

Ethical Issues in Cancer Screening and Prevention

ANYA PLUTYNSKI*

The University of Utah, Salt Lake City, Utah, USA

*Address correspondence to: Anya Plutynski, PhD, The University of Utah, Salt Lake City, UT, USA. E-mail: anya.plutynski@philosophy.utah.edu

November 2009's announcement of the USPSTF's recommendations for screening for breast cancer raised a firestorm of objections. Chief among them were that the panel had insufficiently valued patients' lives or allowed cost considerations to influence recommendations. The publicity about the recommendations, however, often either simplified the actual content of the recommendations or bypassed significant methodological issues, which a philosophical examination of both the science behind screening recommendations and their import reveals. In this article, I discuss two of the leading ethical considerations at issue in screening recommendations: respect for patient autonomy and beneficence and then turn to the most significant methodological issues raised by cancer screening: the potential biases that may infect a trial of screening effectiveness, the problem of base rates in communicating risk, and the trade-offs involved in a judgment of screening effectiveness. These issues reach more broadly, into the use of "evidence-based" medicine generally, and have important implications for informed consent.

Keywords: *base rate fallacy, cancer screening, risk communication, trade-off neglect, USPSTF*

I. INTRODUCTION

Cancer is the second leading cause of death in the United States. Understandably, with so many peoples' lives affected by this devastating disease, early detection has become the mantra of many cancer activists. Screening for disease in asymptomatic individuals has an intuitive appeal. If all cancers have a linear progression toward morbidity and mortality, then halting the progress of the disease early seems the best strategy. Many are hopeful that

more and better screening modalities will be provided to a wider public. If not all cancers have linear progression (as seems clear, cf. Croswell, Ransohoff, and Kramer, 2010), more, however, may not always be better. The recent upset over the USPSTF's recommendations for screening for breast cancer brought the intuitive thinking behind early detection into question and raised a host of ethical issues, besides. The object of this essay will be to examine the intuitive thinking behind support for frequent screening, consider the nature of the evidence for the USPSTF's recommendations, and whether the recommendations can be made consistently with respect for autonomy and beneficence. I will argue that the recommendations can and should be followed consistently with respect for autonomy and beneficence but argue that safeguards need to be put in place to ensure this.

The USPSTF recommended that women under 50 consult with their clinicians about whether to have annual screening (US Preventive Services Task Force, 2009). A meta-analysis of multiple clinical trials appeared to show that the benefit for most women under 50 may be relatively small, and the cost—in terms of false positives, unnecessary biopsies, and overdiagnosis (unnecessarily treating cancers that never would have caused illness during that person's lifetime), is relatively high. For women over 50, they found that biannual screening may be as effective as annual screening for reducing age-adjusted mortality (Mandelblatt et al., 2009; Nelson et al., 2009). As might have been predicted by more savvy publicists, the response was not positive. Many breast cancer activists, and the American Radiological Association, as well as the American Cancer Association dismissed the findings. Some argued that the USPSTF was condemning women to early death. Moving past the outrage, we need to look closely at both the salient scientific and ethical issues at stake in this heated exchange.

The USPSTF recommended individual women in their forties decide whether to screen in consultation with clinicians. Why does this create such outrage? Perhaps some worry that insurance companies will not cover screening if screening is not standard care for some age group, or that even those patients at elevated risk (perhaps due to family history) will not seek screening based on recommendations for the population at large. However legitimate these concerns, the matter of whether or not *most* women under 50 *should* get screened should depend upon whether the risk to those screened is outweighed by the benefit. What is the evidence for the benefits for women in this age group? We should not simply assume that more screening is better, and that all screening modalities for all cancers are risk free for all patients. As the authors of a recent article put it, we should be sure to “protect ourselves from our intuitions” (Croswell, Ransohoff, and Kramer, 2010). Not all cancers progress linearly, and some may never cause clinical symptoms—in other words, preventative screening methods may not always be for the benefit of patients, for some cancers and in some populations. Although many of us assume that in the era of modern medicine, no one prescribes either tests or treatments that do more harm than good, we should make sure to check the evidence.

