mislabeled as abnormal. Table 2 shows the probability that at least one test will be abnormal for various numbers of tests performed. For a panel of 12 tests (common in modern practice), there is a 46% chance that at least one test will be abnormal even in completely healthy persons (10, 66). This phenomenon has led to the cynical saying that “the only normal person is someone who hasn’t had enough tests.” Unfortunately, both patients and naive physicians may fail to appreciate the statistical basis for the normal ranges of chemical tests, underestimate the likelihood of false-positive results, and fail to recognize that many abnormalities will not represent disease (50).

Underestimating the Likelihood of False-Positive Results

This is closely related to the problem of shotgun testing. Many clinicians are unfamiliar with the concept of positive predictive value and fail to appreciate the high probability that positive results among patients with a low probability of disease are likely to be false positives. This occurs most frequently when a physician searches for an uncommon condition.

As a resident, one of my colleagues embarked on an evaluation for acute intermittent porphyria (AIP) in a patient with abdominal pain. Acute intermittent porphyria is a very rare disease, for which common screening tests have only moderate sensitivity and specificity. The prevalence of AIP is estimated to be one to two per 100,000 in Europe, where it is more common than in the United States (35). A widely used screening test for emergency situations is the Watson-Schwartz test for an abnormal urinary excretion product in these patients who have a metabolic abnormality. The screening test is estimated to be about 32% sensitive and 82% specific (13). Table 3 shows the consequences of screening a large number of patients with abdominal pain using the Watson-Schwartz test. The moderate specificity means that out of 9990 patients who do not have AIP, 8192 will have a negative test. Unfortunately, there are 1798 who will have a positive test but no disease. Of the ten patients who have AIP, four will have a positive test. Thus, the predictive value of a positive test, which is the probability of AIP given abdominal pain and a positive Watson-Schwartz test, is 0.002, or 2/10 of
TABLE 3  Positive predictive value of a screening test for Acute Intermittent Porphyria (AIP), a rare metabolic abnormality, among patients with abdominal pain. The Watson-Schwartz test has a sensitivity of 38% and specificity of 82%. This table assumes that the prevalence of AIP among patients with abdominal pain is 10/10,000, much higher than the general population prevalence of 2/100,000

<table>
<thead>
<tr>
<th>Actual diagnosis</th>
<th>AIP</th>
<th>Not AIP</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test Result</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive</td>
<td>4</td>
<td>1,798</td>
<td>1,802</td>
</tr>
<tr>
<td>Negative</td>
<td>6</td>
<td>8,192</td>
<td>8,198</td>
</tr>
<tr>
<td>Total</td>
<td>10</td>
<td>9,990</td>
<td>10,000</td>
</tr>
</tbody>
</table>

Probability of AIP given abdominal pain and a positive Watson-Schwartz test = 4/1802 = .002 (2 tenths of 1%).

1%. When my colleague discovered a positive Watson-Schwartz test, however, he was off and running with further tests. Certain that the patient had this rare condition, he embarked on a lengthy and expensive diagnostic evaluation that yielded some additional ambiguous results until more definitive testing proved repeatedly negative. Unfortunately, this was a predictable series of events.

Inappropriate Screening

Well-meaning clinicians are sometimes prompted to undertake screening tests in asymptomatic patients whom they perceive to have a high risk of a serious disease. A common example is the smoker for whom a physician has concerns about lung cancer. Although many physicians have been tempted to obtain periodic chest X-rays, even as frequently as every six months, randomized trials have failed to show any benefit for such aggressive lung cancer screening. In one trial, five-year survival was 23% for lung cancers diagnosed in the screening group and 0% for those diagnosed in the control group (6). However, the improvement was entirely attributable to lead time and length biases because overall mortality from lung cancer was actually higher in the screened group, which indicated that the aggressive screening and subsequent intervention was neutral or even harmful. Similar concerns may be raised regarding other common practices, such as the use of EKG to screen for heart disease or exercise testing to screen potential runners for heart disease (described above).

