Published in final edited form as:

Clin Gastroenterol Hepatol. 2019 February ; 17(3): 560–562.e2. doi:10.1016/j.cgh.2018.05.002.

Lorenz Curves and Gini Coefficient Analyses Indicate Inefficiencies in Esophageal Adenocarcinoma Screening

Chin Hur^{*,‡}, Tiannan Zhan^{*}, Aaron P. Thrift[§], Thomas L. Vaughan^{||}, and Eric J. Feuer[¶]

^{*}Gastroenterology Division, Massachusetts General Hospital, Boston, Massachusetts; [‡]Harvard Medical School, Boston, Massachusetts; [§]Section of Gastroenterology and Hepatology, Department of Medicine, Baylor College of Medicine, Houston, Texas; ^{||}Program in Epidemiology, Fred Hutchinson Cancer Research Center, Seattle, Washington; [¶]Surveillance Research Program, Division of Cancer Control and Population Sciences, National Cancer Institute, Bethesda, Maryland

Rates of esophageal adenocarcinoma (EAC) have increased rapidly in the United States and much of western Europe, and 5-year survival continues to be poor.¹ Prevention and early detection efforts for EAC have focused on identifying persons with EAC precursor state, Barrett's esophagus, but the survival benefit has been disappointingly low.

Precise risk stratification could improve early identification of individuals who will develop EAC, allowing for such high-risk individuals to be targeted for new and evolving preventive and minimally invasive screening technologies. This underlying issue of efficiency is conceptually similar to the Gini coefficient, a widely accepted method to depict the income distribution of a group to assess wealth inequality. A Lorenz curve plots percentage of wealth against the percentage of the population; perfect wealth distribution would result in a slope of 1 and a Gini coefficient of 0, whereas severe wealth inequity results in higher Gini values closer to 1. Some examples of Gini coefficients are shown in Supplementary Figure 1. In this article, we use relative risks of stratified groups to construct Lorenz curves and quantitatively analyze the distribution of EAC burden in the US population, for which a higher Gini coefficient indicates more efficient risk stratification.

Methods

Lorenz curves and Gini coefficients were constructed for 3 strategies: (1) perfect compliance² and (2) current compliance to using only symptomatic gastroesophageal reflux disease (GERD) to triage patients for endoscopic EAC screening²; and (3) potential risk stratification using 5 risk factors including race, sex, smoking status (never smoker vs ever

Reprint requests: Address requests for reprints to: Chin Hur, MD, 101 Merrimac Street, 10th Floor, Institute for Technology Assessment, Boston, Massachusetts 02114. chur@mgh.harvard.edu; fax: (617) 726-4103.

Supplementary Material

Note: To access the supplementary material accompanying this article, visit the online version of *Clinical Gastroenterology and Hepatology* at www.cghjournal.org, and at <https://doi.org/10.1016/j.cgh.2018.05.002>.

Conflicts of interest

The authors disclose no conflicts.

smoker), obesity (body mass index categories with normal, overweight, and obese), and GERD symptoms that resulted in 48 subgroups. For the third strategy, we estimated the relative risk for EAC and percent population of the subgroups from published data^{3–5} and databases.^{6–8} Calculation details and assumptions can be found in Supplementary Methods and Supplementary Table 1.

Results

Figure 1A presents Lorenz curves with cumulative EAC cases on the x-axis and cumulative US population on the y-axis. First, we depicted the Lorenz curves that are consistent with the previously recommended strategy targeting the presence of GERD symptoms as the primary risk factor to triage patients for endoscopic screening. The yellow curve represents perfect adherence with a Gini coefficient of 0.39, where 17.5% of the population with symptomatic GERD can be used to identify approximately 56% of EAC cases. The red curve depicts the GERD screening strategy with current adherence rates with Gini coefficient of 0.05, where 7% of EAC cases are identified within approximately 2% of the population.

The potential risk-stratification method, depicted in green, to improve the identification of high-risk patients using 5 risk factors (race, sex, smoking status, obesity, and GERD symptoms) resulted in a Gini coefficient of 0.68. Figure 1B lists the 48 subgroups representing all combinations of the 5 risk factors, ordered from highest to lowest relative risk of EAC compared with the reference risk group (nonwhite women who do not smoke, normal weight, without GERD symptoms). Each row with diminishing relative risk in Figure 1B (from top to bottom) corresponds to a point on the green curve of the Lorenz curve in Figure 1A (from left to right). Some specific green points are highlighted because they reflect unique profiles, such as the first nonwhite group, first group without GERD symptoms, and first female group.

