EXPLORING NON-PROPORTIONAL HAZARDS MULTI-LEVEL NETWORK META-REGRESSION (ML-NMR) FOR SURVIVAL EXTRAPOLATIONS IN ONCOLOGY

MSR104

Natalie Dennis[1], Karin Butler[2], Keyur Patel[2], Jinjie Liu[3], Chloe Spalding[2], Farhan Mughal[2]

[1]Daiichi Sankyo Oncology France SAS, Rueil-Malmaison, France, [2]Daiichi Sankyo UK Ltd., Uxbridge, United Kingdom, [3]Daiichi Sankyo, Inc., Basking Ridge, New Jersey, United States

BACKGROUND

- **Indirect treatment comparisons** (ITCs) are often used to evaluate treatments not directly compared in head-to-head trials.
- Traditional ITC methods, such as network meta-analysis (NMA) and matching adjusted indirect comparisons (MAIC), rely on the **proportional hazards** (PH) assumption for time-to-event endpoints. In such cases, a single hazard ratio (HR) may not capture the true treatment effect over time, necessitating more flexible approaches.
- **Multi-level network meta-regression** (ML-NMR) extends ITC methodology by incorporating both aggregate and individual patient data (IPD) in a network, while adjusting for treatment-effect modifiers and allowing for the baseline hazard to vary by treatment arm for **non-PH models** [1].
- ML-NMR is increasingly recognized as a tool for evaluating treatment effectiveness in **Health Technology Assessment** (HTA) [2], allowing for more comprehensive adjustments for heterogeneity and enabling effect estimates to be generated for any specified target population.

OBJECTIVES

This study explores the application of **non-PH ML-NMR** models to enhance survival extrapolations in oncology.

METHODS

- Simulated IPD and aggregate data (AgD) were used from the vignette *Example: Newly diagnosed multiple myeloma* from the *multinma* R package [3], which includes five trials comparing lenalidomide (Len) and thalidomide (Thal) to placebo (Pbo).
- The outcome of interest was **progression-free survival** (PFS).
- Fixed-effects ML-NMR models were constructed using the following distributions: **cubic M-spline**, **Weibull**, **log-logistic**, and **log-normal**.
  - ➢ To accommodate non-PH, a regression model was applied to the shape of the **baseline hazard**. For example, for the Weibull model, we included treatment effects on the shape parameter of the model. This allows for treatment-specific variations while maintaining the ability to predict **absolute treatment effects** in any target population [1].
  - ➢ Model fit was compared using the leave-one-out information criterion (**LOOIC**).
- **Shared-effect modification** was assumed between the two active treatments in the network.
- Population-average marginal **survival probabilities** were generated over a time horizon of 10 years for each target population.

RESULTS

- The network diagram can be found in **Figure 1**.
- Among the non-PH models tested, the log-logistic model had the lowest LOOIC (**Table 1**).

Figure 1. Network diagram

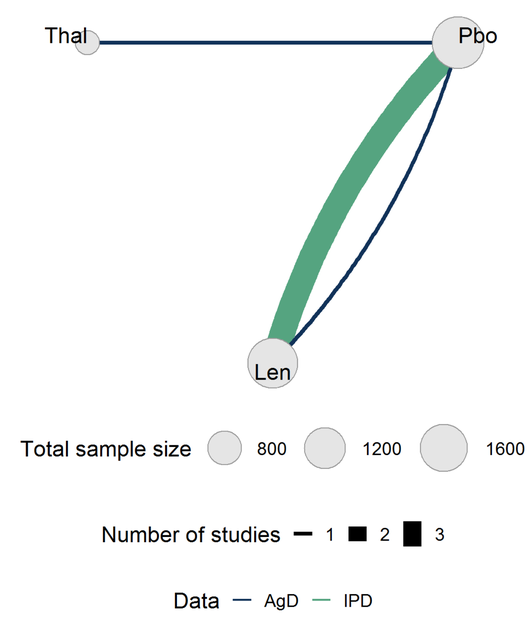

Table 1. LOOIC model comparison for each distribution

Model	LOOIC	ELPD	p_LOO
Log-logistic	24756.4	-12378.2	22.1
Log-normal	24769.1	-12384.6	23.2
Cubic m-spline	24782.3	-12391.1	38.7
Weibull	24807.9	-12403.9	22.9

LOOIC = leave-one-out information criterion; ELPD = expected log pointwise predictive density; p_LOO = effective number of parameters

- As non-PH models were prioritized, population-average conditional hazard ratios (HR) are not provided. Instead, the estimated population-average marginal survival probabilities for each distribution allowed for the HR to vary over time (**Figure 2**).
- The results show that depending on the target population, the estimated median survival times vary, ranging from a median PFS of 40.3 (Palumbo 2014) to 48.5 (Jackson 2019) for lenalidomide based on the log-logistic model (**Table 2**).