Although it is surely correct that clinicians and patients (and insurance companies) find it easier to follow general recommendations, when screening will only benefit a minority at greater risk, it seems that individual-based decision making, rather than universal screening, is a better option. Indeed, the USPSTF was clear that women should consult individually with physicians; their recommendations were *not* “broadly utilitarian,” in the sense of simply considering costs for insurance companies, but instead, emphasized individual assessment and weighing of risks, with clinicians. However, individual-based decision raises another suite of issues about risk communication and decision making under uncertainty. Some have argued that patients are simply bad at making choices based on, for example, misunderstanding of the differences between absolute versus relative risk (Schwartz and Meslin, 2008). As Schwartz and Meslin explain, pamphlets encouraging mammography tend to cite relative over absolute risk as the apparent “risk reduction” will always look higher described as relative risk. Moreover, they support this strategy. They explain:

Absolute risk reduction (ARR) is calculated by subtracting the mortality rate in the screened group from that of the unscreened group. Thus, according to one study, the ARR of performing mammography regularly for 20 years is 4.9 per 1,000 (or 0.49%), i.e., 14 per 1,000 (1.4%) minus 9.1 per 1,000 (0.91%) (Barrat et al., 2005). In contrast, Relative Risk Reduction (RRR) is calculated by *dividing* the decrease in morbidity or mortality by the baseline rate. Using the same model as above, mammograms performed over 20 years would provide a RRR of 36%, i.e., 4.9 per 1,000 decrease in death from breast cancer divided by the baseline rate (in the unscreened group) of 14 per 1,000. In this case as in others, the RRR (36%) is a much larger number than the ARR (0.5%), and so it is not surprising that brochures that aim to encourage patients to get regular mammograms generally present RRR if numbers are given at all. (Schwartz and Meslin, 2008, 868)

As mentioned above, the difference apparent reduction of risk for relative versus absolute risk reduction may be quite stark, particularly when screening for disease that is not very prevalent. Schwartz and Meslin endorse the practice of describing relative risk reduction, as opposed to absolute risk reduction, to patients, to encourage mammography. However, we should be careful to not slip into paternalism, particularly in “gray areas” where the risk reduction is small, and the costs of screening may be relatively high (Quanstrum, 2009). Although a narrow utilitarian or expected utility perspective might simply attach values to lives saved, and so use any strategy (including representing risk deceptively, or discounting or hiding cost to patients due to unnecessary biopsies or overdiagnosis), there are broader issues at stake. Questions of risk and benefit need to be supplemented by some discussion of the reasonable variability in values patients attach to different risks, the norms of clinician-patient relationships, and what respect for autonomy and informed consent requires. Questions about risk and benefit can be better answered once we know how sensitive or specific are our screening tools,

how prevalent the disease, and thus what the risks are of false positives. However, we also need to assess the values behind general versus individually tailored recommendations, and arguments for and against individual consultation with clinicians, versus general recommendations that may benefit some, but not *most* patients.

Even the best clinical trials often leave some gray areas where it is not clear that benefit outweighs cost (Quanstrum and Hayward, 2009). Evidence-based medicine is surely important, but it cannot always draw a bright clear line between when universal screening is advisable and when it is not. In this paper, I will argue, along with Malm (1999), that public health recommendations should be such that the benefit on balance for the patient who receives a recommendation is greater than the harm.¹ By this, I mean that that benefit for any particular patient is higher than risks of harm, taking into account the probability that this harm will occur. However, of course, how risks or benefits are assessed is not merely a matter for science to decide. Although EBM may give us information that we can use to make judgments about policy, in some cases, the evidence underdetermines the recommendations. In just such cases, to make public policy recommendations, absent consideration of the patient's (a) particular risk, and (b) weighting of preferences, is unsound.

Further, we need to also consider the role and limitations of risk communication. When communicating with a patient about their preferences, we should make clear when the recommendations are not 'black and white'; some women need to receive false positive tests, have unnecessary biopsies, or be unnecessarily treated on balance to save lives. How often will women be willing to take the risk of being in the first category to offset the chance of landing in the latter? Similarly, how many men should receive diagnoses of "indolent" prostate cancer that may never have presented clinically to save one life? Arguably, individual patients need to decide for themselves whether they will be willing to go through the risk of unnecessary post-screening tests.

