Bridging the Gap: Demographic Reweighting of Real-World Survival Data Versus RCTs in Acute Myeloid Leukemia

Emma van Eijndhoven, MSc, MA, Dhanish Revanth, MPH, Ambarish J. Ambegaonkar, PhD
APPERTURE LLC, Marlboro, NJ, USA

Corresponding Author: ambi@apperturehealth.com

APPERTURE
FOCUSED VALUE

RWD155

BACKGROUND

- Randomized controlled trials (RCTs) are the gold standard for evaluating treatment efficacy and serve as primary evidence for regulatory, health technology assessment, and payer decision-making[1]
- Strict eligibility criteria and differences in demographics and clinical characteristics between trial and real-world populations can limit the generalizability of RCT findings[2]
- Prior work shows that emulating RCTs using real-world data (RWD) does not necessarily eliminate differences in outcomes[2-6]
- Demographic reweighting has been widely used to align RCT and real-world populations, yet whether this translates into comparable survival outcomes remains unclear
- This challenge is especially relevant in Acute Myeloid Leukemia (AML), a rapidly progressing disease, which has been associated with age and racial disparities in trial enrollment[7]

OBJECTIVE

- To evaluate whether demographic reweighting of Surveillance, Epidemiology, and End Results (SEER) data aligns survival outcome with RCT standard of care (SoC) arms across three AML RCTs

METHODS

Construct RCT cohorts

Identify ALM RCTs and reconstruct RCT SoC cohorts

- Three Phase III AML RCTs with reported OS, SoC control arms, and published Kaplan-Meier curves were identified: ADMIRAL [Xospata (gilteritinib)][9], AGILE [Tibsovo (ivosidenib)][10], and VIALE-A [Venclexta (venetoclax)][11] (**Table 1**)
- **Patient-level data for SoC arms were reconstructed from KM curves** using the Guyot et al. (2012) method[8], as SoC arms are more representative of real-world ALM treatment patterns

Construct SEER cohort

Construct SEER cohort

- SEER data from 2014-2017 were used to enable up to 5 years of follow-up for the real-world AML cohort
- Patients receiving routine SoC chemotherapy were included, while those without follow-up data were excluded

Reweight SEER cohort

Reweight SEER to match RCT demographic characteristics

- SEER data were **reweighted using a raking approach**, which preserves the original sample size while **matching available baseline age, sex, and race distributions** to the RCT SoC arms
- SEER data were selected for reweighting rather than RCT data because patient-level demographic information required for reweighting was available in SEER but not in the reconstructed RCT cohorts

Compare survival

Compare survival across cohorts

- **Restricted Mean Survival Time (RMST)**, which measures average survival time **up to a fixed timepoint (24 months)**, was used to enable direct survival comparison across trials with differing follow-up periods
- RMST estimates were compared across RCT SoC, SEER Unweighted, and SEER Weighted cohorts for each RCT

Table 1. Description of selected AML RCTs used in this study

Drug *(Generic)*	RCT Name	RCT Period	Endpoint (Median OS)	Median OS	Median Follow-up
Xospata *(gilteritinib)*	ADMIRAL[9]	2015-2018	Primary	5.6 months	17.8 months
Tibsovo *(ivosidenib)*	AGILE[10]	2017-2021	Secondary	7.9 months	15.1 months
Venclexta *(venetoclax)*	VIALE-A[11]	2017-2021	Primary	9.6 months	20.5 months

OS: Overall Survival; RCT: Randomized Controlled Trial

Table 2. Baseline characteristics of RCT SoC and SEER cohorts before and after demographic reweighting by trial

Demographic Characteristics	SEER Unweighted	ADMIRAL Xospata (gilteritinib)		AGILE Tibsovo (ivosidenib)		VIALE-A Venclexta (venetoclax)	
		RCT SoC	SEER Weighted	RCT SoC	SEER Weighted	RCT SoC	SEER Weighted
	n = 5,639	N = 124	n = 5,639	n = 74	n = 5,639	n = 145	n = 5,639
Age*							
≥ 65	55%	40%	40%				
≥ 75	26%			58%	57%		
Mean (SD)	62.5 (18.2)					75 (5.7)	62.5 (18.2)
Sex							
Female	42%	57%	57%	51%	51%	40%	40%
Race							
White	79%	60%	61%	35%	35%	75%	75%
Black	10%	6%	6%	6%	6%	1%	1%
Asian	10%	27%	27%	56%	56%	23%	23%
Other	1%	7%	6%	3%	3%	1%	1%

RCT: Randomized Controlled Trial; SD: Standard Deviation, SEER: Surveillance, Epidemiology, and End Results; SoC: Standard of Care
* Age variables were reported differently across RCTs. VIALE-A reported only mean age (SD) rather than age categories; as such reweighting for VIALE-A was limited to sex and race

