MSR27

Evaluating Complete Response (CR) and Progression-Free Survival (PFS) as Surrogate Endpoints (SEs) for Overall Survival (OS) in Relapsed/Refractory (R/R) Chronic Lymphocytic Leukemia (CLL)

Paul Serafini, Mir Sohail Fazeli, Mir-Masoud Pourrahmat, Murat Kurt
Evidinno Outcomes Research Inc., Vancouver, BC, Canada

Background

- CLL is a type of cancer that develops in white blood cells in the blood, bone marrow, lymph nodes, or spleen.[1] It is the most common form of leukemia in adults, accounting for 25–30% of leukemia cases globally with 100,000 new cases and 40,000 deaths annually. The 5-year survival rate across all lines is 85%, with a mortality rate of 4.5%.[2]
- In the past decade, treatment of CLL has been transformed by the transition from chemo-immunotherapy to novel targeted therapies, including BTK and BCL2 inhibitors in the second-line setting and CAR T-cell therapy in the R/R setting.[3] However, there is no cure for CLL, and patients with R/R CLL experience progressively worse outcomes as they exhaust the available treatment options. This highlights a substantial unmet need.
- Although OS is considered the gold-standard endpoint by regulatory agencies when assessing new therapies in solid and liquid malignancies,[4] evaluating OS in clinical trials can be inefficient because long follow-up periods or large sample sizes are often required to demonstrate a statistically meaningful treatment effect. This may delay patient access to potentially life-saving therapies.
- SEs are intermediate measures that can be collected earlier and used to predict long-term clinical benefit of treatment in both randomized trials and in the real-world.[5] SEs must be validated by establishing biological plausibility,, individual-level association, and trial-level association between treatment effects,[5] preferably across randomized controlled trials (RCTs).[6]
- Two clinically relevant SEs in R/R CLL are CR (complete eradication of CLL/SLL in all disease compartments[7]) and PFS (time to progression or death). As CR and PFS reach statistical maturity earlier than OS, they may predict durable remission, and provide early insights on the long-term survival benefits of curative treatments for CLL.
- For example, the US FDA recently approved lisocabtagene maraleucel based on CR results from the TRANSCEND CLL 004 trial[8,9] and acalabrutinib with venetoclax based on PFS results from the AMPLIFY trial.[10]
- Study-level CR-OS association in R/R CLL has been investigated previously by two studies employing proportional hazards models,[11,12] finding significant OS improvements associated with 10% and 20% improvements in CR, respectively. Additionally, one study investigated study-level PFS-OS association in R/R CLL,[13] finding a statistically significant correlation between median PFS and OS.
- Limitations of the studies investigating CR-OS correlation at the study level include outdated evidence bases which do not reflect recent transformations in the treatment landscape,[11,13] inclusion of single-arm and observational studies,[11] treating arm-level CR rates homogeneously for all patients,[11,12] pooled use of reconstructed PFS and OS data, use of unweighted correlation measures, and the investigation of medians as opposed to hazard ratios (HRs) for treatment effects.[13]
- None of the studies employed bivariate random effects meta-analysis (BRMA). BRMA may be preferred over other surrogacy models because of its ability to capture the hierarchy between within- and between-study levels, which enables it to distinguish within- and between-study correlation, and because it can account for the variances of both endpoints simultaneously.[6]

Objectives

- **Primary Objective:** To evaluate CR and PFS as SEs for OS in R/R CLL at the trial-level and to compare their performance in earlier prediction of OS benefit.
- **Secondary Objectives:** To investigate PFS as a bridging SE between CR and OS and to evaluate the ability of CR to predict OS benefit through its association with PFS.