Errors in Data Interpretation

Misinterpretation of a diagnostic test may lead to unnecessary interventions and their potential complications. As one example, coronary angiograms (from cardiac
catheterization) from 308 randomly selected patients were reviewed by a blinded panel of three experienced radiologists and compared with the original interpretations. Technical deficiencies were found in over half the cases. The panel readings showed less significant disease, less severe stenosis, and lower extent of disease than the original interpretations. The classification of subsequent coronary bypass surgery changed from necessary/appropriate to uncertain/inappropriate for 17% to 33% of cases when the panel readings were used. Using an expert set of appropriateness criteria, it appeared that some 17% of coronary bypass procedures and 10% of angioplasty procedures recommended on the basis of these films would have been inappropriate (40).

Overestimating Benefits or Underestimating Risks

An example of this problem comes from consideration of carotid endarterectomy. Clinical trials on the efficacy of carotid endarterectomy have shown that the benefit of surgery must be carefully weighed against complication rates in order to judge whether operative benefit outweighs the risk. In a study of academic medical centers, an expert panel rated the indications for performing endarterectomy, and the charts of 1160 randomly selected endarterectomy patients were abstracted. The expert panel defined acceptable operative risk for the patients under consideration to be less than 3% for death, stroke, or heart attack. Using that definition and the actual hospital mortality rates, only 33% of the 1160 procedures could be classified as appropriate. Even when a more liberal definition of 5% risk for bad outcomes was used, only 58% of the procedures would have been judged appropriate (42). Thus, clinicians recommending carotid endarterectomy in many of these institutions must have overestimated benefits and underestimated risks in making their decisions.

In assessing criteria for appropriateness, physicians’ opinions vary. In a study of six different surgical procedures, 45 panelists rated the appropriateness of various indications both independently and after a two-day discussion. Performers of the surgical procedures had the lowest threshold for appropriateness, followed by physicians in related specialties, and trailed by primary care providers. Approximately one fifth of actual procedures were for indications rated less than appropriate by primary care physicians, but appropriate by performers of the procedures (34).

The problem of testing and treatment in elderly patients has previously been discussed. Because of competing risks from other comorbid conditions, the value of identifying and treating a new condition may be substantially less in older patients than in younger patients. Furthermore, the side effects of most procedures are higher in elderly patients than in younger patients, so the result is sometimes a dramatically different benefit-to-risk ratio than in younger adults (64).

In some cases, diagnostic tests themselves carry iatrogenic risks, regardless of subsequent treatments. Perforation of the colon during colonoscopy or fetal injury during amniocentesis are just two examples (66).
Defensive Medicine

Physicians sometimes request unnecessary tests or treatments in order to avoid medicolegal liability for a missed diagnosis or treatment opportunity. As the examples here suggest, however, poorly thought-out testing may sometimes lead to more patient harm than good, paradoxically increasing medicolegal risks.

Formal decision analysis produces the somewhat surprising result that defensive medicine necessarily reduces the overall quality of patient care. This finding contradicts arguments that defensive testing may further the interests of both doctor and patient. The reason is related to the problem of false-positive and false-negative tests. Decision analysis allows calculation of an optimal testing threshold: a pretest likelihood of disease that makes testing the preferred strategy over simply treating without a test or not treating at all. This threshold depends in part on the true- and false-positive rates of the test, and the consequences of those errors. If a physician widens the range of possibilities over which he or she prefers testing in order to reduce liability, then some patients who would be better left untreated will instead be tested and treated if the test is positive. Similarly, some patients who should be treated will instead be tested, and treatment withheld if the test is negative.

The argument is theoretical, and actual practice may rarely conform to the optimal strategy in any event, for a number of reasons. Nonetheless, the analysis suggests that defensive medicine is not merely a problem of increased cost, but also one of reduced quality-of-care (14).

Patient Demand

Commenting on excessive use of coronary catheterization and revascularization, Lange & Hillis suggest that some patients and families may insist on aggressive management in the face of a heart attack. The term conservative management may convey an aura of obsolescence, inadequacy, or inferiority. In an era of managed care, it may even have the connotation of saving money at the expense of quality-of-care. Patients and families may paradoxically be more understanding and forgiving if an aggressive approach is pursued with a bad outcome, even if the management approach contributed to the adverse results (38).

