# Disease-free survival as a surrogate endpoint for overall survival in early-stage resected esophageal/gastroesophageal junction cancer: An analysis of SEER-Medicare data

*Jaffer A. Ajani[1], Lisa Leung[2], Steve Kanters[2], Prianka Singh[3], Murat Kurt[3], Inkyu Kim[3], Mir-Masoud Pourrahmat[2], Howard S. Friedman[4], Prakash Navaratnam[4], Gregory Reardon[5]*

[1] The University of Texas MD Anderson Cancer Center, Houston, United States of America; [2]Evidinno Outcomes Research Inc., Vancouver, Canada; [3]Bristol Myers Squibb, Lawrenceville, United States of America; [4]DataMed Solutions LLC, New York, United States of America; [5]Keck Graduate Institute, Claremont, United States of America

## Introduction

- Overall survival (OS) is the gold standard efficacy measure in oncology; however, obtaining OS data in early-stage cancers poses challenges such as requiring longer follow up times and larger trial populations[1,2]
- A common way to address these challenges is to use statistically appropriate and clinically relevant surrogate endpoints, such as disease-free survival (DFS), that can alleviate the uncertainty in long-term survival[3]
- We have previously evaluated DFS as a surrogate endpoint for OS by modeling the association between the treatment effects, as hazard ratios (HRs) on each endpoint, using aggregate-level data from randomized controlled trials (RCTs) in adults with resectable esophageal cancer (EC) and gastroesophageal junction cancer (GEJC) receiving therapy in the neoadjuvant, adjuvant, and perioperative settings[4]
  - Results suggested a considerable correlation between DFS and OS (Pearson's correlation coefficient 0.83, 95% CI: 0.70-0.90)
  - However, we recognize that patient cohorts from clinical trials may be different as compared to those from the real-world setting. Analysis of real-world data can complement clinical trial evidence by increasing generalizability to larger and more inclusive patient populations, health care providers, and health care systems, as well as to settings that more realistically reflect day-to-day practice.[5]
- Hence, to complement the findings of our previous correlation meta-analysis using RCT data we have set out to evaluate the appropriateness of DFS as a surrogate endpoint for OS using data from the National Cancer Institute's Surveillance, Epidemiology, and End Results-Medicare linked database (SEER-Medicare)

## Objective

To evaluate the appropriateness of DFS as a surrogate endpoint for OS in adults with resected EC/GEJC

## Methods

**Patient Selection**

Inclusion criteria:

- Primary diagnosis of stage 2 or stage 3 EC/GEJC from 2009 to 2017
- Patients ≥ 66 years of age on index date
- Resectable surgery

Exclusion criteria:

- Patients who had an earlier primary malignant cancer diagnosis
- Did not survive the surgical hospitalization
- Had a poor performance status (ECOG score 3-4)
- Did not receive tumor resection surgery on or after the index admission date
- Underwent surgery after 31 December 2017 (to ensure sufficient follow-up time for the time-to-event outcomes to develop)
- Did not have continuous enrollment in both Medicare plans A and B, with no health maintenance organization enrollment, from 1 year prior to and until the index date
- Did not receive chemotherapy in at least one of the pre- or post-index periods

**Statistical Methods**

- Surrogacy was evaluated in two stages:
  - **Individual-level correlation** between the endpoints assessed according to Spearman's rank correlation (we also computed Kendall's tau estimates from copula models as supportive measures), and
  - **Treatment-level correlation** between the treatment effects on DFS and OS assessed according to Pearson's correlation through weighted linear regression (WLR)

## Methods (continued)

- The treatment-level correlation required clusters of similar patients to obtain numerous HR estimate pairings.
  - Patients were classified in clusters based on treatments they received and baseline characteristics (e.g., age, sex, index year, staging, and race/ethnicity)
  - Propensity score matching addressed imbalances in baseline characteristics between the treatment and control groups in the constructed clusters
- Validation of the WLR model was tested via two approaches:
  - Internally via leave-one-out-cross-validation (LOOCV) as described by National Institute for Health and Care Excellence (NICE)[6]
  - Externally via predictions made for 26 RCTs identified from our previous literature-based correlation meta-analysis in EC/GEJC[4]
- The utility of the WLR model was evaluated with the surrogate threshold effect (STE), which is the minimum DFS benefit that would translate into statistically significant OS benefit. Calculation of STE is dependent on sample size; for the STE presented here we used the sample size of a representative RCT in this disease area.[7]