Unnecessary tests—test that, by and large, do not benefit the vast majority of patients, and may harm them—should not be touted as a human right and a necessity for health and survival. Screening does have downsides, as well as positive outcomes (see, e.g., Gøtzsche and Nielsen, 2009). Patients should know that they may be exposed to unnecessary radiation, that false positives are more likely when a disease is rare, and that in many cases, screening for early disease is neither always necessary nor beneficial. For instance, many more prostate cancers are detected and treated today than 30 years ago due to the promotion and use of PSA tests, and yet it is not clear that quality of life for the vast majority of men treated for prostate cancer is thereby improved (Esserman, Shieh, and Thompson, 2009). Since most prostate cancers are slow growing, such that many men may never have clinical symptoms of these cancers, and treatment can cause impotence and incontinence, it is not

clear that the vast increase in number of prostate cancers diagnosed benefits *most* patients or extended lives.

However, those who administer and profit from these tests and treatments do benefit. We cannot always assume that makers of mammography and other technologies are motivated entirely by altruism. There must be oversight of medical technologies and procedures to ensure that they help and do not harm patients (Groopman, 2007; Quanstrum and Hayward, 2009). The following is an overview and discussion of the general principles that should be put in place for ethical practice of cancer screening and prevention and a call for areas where we need to be more attentive to the variable biology of different cancers, as well as the variable values of patients.

II. AUTONOMY

Respect for individual decision making and informed consent are central to a commitment to patient autonomy and are, of course, required by law for federal research (National Institutes of Health, 2005). However, respect for autonomy is not simply a matter of giving patients choices; it requires that they be informed about both risks and benefits of screening, and it requires recognizing that patients may value risks and benefits differently. PSA tests should be accompanied with some explanation, for example, of disease prevalence, morbidity and mortality, test sensitivity and false positive rates, and side effects of further screening and treatment. As mentioned above, for many prostate cancers, “catching it early” may not be for the best, at least in terms of quality of life. Many men die of heart disease, but die with prostate cancer, having never experienced clinical symptoms. A better understanding of the distinct pathology and genetics of more and less aggressive forms of such cancers is important to assessing the value of such tests. Clinicians are understandably frustrated when, even when screening does appear to have a proven benefit, some patients refuse screening because they are afraid of medical procedures, do not like being put in the role of a “patient,” or simply do not trust the evidence. In such cases, clinicians might provide the patient with more information or better explanations of benefits and risks of screening. However, they also must recognize that it is not unreasonable to value some choices over others; choices to not be screened may not be due to ignorance but differences in values.

Informing patients about risk and benefit, as many have argued, is not a straightforward matter. Even those with PhDs in statistics are subject to heuristics and biases that make it difficult to understand or distinguish between absolute and relative risk, for instance (Kahneman and Tversky, 1979; Weinstein, 1980; McKenna, 1993; Baron, 2000; Bowling and Ebrahim, 2001). Availability heuristics and base rate fallacies are problems that plague researchers as well as the average Joe. Clinicians may be as subject to bias as patients; a clinician

does not want to think of the possibility that the tests (or treatment) are either not effective or may even harm their patient unnecessarily.

With this in mind, respect for autonomy requires that those who manufacture and profit from diagnostic tests actively seek to document and acknowledge risk of screening tests. Federal regulations should require them to provide effective ways of communicating such risk to both clinicians and patients. Ideally, third parties without an interest in misinformation should assess the science behind such risk/benefit assessments. Most of us assume (correctly) that an X-ray or even a CT or two is not a serious risk. However, it has become apparent that such machines may malfunction due to operator errors as well as improper maintenance. Moreover, a lifetime of tests can add up. Patients must be made aware that multiple CT scans and X-rays, over time, are not “risk free.” Risks should be communicated in a way that is user friendly and transparent, and patients and clinicians should communicate about the cumulative effects of such tests over a lifetime.

For example, a patient admitted for a CT scan should be provided a pamphlet showing exactly how much an excess of a certain number of such scans will, on average, increase risk for cancer. Expression of risk in the simplest possible fashion—for example, 1 in 100 individuals who have more than N CT scans are at elevated risk for cancer to extent x —would be most effective. Clinicians, unfortunately, are incentivized to offer more tests, and many do not have accurate knowledge about the risks of such tests, either because manufacturers play down the risks or because there is simply insufficient data. Either this incentive structure for clinicians needs to change or independent parties need to provide some form of evidence-based regulation of unnecessary and risky tests.