In a study of women’s understanding of the mammography screening debate (whether to begin screening at age 40 or age 50), 83% of women believed that mammography had proven benefit for women aged 40 to 49, and 38% believed that its benefit was proven even for women younger than 40 years. In response to an open-ended question about why mammography has been controversial, the most common response was that the debate was about saving money, rather than a question of benefit (65). Also, most women were unaware that screening can detect cancers that never progress (55). Such misunderstandings probably contribute to patient demand for screening and diagnostic tests. In some studies, inappropriate diagnostic tests were most likely to be done among patients with the strongest perception of need (18).
Patients seem willing to pay for diagnostic certainty even when it has no clinical benefit. For example, in one survey, patients were asked about willingness to pay for a test for peptic ulcer disease that would be risk-free, but for which establishing the diagnosis would not alter the course of symptoms. Patients were far more willing to pay for such information (84%) than were managed care executives (43%) or even physicians (61%) (31). Similarly, patients are very willing to pay for fetal ultrasound during normal pregnancy, even though no benefits have been scientifically established. In part, such demand is a result of technology being generally oversold by health care providers.

Low Tolerance of Ambiguity by Doctor or Patient

One of my own elderly patients once presented to an emergency room complaining of difficulty seeing. Because of his age, the emergency room physician considered a diagnosis of temporal arteritis, a rare disease but one that can result in permanent visual loss. There is no specific blood test for this condition, although one nonspecific test for inflammation of any sort (the erythrocyte sedimentation rate) usually is positive in this condition. The test was done and was normal; but even so, the clinician was so concerned about the possibility that the patient was started on high doses of prednisone, a cortisone-like drug. The patient then suffered the new onset of diabetes related to the use of prednisone, experienced severe metabolic complications, psychiatric complications, two hospitalizations, and a temporal artery biopsy. Although the temporal artery biopsy was negative for temporal arteritis, the pathologist qualified his interpretation by noting that because the patient had been on prednisone, it was possible that the true diagnosis was masked, and he could not rule out the condition. Thus, hospital physicians continued the medication. Finally, I stopped the prednisone in the outpatient clinic, which resulted in prompt resolution of the patient’s diabetes and his psychiatric symptoms, and there were no further visual complaints. In retrospect, it appeared that the original visual complaint was related to presbyopia (difficulty focusing on close objects) and that the cascade of adverse events was precipitated by low tolerance of any ambiguity.

Desire to Legitimize Compensation Claims

Anecdotally, it appears that some patients are driven to obtain diagnostic tests by attorneys who want to help patients establish the legitimacy of disability compensation claims. In the case of MRI for the low back, as noted above, some abnormality is extremely likely, even in normal persons. In the context of returning patients to work, well-meaning clinicians feel obliged to respond to these abnormalities. Perhaps as a result, patients covered by the Workers’ Compensation System receive both more invasive types of back surgery and substantially more repeat back surgery than patients covered by other forms of insurance (59). Repeat back surgery is frequently regarded as an indication of poor outcome from the initial operation.
FACTORS THAT FACILITATE TRIGGERING CASCADES

Though not direct triggers, several factors may facilitate the initiation of cascade effects. One such factor is excess capacity of technology and specialists who use it. For example, the United States has an abundance of facilities for coronary catheterization and revascularization and of physicians trained to perform the procedures, in comparison to Canada and Europe (38, 49, 61). Cardiologists who perform angiography are more likely to recommend cardiac catheterization than are cardiologists who do not perform the procedure, and these in turn are more likely to recommend it than general internists (34). This probably accounts in part for the wider use of catheterization and revascularization in the United States, despite evidence that the higher rates of utilization do not overall result in patient benefit (38, 49).

A related factor is physician ownership or proximity to diagnostic facilities. Physicians who work in hospitals with catheterization facilities are more likely to recommend cardiac catheterization than those without easy access to such a facility (19, 48). Convincing evidence suggests that ownership of imaging facilities leads to greater use of imaging (29, 30).

Other perverse financial incentives may result in additional forms of conflict of interest. Identifying more disease means more business. This may partly explain screening campaigns by hospitals or health care systems and the aggressive marketing of diagnostic tests. The concept of physician-induced demand is a long-standing principle of health services, which suggests that physicians can order return visits, perform diagnostic tests, or recommend procedures to some desired level, especially when indications are vague or controversial. The availability of insurance reimbursement also affects health care utilization, as demonstrated in the use of pulmonary artery catheters. In a study of 27 hospitals, private insurance coverage was associated with a 73% greater likelihood of receiving a pulmonary artery catheter (51).

