Annals of Epidemiology xxx (2018) 1–7

Contents lists available at ScienceDirect

Annals of Epidemiology

journal homepage: www.annalsofepidemiology.org

Original article

Relative impact characteristic curve: a graphical tool to visualize and quantify the clinical utility and population-level consequences of implementing markers

Mohsen Sadatsafavi, MD, PhD [a, b, c, *], Paul Gustafson, PhD [d], Zafar Zafari, PhD [e], Don D. Sin, MD [b]

[a] Faculty of Medicine and Faculty of Pharmaceutical Sciences, University of British Columbia, Vancouver, British Columbia, Canada
[b] Faculty of Medicine, University of British Columbia, Vancouver, British Columbia, Canada
[c] Centre for Clinical Epidemiology and Evaluation, Vancouver Coastal Health Institute, Vancouver, British Columbia, Canada
[d] Department of Statistics, University of British Columbia, Vancouver, British Columbia, Canada
[e] Department of Health Policy and Management, Columbia University, New York City, NY

ARTICLE INFO

Article history:
Received 4 January 2018
Accepted 23 July 2018
Available online xxx

Keywords:
Biomarkers
Receiver operating characteristics
Epidemiological methods

ABSTRACT

Purpose: Receiver operating characteristic (ROC) curve analysis is a popular method for evaluating the performance of (bio)markers. However, the standard ROC curve does not directly connect marker performance to patient-related outcomes. Our aim was to fill this gap by proposing a conceptually similar graphical tool that carries information about the clinical uitility of markers.

Methods: We propose a novel graphical tool, the relative impact characteristic (RIC) curve, that depicts the trade-off between the population-level impact of treatment as a function of the size of the treated population for a given marker positivity rule (e.g., a threshold). We establish analogies between the ROC and the RIC curves around the interpretations of shape, slopes, and area under the curve and discuss parametric inference on RIC.

Results: As a case study, we used data from a clinical trial on preventive therapy for exacerbations of chronic obstructive pulmonary disease. We illustrate how the RIC curve can be constructed for a predication score and be interpreted in terms of a marker's ability toward concentrating treatment benefit in the population. We discuss hoe the RIC curve can be used to identify a threshold on the risk score based on the maximal acceptable number-needed-to-treat.

Conclusions: The RIC curve enables evaluation of markers in terms of their treatment-related clinical utility. Its analogies with the standard ROC analysis can facilitate its interpretation, bringing a population-based perspective to the activities of diverse marker development and evaluation teams.

Background

Attempts in finding and implementing novel markers to inform therapeutic choices or other medical decisions have been a major focus of biomedical and clinical research. Recently, under the purview of Precision Medicine and with the arrival of vast "omics" data, the discovery and implementation of (bio)markers have been accelerated [1]. The promise is that markers enable tailoring of disease management to the needs of patients to maximize therapeutic benefit and minimize harm [1]. In this sense, a good marker can be seen as one that results in the "concentration of benefit" from subsequent treatment decisions.

Receiver operating characteristic (ROC) curve and its associated metrics such as the area under the curve (AUC) are ubiquitous in communicating the performance characteristics of markers [2]. The classical ROC curve plots the true positive rate (sensitivity) versus the false positive rate (one minus specificity) of a marker in detecting a binary disease state as a function of marker threshold [3]. Although originally developed for diagnostic markers, the ROC analysis can also be applied to prognostic markers that provide information on the likely course of the disease, as well as predictive

Don Sin is a Tier 1 Canada Research Chair in chronic obstructive pulmonary disease. Mohsen Sadatsafavi is supported by a New Investigator Salary Award from the Canadian Institutes of Health Research and a Scholar Award from Michael Smith Foundation for Health Research.

* Corresponding author. Centre for Clinical Epidemiology and Evaluation, University of British Columbia, 7th Floor, 828 West 10th Avenue, 2405 Wesbrook Mall, Vancouver, Canada V6T 1Z3. Tel.: +1 604 875 5178; fax: +1 604 875 5179.
E-mail address: msafavi@mail.ubc.ca (M. Sadatsafavi).

https://doi.org/10.1016/j.annepidem.2018.07.014

markers that provide information on the response to therapy [4]. The AUC provides an objective summary of marker performance, independent of any threshold [5].

As a marker advances through developmental stages, there is often great interest in understanding the value of its implementation with regard to health outcomes at the population level. The standard ROC analysis is concerned with the impact of a marker in detecting disease states or risks, independent of downstream consequences (e.g., what can be done once the disease is diagnosed). However, the benefit of marker implementation is to modify disease management strategies such that the subsequent treatment is provided to patients who will benefit the most from it. This important aspect of marker implementation is not captured in classical ROC analysis.