## Results

**Patient characteristics**

- The included 925 patients were mainly elderly males of non-Hispanic White ethnicity with relatively even split of stage 2 and stage 3 cancer (**Table 1**)

**Treatment Characteristics**

- From the 925 patients, 766 received neoadjuvant therapy (82.8%) and 218 received adjuvant therapy (23.6%)
  - In both settings, most patients received multi-agent chemotherapy: 705/766 (92%) for neoadjuvant, and 178/218 (81.7%) for adjuvant (**Table 2**)

**Outcome Characteristics**

- Median time-to-events were approximately 28.2 months for OS and 9.8 months for DFS
- Median follow-up time was 26.2 months
  - Follow-up times for single-agent chemotherapy is substantially longer at a median of 34.1 months
- Recurrence events were mostly observed up to approximately 24 months (**Figure 1**)

**Individual-level Correlation Analyses**

- Spearman's ρ was estimated to be 0.76 (95% CI: 0.70, 0.89), suggesting a considerable correlation between DFS and OS
- Supplementary Kendall's tau estimates from copula models ranged between 0.62 and 0.79 (**Table 3**)

**Treatment-level Correlation Analyses**

- Pearson's correlation estimate was 0.96 (95% CI: 0.75, 1.00), suggesting a strong correlation between the treatment effects on DFS and OS
- The computed surrogacy equation was $\log(HR_{OS}) = 0.02 + 1.09 \times \log(HR_{DFS})$ with a corresponding STE of 0.86, meaning a minimum value of 0.86 on the $HR_{DFS}$ is necessary to predict a positive and statistically significant treatment effect on the $HR_{OS}$ (**Figure 2**)

**Internal validation**

- The 95% prediction intervals generated by the model contained the raw OS HRs for 10/11 (91%) clusters in LOOCV (**Figure 3**)

**External validation**

- The 95% prediction intervals generated by the model contained 24/27 (89%) of the observed OS HRs from 26 RCTs in EC/GEJC (**Figure 4**)

**Results in context of previous surrogacy work**

- Findings using SEER-Medicare data were comparable to the previously conducted correlation meta-analysis using aggregate-level data from RCTs (**Table 4**)

**Table 1: Overall patient characteristics**

| Characteristic | Categories | Overall |
|---|---|---|
| Sample Size | | 925 |
| Age (median; $Q_1$, $Q_3$) | Years | 71.8 (68.8, 75.6) |
| Sex (n; %) | Male | 777 (84%) |
| Surgery Index year (median; $Q_1$, $Q_3$) | Years | 2013 (2011, 2015) |
| Clinical staging (n; %) | 2 | 466 (50.4%) |
| | 3 | 459 (49.6%) |
| Race/Ethnicity (n; %) | Non-Hispanic White | 826 (89.3%) |
| | Non-Hispanic Black | 15 (1.6%) |
| | Hispanic | 43 (4.6%) |
| | Asian or Pacific Islander | 4 (0.4%) |
| | American Indian/Alaska Native | 37 (4%) |
| | Other or Unknown | 0 (0%) |

