A Generalized Framework for Eliciting Unreported Subgroup-Specific Events and Survival Outcomes from Aggregate Level Data: Analyses & Insights from Adjuvant & Advanced Stage Gastrointestinal Cancers

PT21

Oguzhan Alagoz, PhD[1], Matthew Dixon, PharmD, PhD[2], Prianka Singh, PharmD[2], Murat Kurt, PhD[2]

[1]University of Wisconsin, Madison, WI, USA; [2]Bristol Myers Squibb, Lawrenceville, NJ, USA

Background

- Survival heterogeneity across subgroups play a critical role in the evaluation of marketing authorization applications by regulatory agencies (e.g. European Medicines Agency, Food and Drug Administration) and reimbursement dossiers by health technology assessment (HTA) agencies (e.g. National Institute for Health and Care Excellence)
- However, analyses of randomized controlled trials (RCTs) typically report:
  - Kaplan-Meier (KM) survival curves for intervention and comparator arms for the overall trial population
  - Subgroup-specific survival information in a summarized form: forest plots displaying hazard ratios (HRs) across selected subgroups, mostly based on stratification factors
- Subgroup-specific KM curves are not extensively reported in clinical publications
- In the absence of clear HTA guidelines, subgroup-specific survival curves are essential for tailored meta-analyses & rigorous economic evaluations to address label and reimbursement restrictions
- Earlier work[1,2] provided alternative modeling frameworks to elicit subgroup-specific survival with stringent assumptions on their structural forms

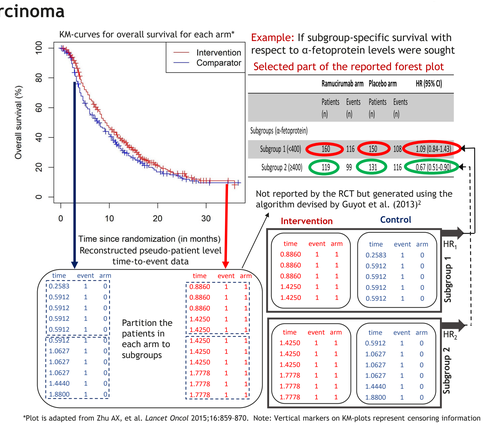
Objective

- Develop an optimization-based approach to elicit unreported subgroup-specific KM curves and underlying cumulative event data using aggregate-level RCT data

Methods

Synopsis: Probabilistically assign each individual patient to one of the two mutually exclusive subgroups in each arm so that resulting HRs for both subgroups are identical or comparable to their reported counterparts (Figure 1)

Figure 1. Schematic overview of the approach: An example from REACH trial of second-line treatment for advanced hepatocellular carcinoma

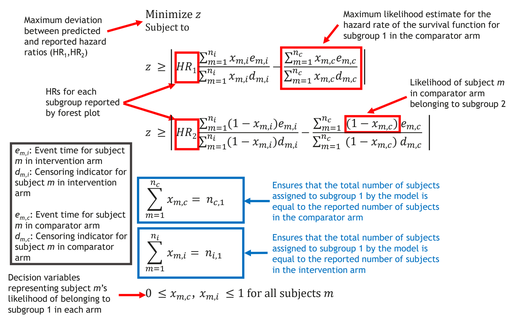


*Plot is adapted from Zhu AX, et al. *Lancet Oncol* 2015;16:859-870. Note: Vertical markers on KM-plots represent censoring information