Discussion

Although there is no correct or established cutoff, one intuitive or visual inflection point is where 80% of patients who will develop EAC are identified, achieved by targeting 29% of the population (shaded in darker green). To identify 90% of EAC cases would require targeting more than 42% of the population.

Our Lorenz curve and Gini coefficient calculations not only highlight the inefficiency of current practice of GERD-only centric strategy, but also exhibit the potential synergic effects among risk factors. For example, the first female group in Figure 1B, with 4 other risk factors present (smoking, obese, GERD symptoms, and white) would still be considered having higher risk of EAC compared with the white male smoking group. The improvements in efficiency shown by our results could be further enhanced with the addition of any independent molecular biomarkers that are currently an area of intense research.

Although our analysis underscores the potential for more efficient screening, it does not suggest or shed light on the optimal screening regimen that should be used or how effective such screening will be. Our observations also serve as a call for more quantitative studies

and analyses on risk stratification to guide targeted screening for EAC and other less common cancer screening.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

References

1. Hur C, Miller M, Kong CY, et al. Trends in esophageal adenocarcinoma incidence and mortality. *Cancer* 2013;119: 1149–1158. [PubMed: 23303625]
2. Vaughan TL, Fitzgerald RC. Precision prevention of oesophageal adenocarcinoma. *Nat Rev Gastroenterol Hepatol* 2015; 12:243–248. [PubMed: 25666644]
3. Hoyo C, Cook MB, Kamangar F, et al. Body mass index in relation to oesophageal and oesophagogastric junction adenocarcinomas: a pooled analysis from the International BEACON Consortium. *Int J Epidemiol* 2012;41:1706–1718. [PubMed: 23148106]
4. Cook MB, Kamangar F, Whiteman DC, et al. Cigarette smoking and adenocarcinomas of the esophagus and esophagogastric junction: a pooled analysis from the International BEACON Consortium. *J Natl Cancer Inst* 2010;102:1344–1353. [PubMed: 20716718]
5. Cook MB, Corley DA, Murray LJ, et al. Gastroesophageal reflux in relation to adenocarcinomas of the esophagus: a pooled analysis from the Barrett's and Esophageal Adenocarcinoma Consortium (BEACON). *PLoS One* 2014;9:e103508. [PubMed: 25075959]
6. The International Barrett's and Esophageal Adenocarcinoma Consortium. Available at: <http://beacon.tlvnet.net/>. Accessed November 1, 2016.
7. Surveillance, Epidemiology, and End Results (SEER) Program. SEER*Stat Database: Incidence - SEER 18 Regs Research Data + Hurricane Katrina Impacted Louisiana Cases, Nov 2015 Sub (2000–2013) <Katrina/Rita Population Adjustment> - Linked To County Attributes - Total U.S., 1969–2014 Counties, National Cancer Institute, DCCPS, Surveillance Research Program. Released April 2016, based on the November 2015 submission. Available at: www.seer.cancer.gov. Accessed November 1, 2016.
8. Centers for Disease Control and Prevention (CDC), National Center for Health Statistics (NCHS). National Health and Nutrition Examination Survey Data (2005–2015). Hyattsville, MD: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention Available at: <https://wwwn.cdc.gov/nchs/nhanes/>. Accessed May 1, 2016.

Figure 1.

(A) Potential risk stratification strategies of EAC cases in US population. The *red curve* represents the GERD screening strategy with current adherence rates, where 7% of EAC cases are identified within approximately 2% of the population. Gini coefficient = 0.05. The *yellow curve* represents perfect adherence where 17.5% of the population with symptomatic GERD can be used to identify approximately 56% of EAC cases. Gini coefficient = 0.39. The *green curve* represents the potential strategy to improve the identification of high-risk patients that incorporates 5 risk factors: race, sex, smoking status (never smoker vs ever smoker), obesity (BMI categories), and GERD symptoms. Gini coefficient = 0.68. (B) Stratified risk groups ranked by relative risk of EAC. Subgroups in US population stratified by 5 risk factors (smoking, BMI, GERD, race, and sex) and ranked by relative risk of EAC to the reference group (*last row*). The relative risk is calculated using proportion and odds ratio of EAC of each risk group, depicted in *blue bar*. BMI, body mass index; G, Gini coefficient.