III. BENIFICENCE

Clinicians and their professional organizations have a vested interest in arguing that their screening and treatment protocols are safe and effective. An independent third party should watch over such claims and make sure that the claimed benefits genuinely outweigh the risks. Sometimes this will involve sponsoring more and better clinical trials. However, epidemiologists may estimate when and whether a test is more or less effective, even without a clinical trial, if they know how prevalent a disease is and how sensitive and specific the available screening modality.

A general rule of thumb is that if a disease is not very prevalent, false positives will be relatively common. A simple example of how prevalence, for instance, is relevant to the effectiveness of screening regimes follows (from Buys, 2011). Suppose that you have a screening test that has very high specificity (the probability that a negative test means the disease is absent) and high sensitivity (the probability that the test will be positive if the disease is

Table 1. Hypothetical disease no. 1—high prevalence

	No. with positive test	No. with negative test	Totals
Disease	475	25	500
No disease	475	9,025	9,500

present). Let us say that the test is 95% specific (95% of positive tests are due to disease) and 95% sensitive (95% of disease is identified). Suppose 10,000 individuals are tested. The performance of the screening test will vary depending upon whether the disease is common (say, 5% prevalence) or very uncommon (.05%).

Consider table 1: For the disease with high prevalence, the positive predictive value of the test (or the proportion of those who test positive who have the condition) = $\frac{1}{2}$. You can calculate this number yourself by dividing the number of true positives by the number of true and false positives (475/950). This means that half of those who screened positive were in fact diseased. That may seem like a very high false-positive rate, but consider table 2 for a much less prevalent disease.

Consider table 2: For the disease with lower prevalence, the positive predictive value of the same test is 1/101. This means that only one out of 101 individuals who screened positive were in fact diseased (see table 2). Thus, the effectiveness of the very same screening modality is significantly less for less common diseases. Cancer, on average, is more prevalent in older populations.

The lesson of this example is fairly simple: when a disease is not very prevalent (as is the case for most cancers in most persons under 50), even the most sensitive and specific screening test is not as effective, and may be more harmful, than not being screened.

What are the costs of a false positive test? Perhaps a few days of worry, followed up by a relatively noninvasive biopsy, are, some argue, not a serious problem. However, sometimes, according to a recent report, unnecessary open biopsies are done when needle aspirated biopsies would be equally effective and safer (Gutwin et al., 2011). Moreover, for many cancers, the only recourse after a positive diagnosis is a very serious surgery. For instance, in the case of ovarian cancer, ovariectomies are not a minor procedure and one with some serious side effects long term. Further, pathologists are not immune to error. So, positive biopsies, that is, claims of malignancy, can be in error. How often does this occur? According to some (Moss, 2005;

Table 2. Hypothetical disease no. 2—low prevalence

	No. with positive test	No. with negative test	Totals
Disease	5	0	5
No disease	500	9,495	9,995

Paci et al., 2006; Welch and Black, 2010) pathologists may overdiagnose some cancers as much as one third of the time. Pathologists' judgments about the "invasiveness" or "metastatic potential" of a tumor vary; one pathologist's atypical ductal hyperplasia, a pathological finding that increases the risk of subsequent cancer but is not itself cancer, may be another pathologist's ductal carcinoma in situ, which may be treated with surgery and radiation therapy. Although on the side of the clinician, it is better to be "safe than sorry," it is not always safer for the patient to undergo ever earlier screening and possibly unnecessary surgical procedures. The upshot is: patients should be able to decide what risk they are willing to accept, and not all screening options are risk free.

IV. EVIDENCE-BASED MEDICINE

In light of these concerns, there needs to be more and better clinical trials of screening modalities and other preventative measures. However, even the best clinical trial is not free of bias. A variety of concerns have been raised about evidence-based medicine, from concerns about meta-analyses and their limitations to the variety of limitations of RCTs, randomized clinical trials. Screening RCTs are subject to a number of systematic biases, and their implications for decision making under risk and uncertainty are not simple.

To start with some simple examples, consider lead-time bias—when detection does not on average extend life but only makes patient aware of a diagnosis earlier. This is likely a confounding factor in many clinical trials of new screening modalities. More sensitive detection methods mean that cancers are detected earlier, which leads to higher estimates of time to mortality since diagnosis. However, if undetectable metastases have occurred prior to the cancer being detected by the screening test, the cancer may cause death at the same time as if it had never been detected. Thus, the measure or "end point" of screening tests should be the "age-adjusted mortality," rather than years of survival after diagnosis.