There have been recent developments in graphical demonstration of the clinical utility and population-level impact of markers [6–8]. In the predictiveness curve, the estimated risk from a risk prediction model is depicted as a function of the risk percentile [6]. As such, it enables the visualization of the classification capacity of the model (the better the model, the more convex the curve) together with its predictiveness in terms of the actual disease risk when applied to the population. The principle underlying Decision Curve Analysis is that choosing a threshold for a marker for test positivity is akin to assigning certain weights for trading off false- and true-positive results [7]. This enables calculating a measure of "net benefit" that can be visualized as a function of the chosen threshold. More recently, the Lorenz curve has been proposed to communicate the concept of "risk concentration" in the population [8]. The curve can facilitate identification of subgroups of individuals in whom disease risk is higher, which can then be targeted for identification (e.g., through screening). The associated Gini index summarizes the inequality in risk distribution into a scalar metric.

The overall purpose of the present work is to broaden the set of such graphical tools by introducing a new graphical method that visually represents the concept of the "concentration of benefit" associated with a marker when the marker is used to inform a specific treatment decision. The proposed graphical tool juxtaposes two fundamental quantities associated with marker implementation: the population-level impact of marker-informed treatment and the relative size of the treated population. Neither of these quantities are communicated through traditional ROC curves, and previous graphical methods do not jointly present these quantities. As such, the proposed method can be seen as an addition to the existing graphical methods for visually communicating various aspects of a marker performance. Our intention is to ensure similarity with the ROC framework given the familiarity of biomarker development teams and many other investigators and knowledge users with the ROC methodology. The resulting relative impact characteristic (RIC) methodology retains many analogies with the ROC analysis yet provides a population-based lens for marker implementation.

The RIC curve: a motivating example

Consider a genomic-based marker that is used to calculate a continuous risk score that predicts the likelihood of disease recurrence in women with breast cancer who are currently in remission. The score can be dichotomized using a threshold so that individuals with a score higher than the threshold would receive chemotherapy. Suppose that the provision of chemotherapy to all eligible patients results in an average improvement in the 5-year survival probability of .04 compared with usual care. Consider that we set the test threshold to the 50th percentile (median) of the marker value in the population. Assume that, given the higher recurrence rate, chemotherapy in patients who score above the threshold will now provide improvements in 5-year survival of 0.06. Therefore, at this threshold, the marker-informed treatment will result in survival improvement of $0.06 \times 0.50 = 0.03$ at a population level. That is, by using this threshold, we provide chemotherapy to 50% of eligible patients yet obtain a "relative impact" of 75% (0.03/0.04) compared with the strategy of providing chemotherapy to all.

The RIC curve illustrates this trade-off, as demonstrated in Figure 1 for this hypothetical marker. The curve is constructed by varying the threshold across its full range and plotting the proportion of individuals whose marker value lies above the threshold (a number between 0 and 1) on the X-axis, and the "relative impact", compared to the outcomes associated with treating all eligible patients, on the Y-axis (a number that is typically between 0 and 1 but theoretically can take any value). When the test threshold is set above the highest possible value, the treatment is not provided to anyone and the relative impact is 0, comprising the left-most point on the curve. When the threshold is set below the lowest possible value, all patients will receive the treatment and the test and treatment bundle has a relative impact of 1, corresponding to the right-most point on the curve. The RIC curve allows one to examine how a marker enables the concentration of benefit: the provision of treatment to a subset of patients who will derive the most benefit from it.

Definition of RIC curve

We consider a single marker (either a biochemical marker or a risk prediction score) that returns a scalar continuous value (or an ordinal value with many levels), which is used to inform a binary treatment decision. Extension of this framework to multiple treatments or sequential use of different markers is conceivable but is beyond the scope of this work. Our focus is on clinical outcomes such as risk or rate of events. We briefly explain in the Discussion section the implication of using policy-related outcomes that consider both the benefits and harms of treatment and marker measurement but exclude this from the focus of the present work. To ease the derivations, and while not strictly necessary, we make the following two assumptions. First, we assume that the marker positivity rule will be based on a threshold so that only individuals who score higher than the threshold will receive the treatment. While we focus on a threshold-based rule for marker positivity, the RIC, in theory, can be constructed using more general rules (e.g., as in Janes et al. [9]) as long as the rule has some free parameters that can result in a varying proportion of the population being eligible for treatment. Second, we assume that treatment outcomes in one patient do not affect disease risk or outcomes in other patients, as would be the case, for instance, in infectious diseases due to disease transmission and herd immunity.

Consider $p\ (x)$ as the proportion of the population that receives the treatment at the threshold value of x on the marker: $p(x) = \int_x^{+\infty} f(y).dy$ with $f\ (.)$ being the probability distribution of marker values in the population. Let $b\ (x)$ be the expected treatment effect (compared with no treatment) for individuals with a marker value of x. The average treatment effect when provided to all is $\overline{b} = \int_{-\infty}^{+\infty} b(y).f(y).dy$. The relative impact, denoted by $q\ (x)$, is the expected effect of treatment in individuals who receive the treatment over the expected effect in all patients: $q(x) = \int_x^{+\infty} b(y).f(y).dy/\overline{b}$. The RIC curve is $q\ (x)$ as a function of $p\ (x)$ as x varies across its entire range.

Please cite this article in press as: Sadatsafavi M, et al., Relative impact characteristic curve: a graphical tool to visualize and quantify the clinical utility and population-level consequences of implementing markers, Annals of Epidemiology (2018), https://doi.org/10.1016/j.annepidem.2018.07.014