Another factor facilitating cascade effects is the attitude that more information is always better. However, in considering a problem such as pancreatic cancer, which has a high mortality but also a high treatment-associated mortality, a decision analysis suggested that certain diagnostic test strategies could actually result in higher mortality rates than not performing a diagnostic test. This result was in part related to the limited positive predictive value of the tests. The authors advocated wider use of decision analysis to help understand when diagnostic testing was really likely to produce overall better outcomes for large groups of patients (57).

Overdependence on surrogate endpoints in clinical trials may also promote ill-advised clinical cascades. Surrogate endpoints are physiologic phenomena such as blood tests or imaging results, which are assumed to be markers of the ultimate outcomes of concern to the patient (such as death, loss of vision, or other severe disability). A now classic example of problems with surrogate markers was the use of the antiarrhythmic drugs encainide and flecainide to suppress premature heartbeats among patients who had suffered a heart attack. These drugs were highly effective...
at suppressing extra beats, and such extra beats were known to be associated with higher mortality rates. Nonetheless, when the drugs were tested in a randomized trial, they resulted in approximately twice the mortality of placebo therapy, which indicated that the surrogate outcome of arrhythmia suppression was not an adequate marker of the outcome of interest (17). Several trials have suggested that shrinkage of malignant tumors is not an adequate marker for length of survival. In some trials, changes in CD4 cell counts have not accurately reflected changes in progression to AIDS or to death among HIV-positive patients (22). Such examples abound in medicine, and they indicate one of the ways in which test results can be overinterpreted.

Finally, a host of factors unrelated to scientific evidence may promote the use of new tests and treatments, often in the absence of strong evidence of their benefit. These factors include direct-to-consumer advertising, aggressive marketing to health care providers, media hype of new technology, political pressure from advocacy organizations, legal decisions, and even legislative action. Figure 1 is a diagram of the many factors that may influence the use of medical technology, sometimes in the face of limited evidence of benefit, or even evidence to the contrary.

CONSEQUENCES OF CASCADE PHENOMENA

Perhaps the most worrisome potential consequences of cascade phenomena are iatrogenic illness, morbidity, and mortality. For example, Starfield has summarized some of the evidence for adverse effects that occur because of iatrogenic injuries...
CASCADE EFFECTS

not associated with recognizable error. These include some 12,000 deaths a year from unnecessary surgery, 80,000 deaths per year from infections in hospitals, and 106,000 deaths per year from non-error-related adverse effects of medications (58). It is impossible to determine what fraction of these are related to cascade effects, but we may speculate that it is a substantial fraction. Data such as those from randomized trials of more aggressive or conservative management for acute coronary syndromes suggest that the aggressive strategy does not improve mortality, and its complications may increase mortality in some settings (38). Expert panels have judged that a substantial fraction of back surgery in the United States is unnecessary and exposes patients to unnecessary complications and even mortality (46). The abundance of technology and specialists in the United States, compared with most other developed countries, has not assured better public health: the United States ranks tenth or below for indicators such as low birth weight percentage, neonatal mortality, years of potential life lost, and life expectancy at age 1, age 15, and age 40 (58).

A less obvious consequence of cascade phenomena may be labeling effects for patients who have no disease. This problem has been demonstrated, for example, among children with benign heart murmurs, who experience greater restriction of physical activity than children without cardiac murmurs, despite having normal hearts (3). Anecdotal experience suggests that many patients who have spinal MRI tests attribute great importance to findings of bulging discs or other degenerative changes, despite evidence that these are as common in asymptomatic patients as among those with back pain. There is some evidence to suggest that simply attaching a diagnosis to patients who were previously unaware of having high blood pressure may result in greater work absenteeism, regardless of whether therapy is begun (28). Thus, labeling effects may be associated with unnecessary disability.

Unnecessary costs are an obvious consequence of cascade effects. The follow-up testing required for unexpected abnormalities, ongoing monitoring, and management of complications for unnecessary procedures are all examples of cost without benefit in the health care system.