Length bias—the fact that while many indolent (slow growing) cancers show up on screening exams, but more aggressive cancers tend to show up between screenings—is a serious problem for some cancers. Length bias leads to the conclusion that screening and therefore early treatment improves survival, whereas screening has simply identified less aggressive tumors.

Finally, overdiagnosis bias (the extreme of length bias) is the most damning problem. It is possible that many people are screened and treated for cancers that may never have caused clinical symptoms. This is particularly true in populations of patients over 80, who are more likely to die of other causes and so may never have suffered any clinical symptoms of a slower growing cancer. Older adults with significant comorbidities should not be subject to the toxic effects of cancer treatment unnecessarily.

What the case of overdiagnosis in older populations illustrates is that context matters. Not all screening tests are best for all patients at all ages, sexes, or ethnicities. Underrepresented groups should be studied for variability in age of onset of different cancers, as well as in aggressiveness of different cancers. It is well known that African Americans (particularly men) have higher rates of mortality from cancer. How much of this is because they suffer from more aggressive cancers and how much of it is due to the fact that African Americans on average do not have adequate access to screening and/or treatment? We need to remedy both inadequate access and inadequate or ill-suited treatment regimens, taking human variability into consideration. More inclusive and comprehensive RCTs of screening for different cancers in different ethnic groups may help resolve some of these dilemmas (for a discussion of the dilemmas of difference, see King, 1992).

In addition to the possible confounding factors of clinical trials, there are a variety of methodological and broader policy concerns that some raise about evidence based medicine.

First, a general worry is that EMB will result in a “one-size-fits-all” approach to setting standards of care. For example, universal screening for a disease starting at some specific age, when individual assessment of risk may be more appropriate. Many worry that EBM and policy will reduce the power of clinicians to make decisions on a case-by-case basis. Moreover, perhaps clinicians need to have their enthusiasms (occasionally) reined in by the evidence. As the case of open biopsies, reported above, and cases of back surgery discussed by Groopman (2007) make strikingly clear, the culture of medicine may be such that clinicians are trained in a practice—a surgical procedure, screening modality, or treatment—that may or may not have evidential support of benefit versus other (perhaps less costly) options. To avoid cases of bias, clinicians may occasionally need to have their practices and choices “second guessed” by epidemiologists.

Second, some may fear that where EBM points to “low” or “gray areas” of benefit, third-party insurers may not be willing to pay for screening or other tests or treatments. This is a legitimate concern, but perhaps this is an argument for, rather than against, more public and open communication about exactly when screening is and is not effective. For instance, if a clinician needs to justify a screening test by giving evidence of increased patient risk, this may actually reduce unnecessary harm to the vast majority of patients who would receive little individual benefit from such a test.

Third, a variety of philosophers have raised questions about the “water-tight” conclusions of RCTs. It is not clear, at least according to Worrall (2002, 2007), that randomized clinical trials are the “final word” on any medical question (they are famously subject to reversals!), and he raises legitimate objections to the claim that randomization (necessarily) eliminates all known and unknown confounders. Further, some (Cartwright, 2010) argue that we need to be careful of “exporting” the results of RCTs into contexts that are

significantly different. It is not clear that the results of even the most carefully conducted RCT at time X in geographical location Y can be exported to an entirely different location, time, or culture. Medical culture, as well as patient population, varies by location, university setting, and rural versus urban areas.

Fourth, critics of meta-analyses argue that it is not altogether clear that they are free of bias (Stegenga, 2011). Bias can enter in at the stage where choice of studies to include in a meta-analysis is made, either in terms of the effect measured or inclusion or exclusion on the bases of assessments of quality and other methodological criteria. Further, different RCTs in different locations (even though they claim to be measuring the same “effect”) may be drawing upon different patient populations, may be different in terms of rates of diagnosis at different institutions and may not have uniform measures of either cause or effect variable (RWJF, 2010). For example, when looking into the USPSTF’s meta-analysis of the effectiveness of screening, different institutions included ductal carcinoma in situ in their “count” of cancer diagnoses and some used different measure of end point; for example, relative risk versus hazard ratio.

Finally, meta-analyses do not have policy recommendations “written on their sleeves.” Even when all the issues surrounding RCTs and meta-analyses are appropriately addressed, an empirical study will always underdetermine policy recommendations. For instance, the USPSTF’s recommendations are based on clinical trials where patients were “invited” to be screened. Of course, compliance was not perfect. So, a recommendation of universal screening for individuals based on less than perfect compliance may be doubtful.

