JAMA Guide to Statistics and Methods

Decision Curve Analysis

Mark Fitzgerald, PhD; Benjamin R. Saville, PhD; Roger J. Lewis, MD, PhD

Decision curve analysis (DCA) is a method for evaluating the benefits of a diagnostic test across a range of patient preferences for accepting risk of undertreatment and overtreatment to facilitate decisions about test selection and use.[1] In this issue of *JAMA*, Siddiqui and colleagues[2] used DCA to evaluate 3 prostate biopsy strategies: targeted magnetic resonance/ultrasound fusion biopsy, standard extended-sextant biopsy, or a combination, for establishing the diagnosis of intermediate- to high-risk prostate cancer. Their goal was to identify the best biopsy strategy to ensure prostatectomy is offered to patients with intermediate- and high-risk tumors and avoided for patients with low-risk tumors.

Related article page 390

Use of the Method

Why Is DCA Used?

When patients have signs or symptoms suggestive of but not diagnostic of a disease, they and their physician must decide whether to (1) treat empirically, (2) not treat, or (3) perform further diagnostic testing before deciding between options 1 and 2. The decision to treat depends on how confident the clinician is that the disease is present, the effectiveness and complications of treatment if the disease is present, and the patient's willingness to accept the risks and burden of a treatment that might not be necessary. A diagnostic test may provide additional information on whether the disease is present.[3] Decision curve analysis is a method to assess the value of information provided by a diagnostic test by considering the likely range of a patient's risk and benefit preferences, without the need for actually measuring these preferences for a particular patient.[1]

A key concept in DCA is that of a "probability threshold," namely, a level of diagnostic certainty above which the patient would choose to be treated. The probability threshold used in DCA captures the relative value the patient places on receiving treatment for the disease, if present, to the value of avoiding treatment if the disease is not present. If the treatment has high efficacy and minimal cost, inconvenience, and adverse effects (eg, oral antibiotics for community-acquired pneumonia), then the probability threshold will be low; conversely, if the treatment is minimally effective or associated with substantial morbidity (eg, radiation for a malignant brain tumor), then the probability threshold will be high.

The net benefit, or "benefit score," is determined by calculating the difference between the expected benefit and the expected harm associated with each proposed testing and treatment strategy. The expected benefit is represented by the number of patients who have the disease and who will receive treatment (true positives) using the proposed strategy.

The expected harm is represented by number of patients without the disease who would be treated in error (false positives) multiplied by a weighting factor based on the patient's threshold probability. The weighting factor captures the patient's values regarding the risks of undertreatment and overtreatment. Specifically, the false-positive rate is multiplied by the ratio of the threshold probability divided by 1 − the threshold probability. For example, if the treatment threshold is 10% (0.1) for a patient with possible pneumonia, then the weighting factor applied to the number of patients without pneumonia treated in error would be 0.1/0.9, or one-ninth, minimizing the effect of false-positive results because the burden of unnecessary treatment is low. Conversely, for a patient with a brain mass that is possibly malignant, the probability threshold might be 90% (0.9), leading to a weighting factor of 0.9/0.1, or 9, and greatly increasing the effect of the risk of false-positive results with any proposed testing and treatment strategy.

Graphically, the DCA is expressed as a curve, with benefit score on the vertical axis and probability thresholds on the horizontal axis. A curve is drawn for each approach that might be taken to establish a diagnosis. Another line is drawn to show what happens when no treatment is ever given (ie, no net benefit), and another curve is drawn as if all patients receive treatment irrespective of test results. For any given patient's probability threshold, the curve with the highest benefit score at that threshold is the best choice.[1]

If one curve is highest over the full range of probability thresholds, then the associated diagnostic approach would be the best decision for all patients, regardless of individual values, and a clinician can use this approach uniformly. If the curves cross, then the optimal approach will depend on the patient's risk tolerance, expressed through their probability threshold.

What Are the Limitations of the DCA Method?

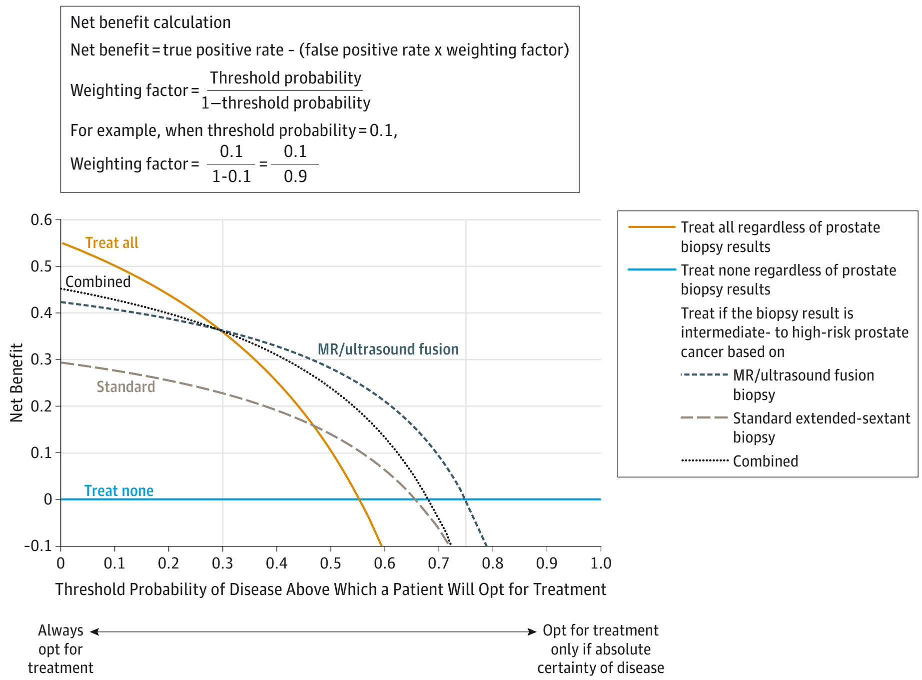
For diseases that are not well studied, there may be insufficient knowledge regarding patient preferences to determine the relevant range of threshold probabilities. Even when the likely range of probability thresholds is known, if the decision curves cross within that range, then the clinician must delve deeper into individual patient preferences to choose a testing and treatment strategy.[3]