One of the mechanisms by which treatment complications may occur is tampering with stable conditions. Tampering occurs when adjustments are made to correct deviations in a system that reflect random variation rather than systematic change. Intervening in response to random variations actually causes a system to become less stable and increases the likelihood of unnecessary treatment and adverse events (5, 21). Modern physicians are flooded with measurements as we monitor a host of physiologic phenomena. Multiple measures from the pulmonary artery catheter are examples in critically ill patients. Other examples include measuring prothrombin times and changing anticoagulant doses; measuring oxygen content of the blood and changing respirator settings; measuring fever and changing antibiotics; measuring blood pressure and changing antihypertensive therapy; measuring blood sugar and changing insulin doses; and a host of other common examples (5).
A false impression of high disease prevalence and great treatment efficacy are additional consequences of cascade effects and may in turn increase their likelihood. Newer diagnostic technology, which detects ever smaller or milder abnormalities, aggravates the problem. Many of the small abnormalities detected with new imaging techniques are clinically irrelevant. Even clinically important abnormalities are detected at much earlier stages than was previously possible. Because of this, their outcomes superficially appear to be improved, when in fact much of the improvement is due to lead time bias and length bias. Higher rates of detection create the impression of higher disease incidence and prevalence, which, along with seemingly improved treatment efficacy, lead to ever more frequent testing and treatment (6). This cycle affects both individual patients and large populations of patients for whom resources are unnecessarily wasted. Such interventions may also lead to unnecessary iatrogenic illness.

Good evidence for a misleading change in apparent disease outcome was described in a study of patients with lung cancer. In 1977, patients with every stage of lung cancer had better outcomes than in cohorts treated in 1953 or 1967 at the same institutions. On record review, it became apparent that this occurred because newer diagnostic techniques identified small metastases that formerly had been silent and unidentified, so that many patients who previously would have been classified in a good stage instead had migrated to a bad stage. The migration of patients out of the good stage meant that those remaining in the good stage had an even better outcome than would have previously occurred. Furthermore, because those who migrated into the bad stage represented the mildest and earliest form of disease in that stage, survival rates in the bad stage also improved. However, individual outcomes had not changed at all between time periods (20). When patients were classified according to symptoms in a fashion that would be unaffected by diagnostic tests, the two cohorts had similar survival rates. This observation was wryly named the “Will Rogers phenomenon,” recalling a comment on the dustbowl era: “When the Okies left Oklahoma and moved to California, they raised the average intelligence level in both states” (20).

WHY IS THERE LITTLE AWARENESS OF CASCADE EFFECTS?

Many have commented that the United States is a culture that assumes more and newer technology must be superior to less and older technology. In a 1994 survey, 33% of Americans, 27% of Canadians, and 11% of Germans thought “modern medicine can cure any illness with access to advanced technology” (7). Furthermore, the use of technology promotes a wider patient expectation of its use. This was shown in a randomized trial in which low-risk patients with back pain were allocated to either early spine X-rays or a brief educational intervention. Baseline beliefs that “everyone with back pain should have an X-ray” were equivalent between the groups and remained stable in the education group, but increased in the group that received X-rays (16).
Patients often do not recognize unnecessary medical care, and patient requests for unnecessary care are common in day-to-day practice. Patients often equate laboratory testing and imaging with high-quality care and assume that the only reason for not performing such tests is financial. At least some fraction of unnecessary surgery appears to be driven by patient demand. Thus, physicians and patients may unwittingly conspire to perform unnecessary tests and therapeutic interventions. Both may perceive this to be in their best interests, and complications or costs are seen as the price of providing good medical care.

The marketing of new technology is designed to maximize demand, and this has reached its extreme expression in direct-to-consumer advertising of prescription pharmaceuticals. Many physicians report that patients request unnecessarily expensive drugs, or even unnecessary drugs, based on effective print or television advertisements.

Many of the factors that initiate cascades, including physician anxiety and fear of litigation, undoubtedly fuel cascades in progress. Many physicians are not well trained in the implications of false-positive and false-negative diagnostic tests and do not incorporate the concept of positive predictive value in making clinical decisions. Thus, from the physician perspective, many initiating events seem perfectly justified and appropriate. As a result, there is little recognition by either patients or physicians of undesirable cascade effects, except in extreme examples.