Methods

- The evidence base consisted of 19 RCTs reporting at least one of OS, PFS, and CR, that were identified in a systematic literature review for a surrogacy analysis in R/R CLL.[14]
- Odds ratios (ORs) of CR were calculated from summary-level event counts without continuity correction, and OS and PFS HRs were estimated by fitting Cox PH models to pseudo individual patient data (IPD) reconstructed from published survival curves.[15]
- CR-OS and PFS-OS associations were explored using four approaches:
  1. BRMA of HR_{OS} and OR_{CR} (CR → OS model)
  2. BRMA of HR_{OS} and HR_{PFS} (PFS → OS model)
  3. A two-step approach in which (1) HR_{PFS} was predicted from OR_{CR} based on a previously published model,[14] then (2) HR_{OS} was predicted from HR_{PFS} using the surrogacy equation derived from BRMA in the second approach above (CR → PFS → OS model).
  4. Previously published multivariate PH model predicting OS using proportion of patients achieving CR, proportion of patients achieving partial response (PR), and proportion of patients with greater than two lines of therapy as covariates (PH model). Unlike the BRMA models, this approach was calibrated by reconstructed time-to-event data for OS and arm-level covariate data from each trial.[12]
- BRMAs were conducted on log-transformed treatment effects, assuming a uniform between 0 and 0.2 (PFS) or −0.2 and 0 (CR) for within-study correlation, per Daniels and Hughes.[16]
- Strength of trial-level association between SEs and OS was evaluated based on the estimated between-study correlation for approaches #1 and #2.
- Predictive performance was evaluated for all four approaches by comparing the mean absolute gap between the observed and predicted HR_{OS} within the evidence bases informing each analysis set.

Results

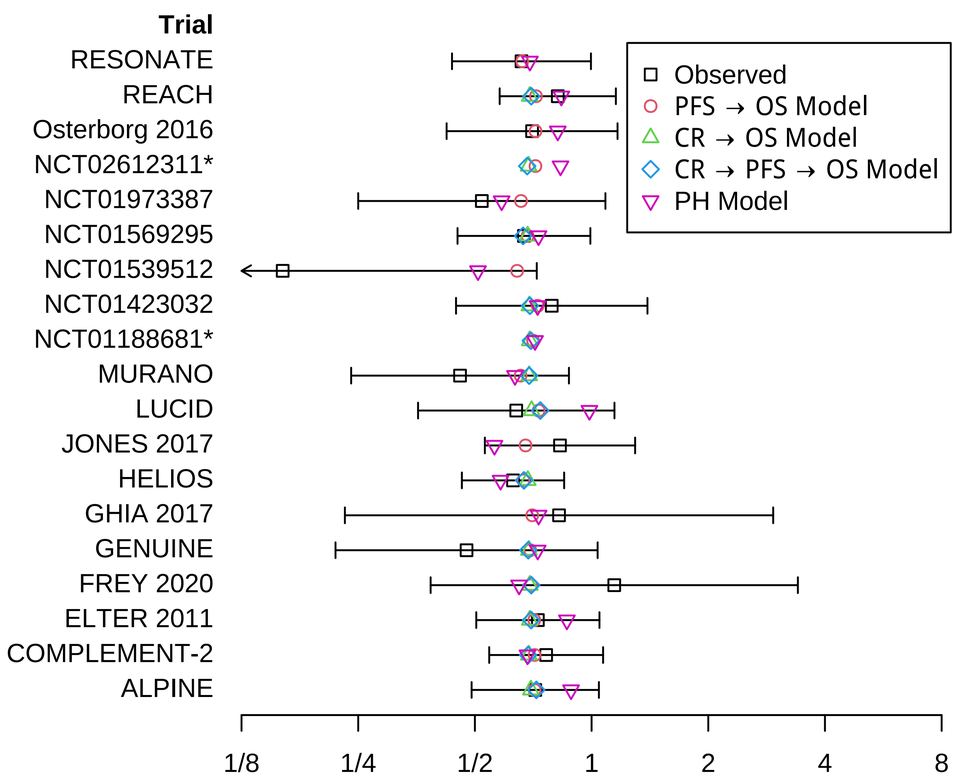
- Of 19 studies included in the evidence base, 17 had evaluable treatment effects on OS, 19 had evaluable treatment effects on PFS, and 13 had evaluable treatment effects on CR. The correlation estimates and resulting parameters from the fitted BRMA models for CR-OS and PFS-OS correlation are presented in **Table 1**. Comparisons of observed HR_{OS} versus predicted HR_{OS} obtained from all four approaches are displayed in **Figure 1**.
- **Summary of CR → OS Model**: All 17 studies with treatment effects on either OS or CR were included in the analysis. The BRMA estimated a weak, negative correlation of −0.16 (95% credible interval [CrI]: −0.95, 0.92) implying positive point estimate of association between improvements in CR and OS. The resulting surrogacy equation had an intercept of −0.36 (95% CrI: −0.59, −0.08) and a slope of −0.01 (95% CrI: −0.20, 0.15). The mean absolute gap between the observed and predicted HR_{OS} across 17 studies was 0.13.
- **PFS → OS Model**: All 19 studies were included in the analysis. The BRMA estimated a moderate, positive correlation of 0.59 (95% CrI: −0.86, 1.00) implying a positive point estimate of association between improvements in PFS and OS. The resulting surrogacy equation had an intercept of −0.31 (95% CrI: −0.48, −0.09) and a slope of 0.05 (95% CrI: −0.06, 0.31). The mean absolute gap between the observed and predicted HR_{OS} across 19 studies was 0.11.
- **CR → PFS → OS Model**: Evidence base for this two-step approach included 11 studies reporting both CR and OS. The mean absolute gap between the observed and predicted HR_{OS} across these 11 studies was 0.12.
- **PH Model**: All 17 studies reporting treatment effects on OS were included in the analysis. When HR_{OS} was predicted from the PH model using reconstructed OS data, the mean absolute gap between the observed and predicted HR_{OS} across 17 studies was 0.17.

