Response to letter to the editor regarding “External validation of a lung cancer-based prediction model for two-year mortality in esophageal cancer patient cohorts” by M. Berbée et al.

Berbée, M.; Muijs, C. T.; Schuit, E.

Published in:
Radiotherapy and Oncology

DOI:
10.1016/j.radonc.2024.110176

Publication date:
2024

Radiotherapy and Oncology 193 (2024) 110176

Contents lists available at ScienceDirect

Radiotherapy and Oncology

journal homepage: www.thegreenjournal.com

Letter to the Editor

Response to letter to the editor regarding "External validation of a lung cancer-based prediction model for two-year mortality in esophageal cancer patient cohorts" by M. Berbée et al.

We thank Populaire et al. for their interest in our study [1]. The authors raise several points, including on the reported association between the mean heart dose (MHD) and 2-year mortality, the representativeness of our patient cohort and delineation of the gross tumor volume, and suboptimal performance of the neo-adjuvant chemoradiotherapy (nCRT) model at the University Hospital Leuven. Below, we will address each of these points separately.

Populaire et al. state that the MHD is no longer associated with 2-year mortality after adjusting for performance status. In our opinion this statement is too strong. Multivariable analysis showed a consistent association with 2-year mortality in terms of the size of the effect (OR between 1.24 and 1.43) while correcting for different sets of prognostic factors (e.g., gross tumor volume (GTV), age, gender, pathology, N-stage, performance status and tumor length) and across different populations (nCRT patients, definitive chemoradiotherapy (dCRT) patients, and the combined cohort of nCRT + dCRT patients). In the nCRT cohort, indeed, the p-value rises above 0.05, but the size of the effect is still very much consistent with lower levels of adjustment (e.g., OR of 1.35 after full adjustment vs. 1.43 after adjusting for GTV only). We believe that this p-value, and thus associated width of the confidence interval, is more related to the sample size, rather than non-existence of an association between the MHD and 2-year mortality. Similarly, the much smaller study cohort of the authors compared to ours (n = 91 vs. n = 568) may explain why the authors found such high p-values for the association of the GTV and MHD with 2-year mortality. Unfortunately, and although recommended [2], the authors do not present any actual effect estimates (e.g., odds ratios) with these p-values, which makes it difficult to assess the associations they found and to what extent the associations in their cohort are different from ours.

The authors question the representativeness of our study population as nearly 20 % of the nCRT patients did not proceed to surgery, which is indeed higher than reported in the CROSS study the authors refer to [3]. However, it is well known that patients included in clinical studies do not necessary reflect clinical practice [4]. Moreover, clinical practice changed since publication of the CROSS study in 2015 and may vary between countries; we used FDG-PET/CTs for re-staging, which was not the case in the CROSS study, the number of patients requesting organ preserving strategies has risen since the publication of the CROSS study, and decision counseling for active surveillance after neoadjuvant chemoradiotherapy has become common in some regions. As such, we believe the patients included in our study are more representative of current clinical practice than those included in the CROSS study. Additionally, most nCRT patients in our study had a performance status of 0, 1 or 2; and only 5/474 (1 %) had a reported performance status of 3 or 4, unlike the 6 % as indicated in the letter by Populaire et al.; it is therefore unlikely that these few patients substantially impacted the results.

The GTVs in the study cohorts were indeed delineated before implementation of the current guideline [5]. However, we question the suggestion that the observed differences in GTV could be explained by this fact, as, in our experience, clinical introduction of the guideline has mainly affected the delineation of the clinical target volume (CTV) opposed to the delineated volume of the GTV.

Populaire et al. show that the nCRT patients model performs poorly at external validation in their University Hospital Leuven population. It is very difficult to fully identify the potential reasons for this poor performance based on the limited information provided in the letter. However, there are several aspects that stand out. Considering the expert status of the UZ Leuven in the treatment of esophageal cancer patients, the number of 91 patients included over a 7-year period (2015–2021) seems low with slightly over 10 patients per year. This low number of included patients, combined with the smaller GTV of patients in their cohort may indicate they may have selected a lower risk population. Additionally, the number of 91 patients is not considered sufficient to allow proper model validation [6], which is exactly the reason why we, in our original publication [1], refrained from emphasizing the validation results in the individual cohorts, and particularly the NKI-AvL cohort, which was of similar size than the Leuven cohort. As such, we feel the authors overstate and overinterpret the results of their underpowered validation study. We do welcome any effort to, potentially jointly, externally validate the nCRT and dCRT models in larger patient cohorts in the future.