AVOIDING CASCADE EFFECTS

Because cascade effects are difficult to identify while they are in action, highly specific recommendations are likely to remain elusive. Nonetheless, certain efforts may help to reduce the likelihood of initiating events. In their seminal article, Mold & Stein suggested the importance of performing a complete patient history and physical examination in order to identify previously recognized conditions, diagnoses, abnormalities, or test results, and to avoid unnecessary duplication of tests. They also pointed out that a complete clinical evaluation is critical to providing an appropriate clinical suspicion (pretest probability) for any subsequent diagnostic testing (44). A related factor is continuity of care. A single primary care provider becomes more familiar with patients and can recognize true deviations from their usual health state in a more sensitive and specific fashion. A health care provider seeing a patient for the first time or only on a single occasion cannot have an equivalent understanding of the patient’s usual health, usual responses to symptoms and illness, and style of using medical care.

Physicians should acquire a more complete understanding of the predictive value of diagnostic tests. While this is unlikely to result in highly quantitative calculations in routine clinical settings, it would help physicians to better understand the probability of false-positive tests, circumstances in which testing is unlikely to be beneficial, and the importance of understanding the performance characteristics
of the tests they order. In a similar vein, diagnostic testing should only be undertaken to answer very specific questions in order to avoid a shotgun approach. Highly selective approaches are more likely to avoid test complications, false-positive results, costs, and anxiety.

Involving patients more completely in clinical decisions may also be important. When patients are well informed of the benefits and risks of alternative approaches to clinical problems, they often have preferences that differ from their physicians’.

For example, patients generally have higher thresholds for beginning antihypertensive therapy or anticoagulant therapy in many clinical circumstances (41, 43). In screening for prostate cancer, well-informed patients are more likely to decline tests than patients receiving routine care (23). Thus, when patients understand the stakes involved with clinical decisions, they may be helpful in averting unnecessary risks.

Researchers may help by better exposing the natural history of increasingly mild disorders detected by advanced technology. Better evaluation of the benefits and harms of treating such mild disease will also facilitate future screening and diagnostic decisions.

Clinicians should insist on proven benefits of therapy for the particular types of patients being considered for treatment. They should be increasingly aware of complication rates in their own hands and in their own facilities. The example of balancing risks and benefits of carotid endarterectomy provides an instructive example (42).

It is important to study the impact of system capacity with regard to tests and interventions. For example, lower capacity for invasive cardiac interventions in Canada reduces their use without apparent detriment to overall health. More deliberate deployment and dissemination of new technology and regionalizing some procedures may have similar benefits. Concentrating high-risk procedures in the hands of centers and physicians with a high clinical volume appears likely to improve outcomes and reduce complication rates (25, 47, 37).

Clinical guidelines, now promoted in many quarters, may help to reduce the unnecessary use of new technology and offer health benefits as well. Decision aids based on clinical guidelines, such as computer-based decision support systems, appear to help reduce the use of unnecessary tests (62).

While the use of second opinions in planning surgical procedures has become fairly common, it is less common to seek second opinions in test interpretation. Nonetheless, second opinions in the interpretation of complex diagnostic tests such as coronary angiograms and pathologic specimens may be warranted in certain situations where the consequences of those interpretations have major implications for both patient benefit and risk.

Physicians and patients alike must recognize that newer and more is not the same as better. Innovators, researchers, and early adopters of new technology should be alert to unanticipated adverse effects. For new medications, premarketing evaluation is quite rigorous, but this is less true of devices and procedures. For all of these new technologies, postmarketing surveillance has been relatively informal
and poorly organized, and better surveillance (e.g., with disease or treatment registries) may help to identify problems at an earlier stage of technology dissemination. Registries for new devices, for example, may serve to identify unexpected complications or a high rate of complications sooner than unmonitored routine practice.

ACKNOWLEDGMENT

Thanks to Dr. Linda Pinsky and Dr. Scott Ramsey for critical reviews of an earlier draft. Pam Hillman provided expert clerical support. Supported in part by grant #042251 from the Robert Wood Johnson Foundation’s Investigator Awards in Health Policy Research Program.