Key features of the generalized set-partitioning approach

- It uses the closed-form maximum-likelihood estimate of the hazard rate for exponentially-distributed event/censoring times to express the HR between the arms in each subgroup. This enables the expression of HRs in a compact linearizable form for any given split of the patients between the subgroups (Figure 2)
- It minimizes the maximum deviation between the estimated HRs resulting from the assignment and the reported HRs by the forest plots across both subgroups (Figure 2)
- If the trial reports the total number of events for each subgroup in each arm, then optimization problem in Figure 2 could be formulated as a linear program (LP) – i.e., the objective function and constraints can be expressed as linear functions of the decision variables $\{x_{m,c}, x_{m,i}\}$ – with a guaranteed global optimal solution, which can be obtained in a reasonably short amount of time using an open-source software.
- However, many RCTs do not report the number of events in each arm for each subgroup. In such cases, the LP in Figure 2 is first parameterized with respect to combinations of numbers of events in either of the subgroups across the arms and then solved iteratively under all such combinations. Total number of events in each arm of the RCT determines the scale of the number of iterations for the LP.
- Across all solutions, the one from the LP with the number of events combination that provides the closest match to the reported 95% confidence intervals for subgroup-specific HRs is selected as the best-fit.
- Since the number of decision variables in the model far exceeds the number of constraints, the LP may admit multiple optimal solutions. Therefore, alternative optimal solutions need to be sampled to refine those with less desirable features.
- A secondary LP constrained to generate the minimum gap between the model-predicted and reported HRs for each subgroup as determined by the primary LP was repeatedly solved with differing objective function coefficients, which were randomly and uniformly generated between -1 and 1. The variables in the objective function of the secondary LP represented patients' probability of being in subgroup 1 in each arm.
- The ties among the alternative optima sampled by the secondary LP were broken by the following steps:
  - For each sampled solution, a Cox proportional hazards model was used to estimate the corresponding 95% confidence intervals (CIs) for each HR
  - Estimated boundaries of 95% CIs of each HR were compared to their reported counterparts from the RCT
  - The solution providing the minimum aggregate deviation between the boundaries of 95% CIs of estimated and reported HRs was favored
- When sampling alternative optima, to avoid numerical instability issues, the equality constraints that ensure the secondary LP to generate an optimal solution to the primary LP were perturbed by a pre-specified tiny error margin.

Figure 2. Optimization model



Testing the performance of the generalized set-partitioning approach

- All codes were written in R programming language; *lpSolveAPI* package of R was used for solving both primary and secondary LP models
- The method was validated using RCTs of gastro-intestinal cancers reporting subgroup-specific KM curves for overall survival (OS) in metastatic setting and different endpoints in adjuvant settings:
  - Metastatic setting: 18 distinct RCTs with 96 subgroups (Table 1)
  - Adjuvant setting: 12 distinct RCTs with 80 subgroups (Table 2)
- Predictive performance of the approach was evaluated by several measures for a total of 176 subgroups (Tables 1 and 2)
- Model-predicted number of events was compared the actual number of events for the RCTs that reported these outcomes

Table 1. List of studies from the metastatic setting included in the case study

Test case	Tumor type	Trial	Subgroup 1	Subgroup 2
1	HCC	REACH	α-fetoprotein < 400 ng/mL	α-fetoprotein ≥ 400 ng/mL
2		REACH	East-Asian	Non-East Asian
3		REACH-2	REACH-2 trial patients	REACH trial patients with α-fetoprotein ≥ 400 ng/mL
4		RESORCE	Last sorafenib dose 800 mg/day	Last sorafenib dose < 800 mg/day
5		REFLECT	Patient body weight: <60 kg	Patient body weight: ≥60 kg
6	EC/GC	KEYNOTE-181	PD-L1 CPS ≥ 10%	PD-L1 CPS < 10%*
7		KEYNOTE-181	SCC	Adenocarcinoma*
8		KEYNOTE-061	ECOG Status 0	ECOG Status 1
9		REGARD	Age < 65	Age ≥ 65
10		RAINBOW	Age < 65	Age ≥ 65
11		TAGS	With gastrectomy	Without gastrectomy
12		KEYNOTE-590	PD-L1 CPS >10	PD-L1 CPS <10
13		KEYNOTE-590	SCC	Adenocarcinoma
14		GATSBY	HER2 IHC3+:	No HER2 IHC3+: South Korea
15		SOLAR	Japan	
16	CRC	RAISE	KRAS status wild-type	KRAS status mutant
17		RAISE	Tumor side (left)	Tumor side (right)
18		CORRECT	Japanese	Non-Japanese
19		ASPECCT	Prior bevacizumab	No Prior bevacizumab
20		CRYSTAL	KRAS codon 12/13 wt Metastasis LLD	KRAS codon 12/13 wt Metastasis non-LLD
21		CRYSTAL	RAS wt Metastasis LLD	RAS wt Metastasis non-LLD
22		FRESCO	Prior targeted therapy (anti-VEGF/EGFR)	No Prior targeted therapy (anti-VEGF/EGFR)
23		FRESCO	Prior VEGFi	No Prior VEGFi
24		FRESCO	Liver metastasis	No liver metastasis
25		XELAVIRI	Male	Females

*Subgroup-specific KM curves are not reported, these subgroups were not used for validation
HCC: hepatocellular carcinoma, EC: esophageal cancer, GC: gastric cancer, CRC: colorectal cancer