Confounders, such as “healthy volunteer” bias, may also contribute such that only those at less risk are actually seeking opportunities to be screened, so that the effectiveness of screening may be downplayed. Issues surrounding compliance require more empirical follow-up. If compliance is imperfect, some worry that recommendations of less frequent screening may cause some women to choose not to go at all. However, the opposite may be the case as well; if women are told that they only need to go every other year, they may be *better* able to comply. Empirical data on when and why patients seek screening might help resolve this issue. Determining how or whether annual versus biannual recommendations are more likely to be complied with is surely relevant to determining what should be recommended. Whether and why patients fail to go for screening seems to be very important to understand, not simply whether screening itself saves lives.

V. RISK AND PRECAUTION

Apart from the issues surrounding the empirical evidence for and against screening, there is a further matter of risk communication and decision

making under uncertainty. How should physicians communicate about screening, particularly when the benefits are small, or uncertain, and may be outweighed by not insignificant risks? How should patients weigh the risks and benefits of early screening? Ought one to make general clinical recommendations drawing upon a precautionary principle and what does it mean to do so? How, if at all, should individual medical decisions be conducted, as distinct from public health recommendations? Is it necessary for clinicians to communicate both absolute and relative risk to patients? What are we to make of the fact that both patients (and clinicians!) are poor at deliberating on decisions under risk? (see, e.g., Fagerlin et al., 2007; Schwartz, 2009).

As Resnik (2004) and Katz (1984) remind us in this context, we must be attentive to both the science and values at stake in medical decision making. The case of screening is not unique in this respect; many medical decisions are made without complete evidence. Diagnosis and treatment decisions often are made under time constraints; waiting for all the evidence to come in might mean that a decision comes too late. Resnik (2004) argues that where a risk is plausible and serious (e.g., if negative outcomes are irreversible), precaution is in order. This would seem to argue in favor of screening, whatever the potential risks of false positives or overdiagnosis. However, as Resnik remarks, given the sensitivity of the PSA test (and thus the high risk of false positives), the patient has to make an individual decision about whether he or she is willing to risk the need for follow up and possible treatment for a disease that was indolent. Precaution is a double-edged sword.

Schwartz and Meslin (2008) argue that providing patients with absolute rather than relative risk information may cause them to forgo a test or treatment that may be beneficial. As mentioned above, absolute risk is calculated by subtracting the mortality rate in the screened from the unscreened group. Relative risk is calculated by dividing the decrease in morbidity or mortality by the baseline rate. Schwartz and Meslin argue that providing relative risk information is preferable and warn that absolute risk information can cause individuals to forgo screening and the benefits thereof. If patients request further information, they should certainly be provided it (Schwartz and Meslin, 2008).

Schwartz and Meslin acknowledge that their proposal may raise legitimate concerns about informed consent and paternalism. There is, arguably, a principle of “truth in advertising” that clinicians should follow. By way of analogy, were the NRA to argue that carrying a firearm reduces one’s risk of dying in a terrorist attack by 50%, one should be skeptical, and not only because getting statistically significant data to support such a claim would be nearly impossible. Relative risk reduction can be deceptive, particularly when the baseline prevalence of some disease is exceedingly low. Clinicians must be attentive to this and to the fact that there may be more than one reasonable response to risk. Trade-off neglect, or the failure to acknowledge and recognize reasonable trade-offs in decision, is a very common failure in reasoning in such contexts. Different parties value precaution differently and

may value, for example, more years of life without worry of false positives over early aggressive treatment. As we have seen from the examples discussed above, we all (clinicians included) should not assume that our “intuitions” are always correct; early frequent screening may not always benefit all patients. Framing risk reduction in relative terms is, arguably, a form of coercion, particularly when considering cases of gray areas, where one could legitimately choose one of several options and not be entirely “irrational.” Values differ, and putting likely trade-offs “on screen” is at least plausibly a moral obligation, even if putting every possible remote risk on screen for patients is not.