Visit the Annual Reviews home page at www.annualreviews.org

LITERATURE CITED

CASCADE EFFECTS


55. Schwartz LM, Woloshin S, Sox HC,
44  DEYO


CONTENTS

Epidemiology and Biostatistics

HALYs and QALYs and DALYs, Oh My: Similarities and Differences in Summary Measures of Population Health,
Marthe R. Gold, David Stevenson, and Dennis G. Fryback 115

The Importance of the Normality Assumption in Large Public Health Data Sets, Thomas Lumley, Paula Diehr, Scott Emerson, and Lu Chen 151

Gambling and Related Mental Disorders: A Public Health Analysis, Howard J. Shaffer and David A. Korn 171

Thinking Outside the Box: Recent Advances in the Analysis and Presentation of Uncertainty in Cost-Effectiveness Studies, Andrew H. Briggs, Bernie J. O’Brien, and Gordon Blackhouse 377

Environmental and Occupational Health

Socioeconomic Status and Health: The Potential Role of Environmental Risk Exposure, Gary W. Evans and Elyse Kantrowitz 303

Effects of Smoking Restrictions in the Workplace, Ross C. Brownson, David P. Hopkins, and Melanie A. Wakefield 333

The Future of Benefit-Cost Analyses of the Clean Air Act, Alan Krupnick and Richard Morgenstern 427

Public Health Practice


The Effectiveness of State-Level Tobacco Control Interventions: A Review of Program Implementation and Behavioral Outcomes, Michael Siegel 45

Challenges in Motor Vehicle Safety, Patricia F. Waller 93

Effects of Smoking Restrictions in the Workplace, Ross C. Brownson, David P. Hopkins, and Melanie A. Wakefield 333

Tuberculosis, Parvathi Tiruviluamala and Lee B. Reichman 403
## CONTENTS

### SOCIAL ENVIRONMENT AND BEHAVIOR

<table>
<thead>
<tr>
<th>Title</th>
<th>Author(s)</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Direct Marketing of Pharmaceuticals to Consumers</td>
<td>Alan Lyles</td>
<td>73</td>
</tr>
<tr>
<td>Dietary Interventions to Prevent Disease</td>
<td>Deborah J. Bowen and Shirley A. A. Beresford</td>
<td>255</td>
</tr>
<tr>
<td>The Macroeconomic Determinants of Health</td>
<td>S. V. Subramanian, Paolo Belli, and Ichiro Kawachi</td>
<td>287</td>
</tr>
<tr>
<td>Socioeconomic Inequalities in Injury: Critical Issues in Design and Analysis</td>
<td>Catherine Cubbin and Gordon S. Smith</td>
<td>349</td>
</tr>
</tbody>
</table>

### HEALTH SERVICES

<table>
<thead>
<tr>
<th>Title</th>
<th>Author(s)</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cascade Effects of Medical Technology</td>
<td>Richard A. Deyo</td>
<td>23</td>
</tr>
<tr>
<td>Direct Marketing of Pharmaceuticals to Consumers</td>
<td>Alan Lyles</td>
<td>73</td>
</tr>
<tr>
<td>Morbidity and Mortality from Medical Errors: An Increasingly Serious Public Health Problem</td>
<td>David P. Phillips and Charlene C. Bredder</td>
<td>135</td>
</tr>
<tr>
<td>Utilization Management: Issues, Effects, and Future Prospects</td>
<td>Thomas M. Wickizer and Daniel Lessler</td>
<td>233</td>
</tr>
</tbody>
</table>

### INDEXES

<table>
<thead>
<tr>
<th>Type</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subject Index</td>
<td></td>
<td>449</td>
</tr>
<tr>
<td>Cumulative Index of Contributing Authors, Volumes 14–23</td>
<td></td>
<td>473</td>
</tr>
<tr>
<td>Cumulative Index of Chapter Titles, Volumes 14–23</td>
<td></td>
<td>478</td>
</tr>
</tbody>
</table>

### ERRATA

An online log of corrections to *Annual Review of Public Health* chapters (if any have yet been occasioned, 1997 to the present) may be found at [http://publhealth.annualreviews.org/](http://publhealth.annualreviews.org/)