*$Q_1$ first quartile, $Q_3$ third quartile*

**Table 2: Overall treatment characteristics**

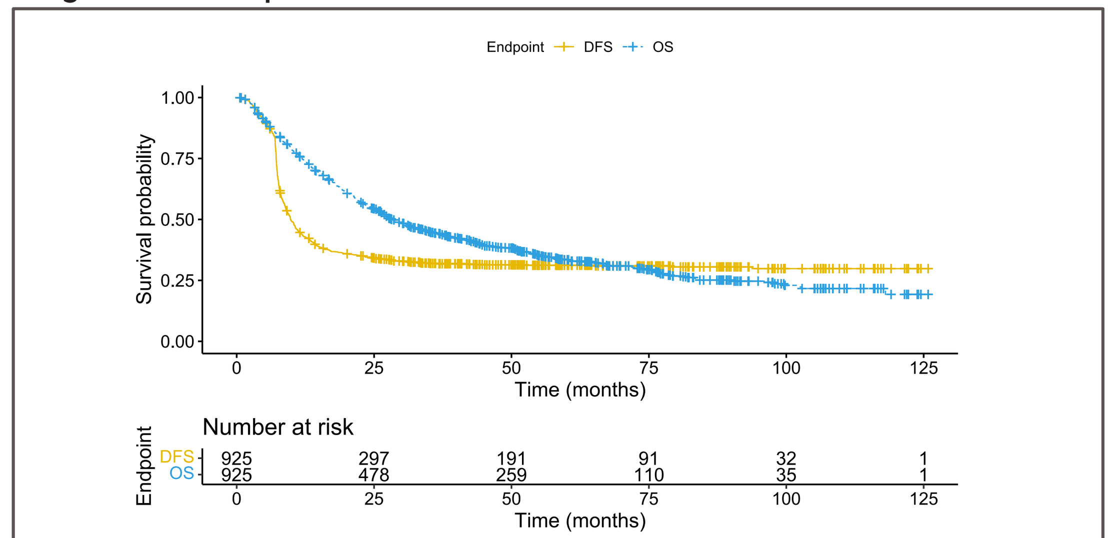
| Characteristic | Categories | Overall |
|---|---|---|
| Sample Size | | 925 |
| Neoadjuvant therapy (n; %) | Single-agent chemotherapy | 61 (6.6%) |
| | Multi-agent chemotherapy | 705 (76.2%) |
| | None | 159 (17.2%) |
| Number of cycles in neoadjuvant setting (n; %) | 1 | 23 (3%) |
| | 2 | 159 (20.8%) |
| | 3 | 88 (11.5%) |
| | 4 | 88 (11.5%) |
| | 5+ | 408 (53.3%) |
| Duration in neoadjuvant setting (months) (median; $Q_1$, $Q_3$) | Duration of treatment cycle | 1.05 (0.92, 1.31) |
| Adjuvant therapy (n; %) | Single-agent chemotherapy | 40 (4.3%) |
| | Multi-agent chemotherapy | 178 (19.2%) |
| | None | 707 (76.4%) |
| Number of cycles in adjuvant setting (n; %) | 1 | 11 (5%) |
| | 2 | 6 (2.8%) |
| | 3 | 15 (6.9%) |
| | 4 | 26 (11.9%) |
| | 5+ | 160 (73.4%) |
| Duration in adjuvant setting (months) (median; $Q_1$, $Q_3$) | Duration of treatment cycle | 6.05 (2.79, 7.92) |