Using larger cohorts, we will also be able to further improve the models, e.g., by including other characteristics relevant for the prediction of 2-year mortality, or characteristics related to surgery technique or recovery pathway as highlighted by Populaire et al.. This is important as the key aim in the modelling of complications following radiation is to find the most relevant (combination of) dose parameters that drive the increased risk for the outcome. Several efforts are made to collect such cohorts. First, we have initiated the MODELS project (prediction models for complications to improve radiotherapy plan optimization and patient selection techniques). In this consortium project, patient, treatment, and outcome data of over 2000 patients will be collected and used to develop and validate multivariable models for various relevant patient outcomes. In this project we will be able to account for heterogeneity in patients across different centers, e.g., using internal-external cross-validation methods as performed in the construction of the 2-year mortality model used in the national indication protocol for proton therapy in lung cancer patients [7], to develop models that perform consistently across various populations. Similar activities can be performed in a second cohort, the PROTRAIT registration in which all patients who are treated with proton radiotherapy in the Netherlands are

https://doi.org/10.1016/j.radonc.2024.110176
Received 13 February 2024; Accepted 16 February 2024
Available online 18 February 2024

included and for whom patient, tumor, and treatment characteristics as well as complications and treatment outcomes are being collected prospectively.

To summarize, we do agree with the notion that further improvement of the models in our paper is possible, but we do not fully agree with the points that were raised by Populaire et al. and feel that the conclusions based on their external validation study are overstated and likely invalid as they were based on an underpowered (and potentially selected) study.

CRediT authorship contribution statement

M. Berbee: Writing original draft. **C.T. Muijs:** Writing original draft. **E. Schuit:** Writing original draft, Methodology.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] Berbee M, Muijs CT, Voncken FEM, et al. External validation of a lung cancer-based prediction model for two-year mortality in esophageal cancer patient cohorts. Radiot Oncol 2023;190:109979.
[2] Wasserstein RL, Lazar NA. The ASA Statement on p-Values: Context, Process, and Purpose. Am Stat 2016;70:129–33.
[3] Shapiro J, van Lanschot JJB, Hulshof M, et al. Neoadjuvant chemoradiotherapy plus surgery versus surgery alone for oesophageal or junctional cancer (CROSS): long-term results of a randomised controlled trial. Lancet Oncol 2015;16:1090–8.
[4] Higgins LM, Lester-Coll NH, Ades S, et al. 'CROSS'-ing into the 'Real World': a retrospective cohort study of patients receiving trimodality and bimodality therapy for esophageal cancer. J Gastrointest Oncol 2023;14:480–93.
[5] Thomas M, Mortensen HR, Hoffmann L, et al. Proposal for the delineation of neo-adjuvant target volumes in oesophageal cancer. Radiot Oncol 2021;156:102–12.
[6] Riley RD, Snell KIE, Archer L, et al. Evaluation of clinical prediction models (part 3): calculating the sample size required for an external validation study. BMJ 2024;384:e074821.
[7] NVRO. Landelijk Indicatie Protocol Protonen Therapie Longcarcinoom; 2019. https://nvro.nl/images/documenten/rapporten/LIPP_longen_final_01122019.pdf.

M. Berbée *
Department of Radiation Oncology (Maastro), GROW School for Oncology and Reproduction, Maastricht University Medical Centre+, Maastricht, The Netherlands

C.T. Muijs
Department of Radiation Oncology, University Medical Center Groningen, Groningen, The Netherlands

E. Schuit
Department of Epidemiology & Health Economics, Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht University, The Netherlands

* Corresponding author.
E-mail address: maaike.berbee@maastro.nl (M. Berbée).