Table 1: Correlation estimates and parameters of surrogacy equations derived from two BRMA models, and predictive performance across all models.

Analysis	Studies	Correlation (95% CrI)	Mean Gap	Intercept (95% CrI)	Slope (95% CrI)
CR → OS Model	17	−0.16 (−0.95, 0.92)	0.13	−0.36 (−0.59, −0.08)	−0.01 (−0.20, 0.15)
PFS → OS Model	19	0.59 (−0.86, 1.00)	0.11	−0.31 (−0.48, −0.09)	0.05 (−0.06, 0.31)
CR → PFS → OS Model	11	N/A	0.12	N/A	N/A
PH Model	17	N/A	0.17	N/A	N/A

Note: All intervals are 95% credible intervals (CrIs). The mean gap is the mean absolute difference between observed and predicted OS HRs.
Abbreviations: CR = Complete Response, HR = Hazard Ratio, OS = Overall Survival, N/A = Not Applicable, PFS = Progression-Free Survival, PH = Proportional Hazards.

Figure 1: Comparison of observed HR_{OS} and HR_{OS} predictions from all four models.



Trial	Observed HR (95% CI)	Best Fit (Predicted OS HR)
RESONATE	0.66 (0.44, 1.00)	PFS → OS Model (0.66)
REACH	0.82 (0.58, 1.16)	PH Model (0.84)
Osterborg 2016	0.70 (0.42, 1.17)	PFS → OS Model (0.72)
NCT02612311*		
NCT01973387	0.52 (0.25, 1.09)	PH Model (0.59)
NCT01569295	0.67 (0.45, 0.99)	CR → PFS → OS Model (0.67)
NCT01539512	0.16 (0.04, 0.72)	PH Model (0.51)
NCT01423032	0.79 (0.45, 1.39)	PFS → OS Model (0.73)
NCT01188681*		
MURANO	0.46 (0.24, 0.87)	PH Model (0.63)
LUCID	0.64 (0.36, 1.15)	CR → OS Model (0.70)
JONES 2017	0.83 (0.53, 1.30)	PFS → OS Model (0.68)
HELIOS	0.63 (0.46, 0.85)	PFS → OS Model (0.67)
GHIA 2017	0.82 (0.23, 2.94)	PH Model (0.73)
GENUINE	0.48 (0.22, 1.04)	CR → PFS → OS Model (0.69)
FREY 2020	1.14 (0.38, 3.41)	CR → PFS → OS Model (0.70)
ELTER 2011	0.73 (0.50, 1.05)	PFS → OS Model (0.71)
COMPLEMENT-2	0.76 (0.54, 1.07)	PFS → OS Model (0.71)
ALPINE	0.72 (0.49, 1.04)	PFS → OS Model (0.71)

Note: An asterisk indicates that a study did not have an observed OS treatment effect to compare to as a benchmark because they did not report an OS HR and the OS HR was not estimable from reconstructed pseudo IPD. Abbreviations: CI = Confidence Interval, CR = Complete Response, HR = Hazard Ratio, OS = Overall Survival, PFS = Progression-Free Survival, PH = Proportional Hazards.