*$Q_1$ first quartile, $Q_3$ third quartile*

**Table 3: DFS-OS individual-level correlation using copula models**

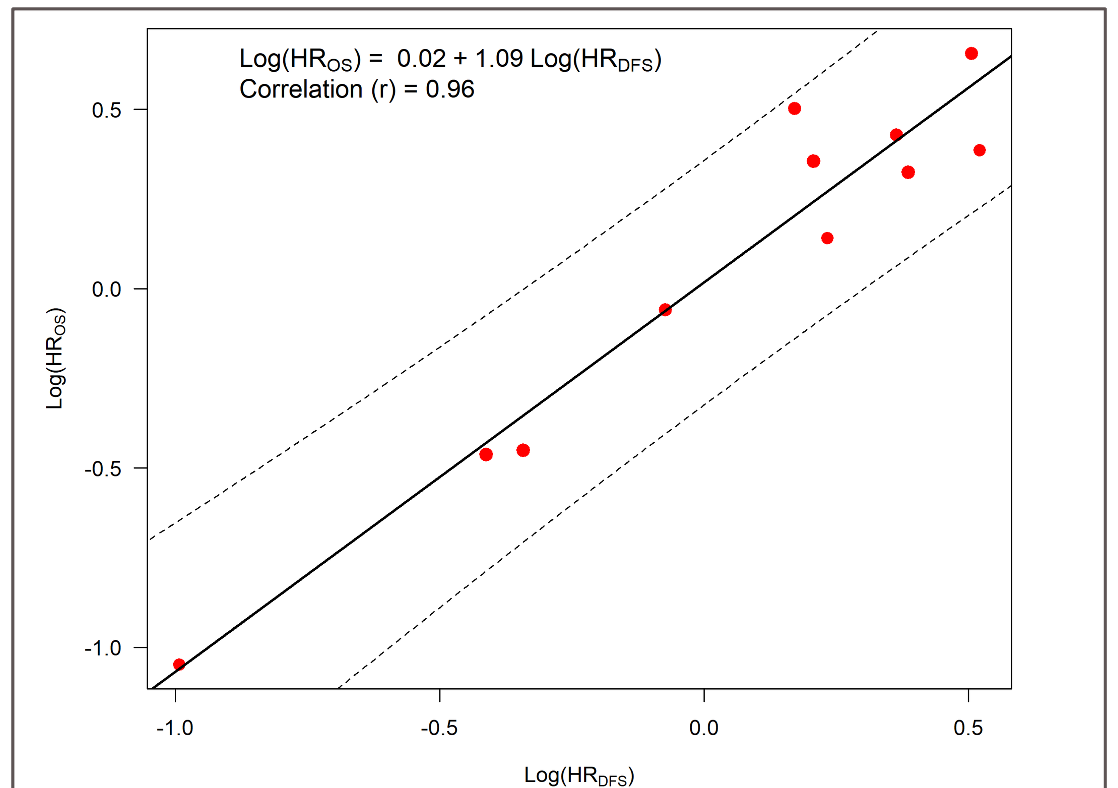
| Copula Model | Estimate (95% CI) |
|---|---|
| Clayton | 0.62 (0.56, 0.66) |
| Gumbel | 0.79 (0.77, 0.81) |

**Figure 1: Raw Kaplan-Meier curves of DFS and OS**



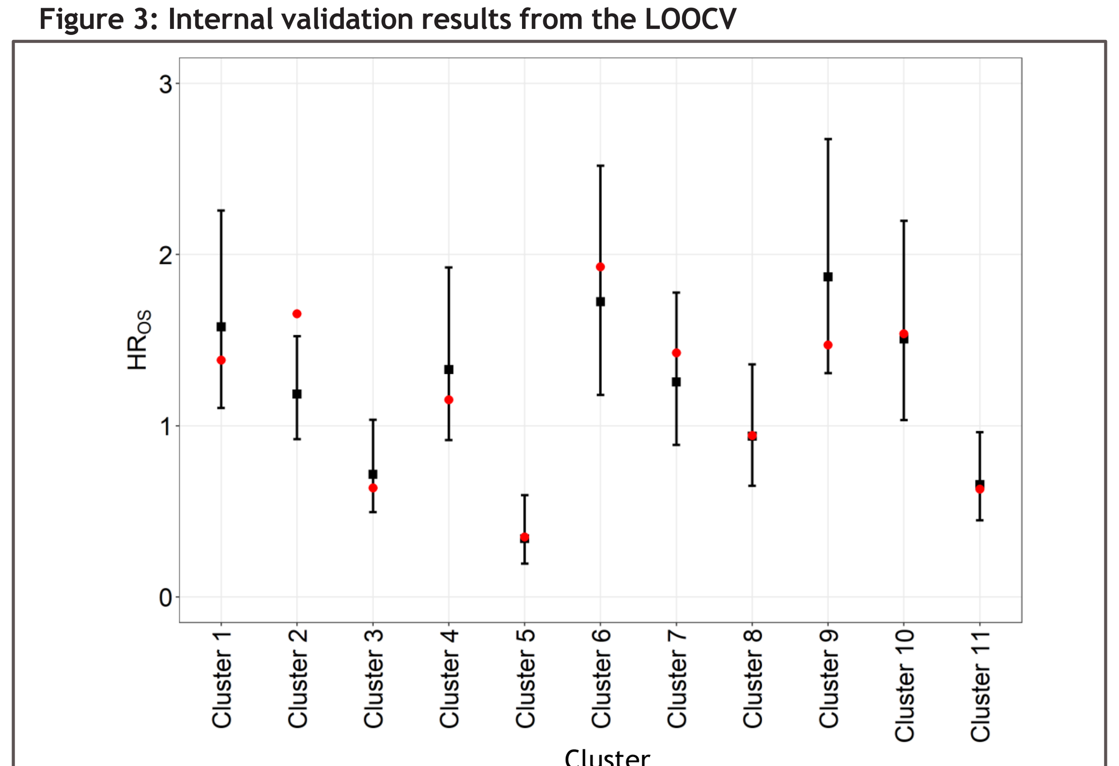

*Note: The Kaplan-Meier curves for DFS and OS endpoint were plotted in yellow and blue correspondingly. Markings of crosses represent censoring on the curves. Abbreviations: OS - Overall survival; DFS - Disease-free survival.*