Table 2. List of studies from the adjuvant setting included in the case study

No	Tumor	Endpoint	Study/ Trial	Subgroup 1	Subgroup 2
1	EC/GC	OS	Burmeister et al. (2005)[5]	SCC	Non-SCC
2		PFS	Burmeister et al. (2005)[5]	SCC	Non-SCC
3		DFS	JCOG9204	Node negative	Node positive
4		OS	SAKK75/08	AC	SCC
5		OS	CROSS	SCC	AC
6		PFS	CROSS	SCC	AC
7		OS	NeoRes I	SCC	AC
8	CRC	OS	MOSAIC	Stage II	Stage III
9		DFS	MOSAIC	Stage II	Stage III
10		OS	NSABP C-07	Age <70	Age ≥70
11		DFS	NSABP C-07	Age <70	Age ≥70
12		OS	NSABP C-07	Stage III	Stage II
13		DFS	NSABP C-07	Stage III	Stage II
14		DFS	SCOTT	FOLFOX	CAPOX
15		DFS	TOSCA	Stage II	Stage III
16		OS	S-AVANT	T1-3	T4
17		DFS	S-AVANT	T1-3	T4
18	HCC	OS	SHARP	BCLC B	BCLC C
19		OS	SHARP	Normal bilirubin	Elevated bilirubin
20		PFS	Kudo et al. (2011)[6]	Japanese	Korean

AC: adenocarcinoma; BCLC: Barcelona clinic liver cancer staging system; CAPOX: Oxaliplatin and capecitabine; CRC: colorectal cancer; DFS: disease-free survival; EC: esophageal cancer; FOLFOX: Folinic acid, fluorouracil and oxaliplatin; GC: gastric cancer; HCC: hepatocellular carcinoma; OS: overall survival; PFS: progression-free survival; SCC: squamous-cell carcinoma;

Performance Measures

For each subgroup, quality of the predictions from the model was assessed visually and evaluated statistically to the data estimated/reported from the RCTs via:

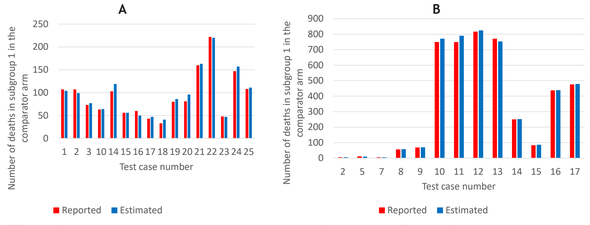
1. *Survival prediction accuracy:* % of times in which model-predicted survival rate fell within the 95% CI of the reported survival rate
2. *Restricted mean survival time (RMST) gap (ΔRMST):* % gap between the model-predicted RMST and the RMST estimated from the reported KM-curve
3. *RMST alignment:* model-predicted RMST vs. 95% CI of the RMST estimated from the reported KM-curve
4. *Average (avg.) survival rate gap [ΔS(t)]:* Avg. absolute difference between model-predicted and reported survival rates

Calculation of all performance metrics was based on the reported follow-up durations of the RCTs whereas performance metrics 1 and 4 were estimated on a monthly basis.

Results

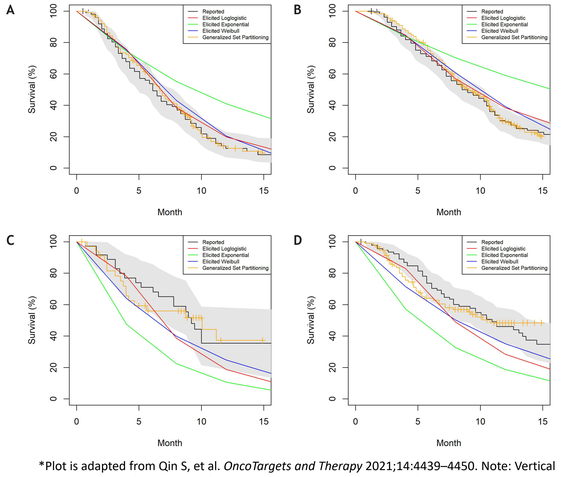
- Generalized set-partitioning approach accurately estimated number of events from subgroup 1 in each arm, largest difference between number of deaths is less than 20 (Figure 3).
- Predicted subgroup-specific survival curves were comparable to the reported curves visually (Figure 4).
- Across all subgroups, average absolute gap between predicted and reported monthly survival rates was ≤5% (Table 3).
- Predicted survival curves laid within the 95% CIs of the reported survival curves in 83% (metastatic) and 79% (adjuvant) of the time (Table 3).
- Predicted RMSTs were within the 95% CIs of their reported counterparts in 68 (metastatic) and 63 (adjuvant) subgroups (Table 3).
- Average relative gap between predicted and reported number of deaths across all subgroups was 8.5% (metastatic) and 2.3% (adjuvant) (Table 3).
- For the test cases in the metastatic setting, generalized set partitioning method performed better than RMST-based exponential and Weibull models whereas it performed comparable to the RMST-based loglogistic model.
- For the test cases in the adjuvant setting, generalized set partitioning method performed substantially better than RMST-based parametric models