Conclusions

- BRMAs estimated the association between OR_{CR} and HR_{OS} as weak, and the association between HR_{PFS} and HR_{OS} as moderate. Credible intervals around the estimates were wide, indicating substantial uncertainty for the association between candidate SEs and OS.
- The BRMA models associating OS directly with CR or PFS had similar predictive performance to the two-step approach using PFS as a bridge between CR and OS.
- When compared with HR_{OS} estimates, predictions from the PH model were less accurate than those from the other three approaches, despite accounting for additional information on PR, line of therapy, and temporal OS trends. This may partly be due to calculation of OS HRs from the PH model for a hypothetical population with an average PR rate and average proportion of patients with >2 lines of therapy observed in the evidence base rather than specific characteristics of each trial.
- Restricted trial evidence limited exploration of SE-OS correlations by mechanism of action (targeted vs non-targeted therapies). Further validation of broader findings is also warranted due to the limited number of trials in the evidence base.

References

1. Chen, C., et al. (2025). Cancer Control, 12, 1–10. https://doi.org/10.1177/10732748251397071
2. Pektas, G., et al. (2025). Diagnostics (Basel). 15(6), 728. https://doi.org/10.3390/diagnostics15060728
3. Wierda, W. G., et al. (2026). Journal of the National Comprehensive Cancer Network, 24(3), 68–80. https://doi.org/10.6004/jnccn.2026.0012
4. Driscoll, J. J., & Rixe, O. (2009). The Cancer Journal, 15(5), 401–405. https://doi.org/10.1097/ppo.0b013e3181bdc2e0
5. Taylor, R., & Elston, J. (2009). Health Technology Assessment, 13(8). https://doi.org/10.3310/hta13080
6. NICE (2019). Technical Support Document 20. https://sheffield.ac.uk/sites/default/files/2022-02/TSD20-mvmeta-final.pdf
7. Hallek, M., et al. (2018). Blood, 131(25), 2745–2760. https://doi.org/10.1182/blood-2017-09-806398
8. Frellick, M. (2024). Oncology News Central, March 15. https://www.oncologynewscentral.com/article/fda-approves-first-ever-car-t-therapy-in-cll-sll
9. Siddiqi, T., et al. (2023). The Lancet, 402(10402), 641–654. https://doi.org/10.1016/s0140-6736(23)01052-8
10. Syrek, R. (2026). Oncology News Central, February 15. https://www.oncologynewscentral.com/drugs/info/fda-approves-venetoclax-plus-acalabrutinib-for-previously-untreated-cll
11. Ektare, V., et al. (2016). Blood, 128(22), 2045–2045. https://doi.org/10.1182/blood.V128.22.2045.2045
12. Serafini, P., et al. (2026). Value in Health, 29(S6).
13. Beauchemin et al. (2015). Current Oncology, 22(3), 148–156. https://doi.org/10.3747/co.22.2119
14. Wang, L., et al. (2025). Leukemia Research, 158, 108113. https://doi.org/10.1016/j.leukres.2025.108113
15. Guyot, P., et al. (2012). BMC Medical Research Methodology, 12(1). https://doi.org/10.1186/1471-2288-12-9
16. Daniels, M. J. & Hughes, M. D. (1997). Statistics in Medicine, 16, 1965–1982, https://doi.org/10.1002/(SICI)1097-0258(19970915)16:17%3C1965::AID-SIM630%3E3.0.CO;2-M

Acknowledgments

This study was conducted by Evidinno Outcomes Research Inc. Paul Serafini, Mir Sohail Fazeli, and Mir-Masoud Pourrahmat report employment with Evidinno Outcomes Research Inc. Murat Kurt reports current employment with Iovance Biotherapeutics. The conduct of this research is associated with prior contractual engagement between Murat Kurt and Evidinno Outcomes Research Inc. and therefore does not represent the view of Iovance Biotherapeutics on the subject. Authors report no other conflict of interest.

EVIDINNO a UBC company

Presented at ISPOR Europe 2026; November 8-11; Vienna, Austria

Email: mpourrahmat@evidinno.com

Copies of this poster are for personal use only and may not be reproduced without written permission of the authors.