Even clinicians may disagree as to the value of a treatment, procedure, or screening modality, depending on their knowledge and values. As Quanstrum argues, radiologists may be overly confident about the benefits of screening (Quanstrum and Hayward, 2009). Especially where there are gray areas concerning benefit, clinicians should watch that their own biases do not cause them to frame the issues so that they gently urge one course of action. However well meaning, such patterns run the risk of leading to overtreatment and overdiagnosis, no small risk, particularly for older patients. What is at risk is not simply worry or fear of cancer, but overtreatment, which is much more serious, particularly for elderly women and men who may well die from other causes before even experiencing clinical symptoms of disease.

VI. CONCLUSIONS

Although clinical trials of the effectiveness of screening are important, understanding cancer’s basic biology and natural history is as, if not more, important to both the development of smarter screening modalities and better treatment. There needs to be more swift progress in understanding the variable prognosis for varieties of “the same” cancers. The assumption that all cancers progress in a linear fashion at the very same speed toward morbidity and mortality is no longer reasonable (Zahl et al., 2008; Croswell, Ransohoff, and Kramer, 2010). Although it is not news that not all breast cancers are alike, as we learn more about the variety of the major forms of cancer, it appears that even some very aggressive and uniformly deadly cancers, such as pancreatic cancer, vary in their genetic profile and prognosis. This variation could make significant differences in recommendations for both screening and treatment. Although sequencing the “genome” of these different cancers promises much, it is also the case that targeted therapies need not await a complete “genetic diagnosis” of each and every therapy. Some of the best available targeted therapies were made possible without a complete understanding of the genetic bases of the disease—understanding not only the genetic but also molecular character, epigenetics, and/or developmental context of different cancers can suggest innovative treatment as well.

Cancer is not one disease, but several. Not all cancers progress in the same fashion. Some are “indolent”—they never or only at long last produce clinical

symptoms. Such slow-growing cancers may call for a “wait and see” strategy. With better and more evidence-based pathological information about the prospective course of different varieties of this disease, as well as the development of better, more targeted and less toxic therapies, this may be a case where less is more.

NOTES

1. As she puts it: “A necessary condition for an ethically justified screening recommendation is that the test reasonably can be expected to be beneficial on balance for the person taking it. The qualifier ‘on balance’ takes into account the benefits and harms of the tests and outcomes as well as the probabilities that those benefits and harms will occur. Thus it does not require that a given test be actually beneficial to everyone who takes it. Such a requirement would be unreasonable since, given the nature of screening, the vast majority of people screened will not have the disease in question and thus will not be able to benefit from the screening (except by being told that they do not have a disease that they had no reason to suspect they had in the first place). Instead, the standard requires only that the size of the benefit, taking into account the probability that it will occur, outweighs the size of the risks, taking into account the probability that they will occur. (Thus, an unlikely but great benefit can outweigh a very likely but small harm.) When this is the case, the test can be said to be good for the recipient, all things considered or ‘beneficial on balance’” (Malm, 1999, p. 27).

REFERENCES

- Baron, J. 2000. *Thinking and deciding*. 3rd ed. Cambridge: Cambridge University Press.
- Bowling, A., and S. Ebrahim. 2001. Measuring patients’ preferences for treatment and perceptions of risk. *Quality Health Care* 10(Suppl 1):i2–8.
- Buys, S. 2011. “Why do we die from cancer?” University of Utah Medical School Lecture notes, Cells, Molecules and Cancer Course, Spring 2011.
- Cartwright, N. 2010. Philosophy of Science Association presidential lecture. Philosophy of Science Association Meeting, Montreal, Canada.
- Croswell, J. M., D. Ransohoff, and B. S. Kramer. 2010. Principles of cancer screening: Lessons from history and study design issues. *Seminars in Oncology* 37:202–15.
- Esserman, L., Y. Shieh, and I. Thompson. 2009. Rethinking screening for breast cancer and prostate cancer. *Journal of the American Medical Association* 302:1685–92.
- Fagerlin, A., B. J. Zikmund-Fisher, and P. A. Ubel. 2007. ‘If I’m better than average, then I’m OK’: Comparative information influences beliefs about risk and benefits. *Patient Education and Counseling* 69:140–4.
- Götzsche, P. C., and M. Nielsen. 2009. Screening for breast cancer with mammography. *Cochrane Database Systematic Review* 4:CD001877.
- Groopman, J. 2007. *How doctors think*. New York: Houghton Mifflin.
- Gutwin, L. G., D. Gutwein, N. Ang, H. Liu, J. K. Marshall, S. N. Hochwald, E. M. Copeland, and S. R. Grobmyer. 2011. Utilization of minimally invasive breast biopsy for the evaluation of suspicious breast lesions. *American Journal of Surgery*. Available: [http://www.ajsfulttextonline.com/article/S0002-9610\(10\)00611-2/fulltext](http://www.ajsfulttextonline.com/article/S0002-9610(10)00611-2/fulltext) (Accessed March, 2011).
- Kahneman, D., and A. Tversky. 1979. Prospect theory: an analysis of decision under risk. *Econometrica* 47:263–92.
- Katz, J. 1984. *The silent world of doctor and patient*. New York: The Free Press.
- King, P. 1992. The dangers of difference. *Hastings Center Report* 22:35–8.