**Figure 2: WLR model on treatment comparisons developed for treatment-level correlation**



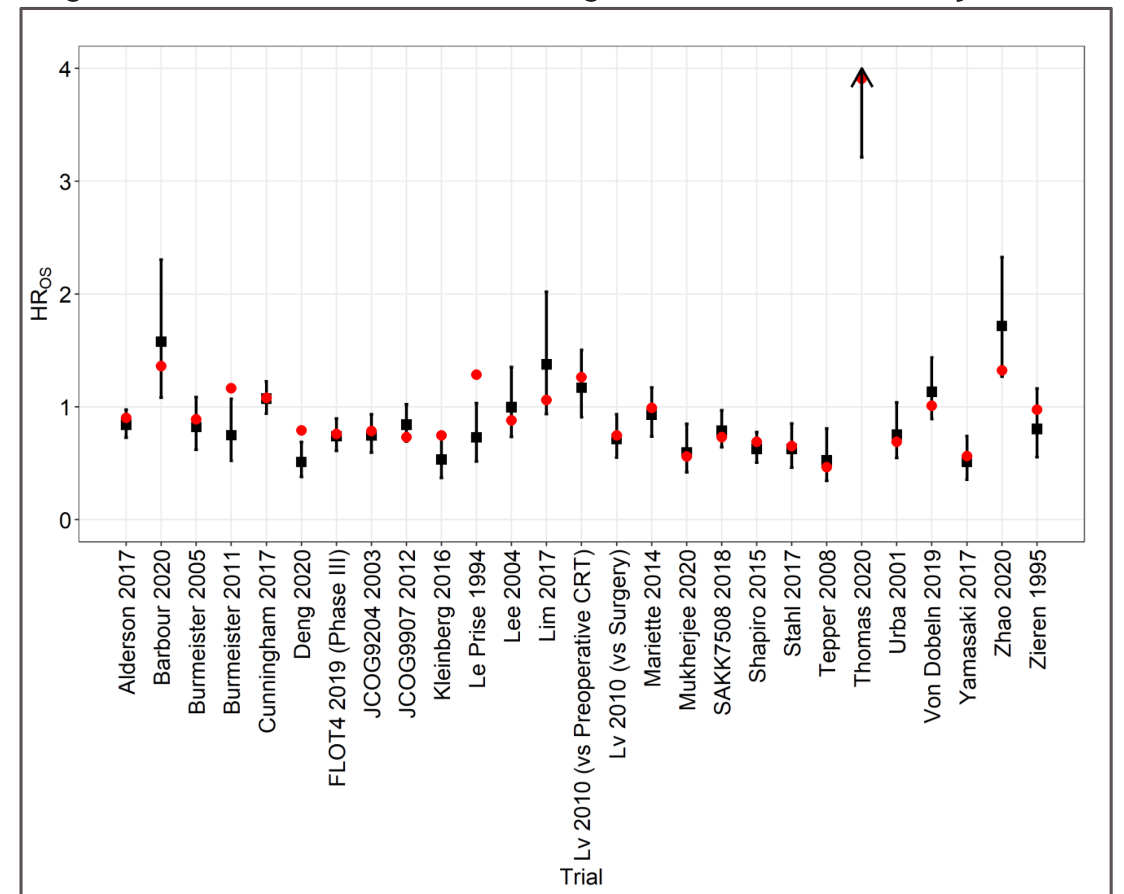

***Legend**: The weighted linear regression is graphed as a solid straight line with its corresponding 95% predictive interval boundaries as dotted curved lines. The red dots are plotted using measures from the reported $\log(HR_{DFS})$ on the x-axis against measures of $\log(HR_{OS})$ on y-axis for each treatment comparison per treatment cluster. Sizes of the dots correspond to the weights associated within the surrogacy equation. Abbreviations: HR - Hazard ratio; OS - Overall survival; DFS - Disease-free survival.*