RESULTS

Study Cohorts and Baseline Characteristics

- A total of 124, 74, and 145 patients were included in the RCT SoC cohorts for ADMIRAL, AGILE, and VIALE-A, respectively
- The SEER cohort included 5,639 patients with AML from 2014-2016
- **Baseline demographic differences were observed** between RCT SoC cohorts and Unweighted SEER cohort (**Table 2**)
  - Patients were **younger in ADMIRAL**, and **older in AGILE and VIALE-A** relative to SEER
  - **Female representation was higher in ADMIRAL and AGILE**, while comparable in VIALE-A to SEER
  - **Trials** enrolled a **higher proportion of Asian** and a **lower proportion of Black patients**

Demographic Reweighting

- Demographic reweighting **improved alignment** between SEER and RCT SoC demographic characteristics across each RCT (**Table 2**)
- Age-based reweighting was not feasible for VIALE-A due to missing age categorization

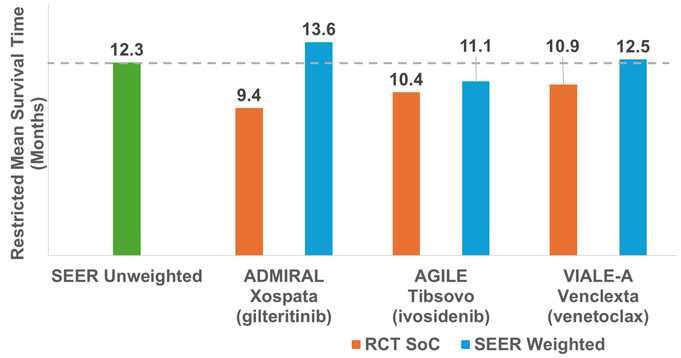
Survival Comparison

- **RMST was longer** in Unweighted SEER compared to RCT SoC cohorts across all RCTs (**Fig 2**)
- Reweighting **did not consistently reduce differences** in survival between SEER and RCT SoC cohorts
  - Survival differences increased for **ADMIRAL (RMST Δ: +1.3 months)** and **VIALE-A (RMST Δ: +0.2 months),** albeit minimally for the latter
  - Survival differences decreased for **AGILE (RMST Δ: -1.2 months)**

LIMITATIONS

- The RCTs enrolled global populations, which may limit comparability with the US-based SEER population due to differences in clinical practice patterns
- SEER does not include key clinical variables needed to replicate trial eligibility criteria (e.g., disease severity, treatment history), which may lead to residual confounding
- Clinical Trial study periods did not fully overlap with the SEER study period (2014-2016)
- RCT SoC cohorts were reconstructed from published KM curves using the Guyot et al. (2012) method,[8] which may introduce approximation error in survival estimates

Fig 2. RMST at 24 months across RCT SoC arm, SEER (Unweighted), and SEER (Weighted) cohorts by trial



RCT: Randomized Controlled Trial; SEER: Surveillance, Epidemiology, and End Results; SoC: Standard of Care

CONCLUSIONS & FUTURE RESEARCH

- **Mean survival time was longer in the real-world AML population** compared to RCT SoC cohorts
- **Demographic reweighting aligned baseline demographic distributions** between SEER and RCT SoC populations but **did not consistently reduce differences in survival across RCTs**
- Future studies should **incorporate clinical confounders** and other **sources of heterogeneity** to improve the applicability of RCT evidence to real-world populations

REFERENCES

1. Wilson, B. E. et al, (2024).https://doi.org/10.1016/j.eclinm.2024.102915
2. Orcutt, X. et al, (2025). https://doi.org/10.1038/s41591-024-03352-5
3. Colnet, B. et al, (2025) https://doi.org/10.1093/jrsssa/qnae043
4. Incerti, D. et al, (2024) https://www.ispor.org/publications/journals/value-outcomes-spotlight/vos-archives/issue/view/heor-innovation-and-impact/a-new-frontier-in-health-economics-and-outcomes-research--generalizing-clinical-trial-outcomes-to-real-world-settings
5. Heyard, R. et al, (2023) https://doi.org/10.1101/2023.07.13.23292601
6. Vincent, D. et al, (2025) https://www.ispor.org/heor-resources/presentations-database/presentation-cti/ispor-europe-2025/poster-session-2-2/comparing-individualized-treatment-effect-inference-methods-through-a-simulation-study
7. Loeb, N., et al, (2025) https://doi.org/10.1016/j.bneo.2025.100070
8. Guyot, P. et al, (2012) https://doi.org/10.1186/1471-2288-12-9
9. Perl, A.E. et al, (2019) https://doi.org/10.1056/NEJMoa1902688
10. Montesinos, P et al, (2022) https://doi.org/10.1056/NEJMoa2117344
11. DiNardo, C.D et al, (2020) https://doi.org/10.1056/NEJMoa2012971