- McKenna, F. P. 1993. It won't happen to me: Unrealistic optimism or illusion of control? *British Journal of Psychology* 84:39–50.
- Malm, H. 1999. Medical screening and the value of early detection: When unwarranted faith leads to unethical recommendations. *Hasting Center Report* 1:26–37.
- Mandelblatt, J., K. A. Cronin, S. Bailey, D. A. Berry, H. J. de Koning, G. Draisma, H. Huang, *et al.* 2009. Effects of mammography screening under different screening schedules: Model estimates of potential benefits and harms. *Annals of Internal Medicine* 151:738–47.
- Moss, S. 2005. Overdiagnosis and overtreatment of breast cancer: Overdiagnosis in randomised controlled trials of breast cancer screening. *Breast Cancer Research* 7:230–4.
- Nelson, H. D., K. Tyne, A. Naik, C. Bougatsos, B. K. Chan, and L. Humphrey. 2009. Screening for breast cancer: an update for the U.S. Preventive Services Task Force. *Annals of Internal Medicine* 151:727–37.
- National Institutes of Health, Regulations and Ethical Guidelines. 2005. Title 45 CFR Part 46 Protection of Human Subjects. Code of Federal Regulations, Department of Health and Human Services. <http://ohsr.od.nih.gov/guidelines/45cfr46.html> (Accessed August, 2011).
- Paci, E., G. Miccinesi, D. Puliti, P. Baldazzi, V. De Lisi, F. Falcini, C. Cirilli, *et al.* 2006. Estimate of overdiagnosis of breast cancer due to mammography after adjustment for lead time. A service screening study in Italy. *Breast Cancer Research* 8:R68.
- Quanstrum, K., and R. A. Hayward. 2009. Lessons from the Mammography Wars. *New England Journal of Medicine* 363:1076–9.
- Resnik, D. 2004. The precautionary principle and medical decision making. *Journal of Medicine and Philosophy* 29:281–99.
- RWJF. 2010. *Who you are, where you live, affects medical treatment. Dartmouth study finds* [On-line]. Available: <http://www.rwjf.org/healthreform/product.jsp?id=20855> (Accessed May 17, 2010).
- Schwartz, P. 2009. Disclosure and rationality: Comparative risk information and decision-making about prevention. *Theoretical Medicine and Bioethics* 30:199–213.
- Schwartz, P., and E. Meslin. 2008. The ethics of information: Absolute risk reduction and patient understanding of screening. *Journal of General Internal Medicine* 6:867–70.
- Stegenga, J. 2011. Is meta-analysis the platinum standard of evidence? *Studies in History and Philosophy of Biological and Biomedical Sciences* 42:497–507.
- US Preventive Services Task Force. 2011. *Screening for breast cancer: Recommendation statement*. AHRQ Publication No. 10-05142-EF-2, November 2009. Rockville, MD: Agency for Healthcare Research and Quality. Available: <http://www.ahrq.gov/clinic/uspstf09/breastcancer/brcanrs.htm> (Accessed February 15, 2011).
- Weinstein, N. D. 1980. Unrealistic optimism about future life events. *Journal of Personal Social Psychology* 39:806–20.
- Welch, H. G., and W. Black. 2010. Overdiagnosis and cancer. *Journal of the National Cancer Institute* 102:1–9.
- Worrall, J. 2002. What evidence in evidence-based medicine? *Philosophy of Science* 69: S316–30.
- Worrall, J. 2007. Why there's no call to randomize. *British Journal for the Philosophy of Science* 58:451–88.
- Zahl, P., J. Maehlen, and H. G. Welch. 2008. The natural history of invasive breast cancers detected by screening mammography. *Archives of Internal Medicine* 168:2311–6.