Caution should be used in interpreting DCAs based on published ranges of threshold probabilities, particularly when there are many treatment options available to a patient. A patient is likely to have a different threshold probability if treatment is, for example, radiation rather than prostatectomy. The threshold probability needs to apply to a well-defined path of treatment.

Decision curve analysis does not explicitly account for the costs (monetary costs, time lost, physical or psychological discomfort, etc) associated with the diagnostic test. Further, if the diagnostic test provides information about how to treat as well as whether to treat (eg, a biopsy that yields both a cancer diagnosis and tumor type, allowing the selection of a specific therapy), the decision curve does not incorporate the value of this additional information.

Another challenge in correct implementation of DCA is that the data required for establishing the curve are often difficult to obtain. There must be sufficient study data for the population of in-

Figure. Net Benefit as a Function of a Threshold Probability of Intermediate- to High-Risk Prostate Cancer



Threshold probability refers to the point at which a patient considers the benefit of treatment for intermediate- to high-risk prostate cancer equivalent to the harm of overtreatment for low-risk disease and thus reflects how the patient weights the benefits and harms associated with this decision. The highest curve at any given threshold probability is the optimal decision-making strategy to maximize net benefit. Net benefit was maximized with threshold probabilities of 0%-30% by the "treat all" approach; with threshold probabilities of 30%-75%, net benefit was maximized by the targeted magnetic resonance (MR)/ultrasound fusion approach; and with 75%-100%, net benefit was maximized by the "treat none" approach. (Adapted from Supplement, Siddiqui et al.[2])

terest to whom the diagnostic test has been applied and the true state of the disease known for each patient at the time of the test. A fairly large patient study may be needed to establish estimates of traditional measures of accuracy (sensitivity, specificity).

Why Did the Authors Use DCA in This Particular Study?

There is controversy surrounding the benefits of screening and intervention relative to the costs of unnecessarily treating low-risk prostate cancers.[4,5] Justification for use of MR/ultrasound fusion–guided biopsy with ultrasound-guided biopsy to diagnose prostate cancer must be shown to benefit a broad range of patients.

How Should DCA Findings Be Interpreted in This Particular Study?

The DCA reported by Siddiqui et al[2] showed that for patients with threshold probabilities of 0% to 30%, representing a relative preference for empirical treatment, the net benefit is greatest if all patients are treated and that the diagnostic tests do not add sufficient information to improve care (**Figure**). In this range of threshold probabilities, patients appear to be more concerned about missing a diagnosis of cancer than about receiving unnecessary treatment. For midrange threshold probabilities of 30% to 75%, the targeted biopsy approach is superior to other strategies, including the 2 other diagnostic approaches evaluated. For higher thresholds (>75%) at which patients may be more concerned about unnecessary treatment than missed cancer, the option to not treat is preferred and neither diagnostic test has value.

Caveats to Consider When Looking at Results Based on DCA

One shortcoming of this study was the use of a subset of 170 patients who underwent prostatectomy in constructing the DCA. These patients self-selected for prostatectomy after learning the results of their targeted and standard biopsies. This group primarily comprised men who had higher cancer risk, resulting in potential bias when estimating false positives, false negatives, and other diagnostic measures. The patients classified as low risk who still opted for prostatectomy are patients with low probability thresholds, who might also be different from the broader population of men with symptoms or findings suggesting prostate cancer.

ARTICLE INFORMATION

Author Affiliations: Berry Consultants, Austin, Texas (Fitzgerald, Saville, Lewis); Department of Biostatistics, Vanderbilt University School of Medicine, Nashville, Tennessee (Saville); Department of Emergency Medicine, Harbor-UCLA Medical Center, Torrance, California (Lewis); Los Angeles Biomedical Research Institute, Torrance, California (Lewis); David Geffen School of Medicine, University of California, Los Angeles (Lewis).

Corresponding Author: Roger J. Lewis, MD, PhD, Department of Emergency Medicine, Harbor-UCLA Medical Center, 1000 W Carson St, Bldg D9, Torrance, CA 90509 (roger@emedharbor.edu).

Conflict of Interest Disclosures: All authors have completed and submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest and none were reported.

REFERENCES

1. Vickers AJ, Elkin EB. Decision curve analysis. *Med Decis Making*. 2006;26(6):565-574.

2. Siddiqui MM, Rais-Bahrami S, Turkbey B, et al. Comparison of MR/ultrasound fusion–guided biopsy with ultrasound-guided biopsy for the diagnosis of prostate cancer. *JAMA*. doi:10.1001/jama.2014.17942.

3. Sox HC, Higgins MC, Owens DK. *Medical Decision Making*. 2nd ed. West Sussex, UK: John Wiley & Sons; 2013.

4. Froberg DG, Kane RL. Methodology for measuring health-state preferences. *J Clin Epidemiol*. 1989;42(4-7):345-354.

5. Hoffman RM. Clinical practice: screening for prostate cancer. *N Engl J Med*. 2011;365(21):2013-2019.