Figure 3. Difference between reported and predicted number of deaths in the comparator arm among subgroup 1 population in the metastatic setting (panel [A]) and adjuvant setting (panel [B])



These charts include only test cases that reported the number of events for each subgroup

Figure 4. Example visual comparison of model predictions to the reported subgroup-specific survival (with & without liver metastases in panels [A,C] and [B,D], respectively; placebo and fruquintinib arms in panels [A,B] and [C,D], respectively) in the FRESCO trial[7] for the treatment of metastatic CRC who have progressed from second line or above chemotherapy*



*Plot is adapted from Qin S, et al. *OncoTargets and Therapy* 2021;14:4439–4450. Note: Vertical markers represent censoring information, shaded areas represent 95% CI.

Table 3. Summary & comparison of current modeling approach to the reported data & predictions from previously published parametric modeling approaches

			Performance Metrics			
Test Case	Functional Form for Subgroup Survival or modeling approach	Total number of subgroups, n	Survival prediction accuracy*, %	ΔRMST*, %	RMST alignment, n (%)	ΔS(t)*
Metastatic setting	RMST-exponential†,[1]	96	50	20	45 (47)	0.103
	RMST-Weibull†,[2]	96	73	10	66 (69)	0.054
	RMST-loglogistic†,[2]	96	80	10	70 (73)	0.046
	Generalized set-partitioning approach	96	80	10	68 (71)	0.050
Adjuvant setting	RMST-exponential†,[1]	80	30	19	22 (28)	0.141
	RMST-Weibull†,[2]	80	36	19	25 (31)	0.137
	RMST-loglogistic†,[2]	80	54	9	47 (59)	0.063
	Generalized set-partitioning approach	80	79	7	63 (79)	0.047

*: Calculated as average across all subgroups in the corresponding test case. †Indicates the assumed distribution of subgroup-specific survival from previously published work using RMST as an objective criterion.[1,2]

Conclusions

- In the absence of subgroup- and arm-specific number of events data, generalized set partitioning method provided reasonably accurate estimates.
- Proposed set partitioning framework:
  - Provides a distribution free, flexible and computationally tractable approach for eliciting unreported subgroup-specific survival from aggregate level RCT data by generating easy-to-interpret solutions.
  - Is scalable to settings with missing survival data for more than two subgroups
  - Enables indirect efficacy comparisons and meta-analyses across subgroups
- Compared to precedent models that elicit subgroup-specific survival using RMST as an objective criterion and assuming parametric forms (exponential, Weibull & loglogistic) for the survival distributions of subgroups, the set partitioning approach:
  - Generated highly robust and more precise results with a superior performance in almost all performance metrics
  - Produced visually more plausible outcomes while maintaining critical censoring information across all subgroups enabling the derivation of confidence bands for subgroup-specific survival curves
  - Was able to incorporate 95% CI information on the HRs of both subgroups reported by the forest plots into curve selection
  - Was free of convergence issues during optimization and the need for a pre-specified threshold follow-up time for the definition of its objective function.

References

1. Alagoz O, et al. *Value in Health 24 (2021):S3.*
2. Alagoz O, et al. *Value in Health 25 (2022):S10.*
3. Zhu AX, et al. *Lancet Oncology* 2015;16:859-870.
4. Guyot P, et al. *BMC Med Res Met* 2012; 12: 1-13.
5. Burmeister BH, et al. *Lancet Oncology* 2005;6:659-668.
6. Kudo M, et al. *European J of Cancer* 2011;47:2117-2127.
7. Qin S, et al. *Oncotargets and Therapy* 2021;14:4439-4450.

Acknowledgments

• This study was supported by Bristol Myers Squibb.
• All authors contributed to and approved the presentation.

Presented at ISPOR 2024; May 5-8, 2024; Atlanta, GA

Email: alagoz@engr.wisc.edu