**Figure 3: Internal validation results from the LOOCV**




***Legend**: The overall results of the LOOCV are presented in the above graph. In this figure, the red dots are the reported $HR_{OS}$ measures and the black dots are the predicted $HR_{OS}$ measures. Corresponding to the predicted $HR_{OS}$ measures, the 95% prediction intervals are presented as black solid lines per each cluster. To assess the cross-validation results, reported $HR_{OS}$ measures should fall within the plotted interval. Abbreviations: HR - Hazard ratio; OS - Overall survival.*

**Figure 4: External validation results using RCT evidence base from Ajani et al.[4]**




***Legend:** The black dots and lines represent the predicted HR of OS and its 95% prediction interval. The red dots represent what was observed in the reported publication. The outlier, Thomas 2020, had an estimated prediction interval upwards that was extended past the graph due to an extremely high $HR_{OS}$ measure and a relatively small sample size (n=30 patients). Nevertheless, the actual $HR_{OS}$ remains relatively close to the predicted $HR_{OS}$ and is within the prediction interval.*

**Table 4: Comparison with surrogacy work using RCT evidence**

| Analysis set | Correlation Estimate (95% CI) | STE | LOOCV |
|---|---|---|---|
| Current analysis using SEER-Medicare data | 0.96 (0.75, 1.00) | 0.86 | 10/11 (91%) |
| Correlation meta-analysis from Ajani et al. using RCT evidence[4] | 0.89 (0.76, 0.95) | 0.82 | 26/27 (96%) |

*CI - confidence interval; STE - surrogate threshold effect; LOOCV - leave-one-out-cross-validation*

## Conclusions

- This analysis shows consistent and robust correlations between DFS and OS, and between the treatment effects on these endpoints
- The surrogacy equation between the treatment effects of DFS and OS displayed high predictive accuracy and utility (assessed by the validation, and the STE)
- Findings support DFS as a surrogate for OS in early-stage EC/GEJC treatments in the real-world setting

## References

1. Lux MP, Ciani O, Dunlop WCN, Ferris A, Friedlander M. The Impasse on Overall Survival in Oncology Reimbursement Decision-Making: How Can We Resolve This? Cancer management and research. 2021;13:8457-8471.
2. Pazdur R. Endpoints for assessing drug activity in clinical trials. Oncologist. 2008;13 Suppl 2:19-21.
3. FDA. Clinical Trial Endpoints for the Approval of Cancer Drugs and Biologics Guidance for Industry. December 2018. https://www.fda.gov/media/71195/download
4. Ajani JA, Leung L, Singh P, et al. Disease-free survival as a surrogate endpoint for overall survival in adults with resectable esophageal or gastroesophageal junction cancer: A correlation meta-analysis. Eur J Cancer. 2022;170:119-130.
5. Sherman RE, Anderson SA, Dal Pan GJ, et al. Real-World Evidence — What Is It and What Can It Tell Us? New England Journal of Medicine. 2016;375(23):2293-2297.
6. Bujkiewicz S, Achana F, Papanikos T, Riley R, Abrams K. Multivariate meta-analysis of summary data for combining treatment effects on correlated outcomes and evaluating surrogate endpoints. NICE DSU technical support document. 2019;20.
7. Kelly RJ, Ajani JA, Kuzdzal J, et al. Adjuvant Nivolumab in Resected Esophageal or Gastroesophageal Junction Cancer. New England Journal of Medicine. 2021;384(13):1191-1203.

## Acknowledgments

- All authors contributed to and approved the presentation; writing assistance was provided by Evidinno Outcomes Research Inc., funded by Bristol Myers Squibb.
- We thank Mariana Machado, of Evidinno Outcomes Research Inc., for supporting the poster development.

Presented at the Gastrointestinal Cancers Symposium (ASCO-GI); January 17-19, 2023; San Francisco, CA and Virtual

Contact: lleung@evidinno.com