Figure 2. Population-average marginal survival probabilities for different target populations

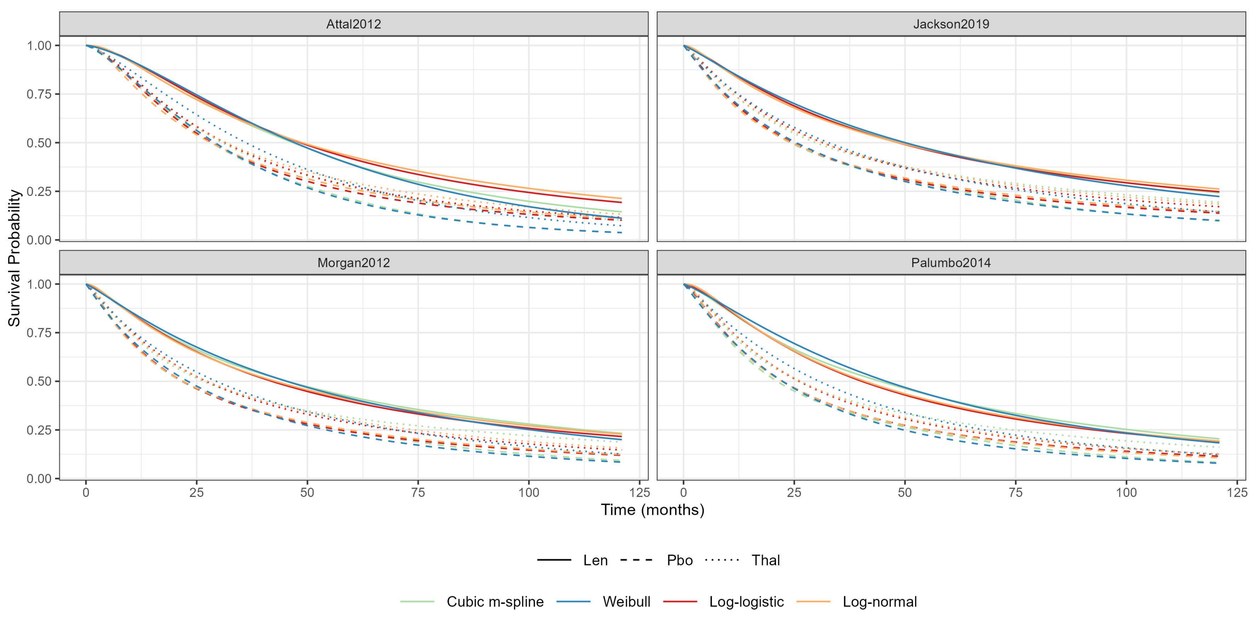

Table 2. Population-average median survival times – Log-logistic model

Study	Placebo	Lenalidomide	Thalidomide
Attal 2012	28.5 (25.5, 31.8)	48.4 (43.2, 54.3)	31.4 (24.1, 40.7)
Palumbo 2014	22.3 (18.8, 27.1)	40.3 (33.8, 48.5)	26.1 (19.7, 35.0)
Jackson 2019	24.7 (22.3, 27.4)	48.5 (44.6, 52.9)	31.4 (24.8, 39.7)
Morgan 2012	21.9 (18.8, 25.6)	42.0 (34.8, 51.3)	27.2 (23.4, 31.8)

CONCLUSIONS

- ML-NMR is a major advancement in ITCs, expanding population-adjusted methods to **larger networks** and enabling treatment effect estimation in any **target population**, which is important to align with the population of interest [2].
- **Relaxing the PH assumption** with ML-NMR allows for more flexible modelling of long-term survival, which is often needed for economic modelling. This has not been largely discussed as an advantage of ML-NMR compared to other population-adjusted methods.
- The extrapolated results demonstrated a **strong fit** to the observed simulated Kaplan-Meier (KM) data, which had follow-up times of ~50 to 75 months. For example, in Attal 2012, the median PFS was 46.3 months for Len and 28.0 months for Pbo, compared to 48.4 and 28.5 months in the log-logistic model of the ML-NMR.
- Some **limitations** and **considerations** when conducting ML-NMR include:
  - ➢**Kaplan-Meier curves** must be published for each study to conduct ML-NMR on time-to-event endpoints.
  - ➢ML-NMR is computationally intensive; **convergence issues** occurred for some parametric distributions in this case study (Gompertz).
  - ➢When IPD is limited, the **shared effect modifier** assumption is required, meaning that population-average conditional treatment effects do not vary across populations.
- This case study shows how ML-NMR can also produce **marginal treatment effects** for different target populations, thereby reducing reliance on the shared effect modifier assumption. This key feature has not been widely emphasized in current literature.
- Distinguishing between marginal and conditional treatment effects is crucial in HTA; **marginal effects** are generally preferred for policy decisions at the population level [4].

REFERENCES

[1] Phillippo D et al. "Multilevel network meta-regression for general likelihoods: synthesis of individual and aggregate data with applications to survival analysis," arXiv [pre-print], 2024.

[2] Member State Coordination Group on Health Technology Assessment, "Methodological Guideline for Quantitative Evidence Synthesis: Direct and Indirect Comparisons," 2024.

[3] Phillippo D. "multinma: Network Meta-Analysis of Individual and Aggregate Data in Stan. Version 0.7.2," R package, 2024.

[4] Remiro-Azócar A et al. " Conflating marginal and conditional treatment effects: Comments on 'Assessing the performance of population adjustment methods for anchored indirect comparisons: A simulation study'". Stat Med. 2021 May 20;40(11):2753-2758.

Disclosures: All authors are current employees of Daiichi Sankyo. CS holds stock with Daiichi Sankyo.

For more information, please contact Chloe Spalding (chloe.spalding@daiichisankyo.com) or Natalie Dennis (natalie.dennis@daiichisankyo.com