MSR110

# A comparison of Stan versus WinBUGS software for conducting Bayesian hazard ratio-based network meta-analysis

Jevdjevic M[1], Youn J-H[1], Petersohn S[1], Gittfried A[1], Ainsworth C[2], Piena M[1]

[1]OPEN Health Evidence & Access, Rotterdam, Netherlands, [2]OPEN Health Evidence & Access, Manchester, UK

OPEN HEALTH

## INTRODUCTION

- Bayesian network meta-analysis is a common tool for evidence synthesis in health sciences. Since its release in 1997, WinBUGS has become the software for conducting Bayesian network meta-analyses (NMA) in health sciences, using Markov chain Monte Carlo (MCMC) methods.[1]
- However, more efficient sampling algorithms to the WinBUGS Gibbs sampling such as Hamiltonian Monte Carlo (HMC) have been proposed as a more efficient alternative. Stan, a programming language released in 2012 and in active development today, uses a variant of HMC, the No-U-Turn Sampler (NUTS).[2] This may offer additional functionalities and ease of use over WinBUGS.
- This poster presents a case study comparing WinBUGS to Stan for the NMA of continuous data, using publicly available effectiveness data of different treatment regimens in patients living with human renal cell carcinoma (RCC).[3]

## OBJECTIVES

- Assessing the differences between and viability of WinBUGS and stan for conducting a hazard ratio (HR) NMA in RCC.
- Establishing a framework of comparison between two different software when conducting a HR-based NMA.

## METHODS

### Data

- This case study used publicly available survival data. A previously published network of evidence for overall survival (OS) of patients on first-line (1L) treatment for RCC was replicated for this case study.[3] The star-shaped network contained 6 trials and 7 treatments and can be found in Figure 1.

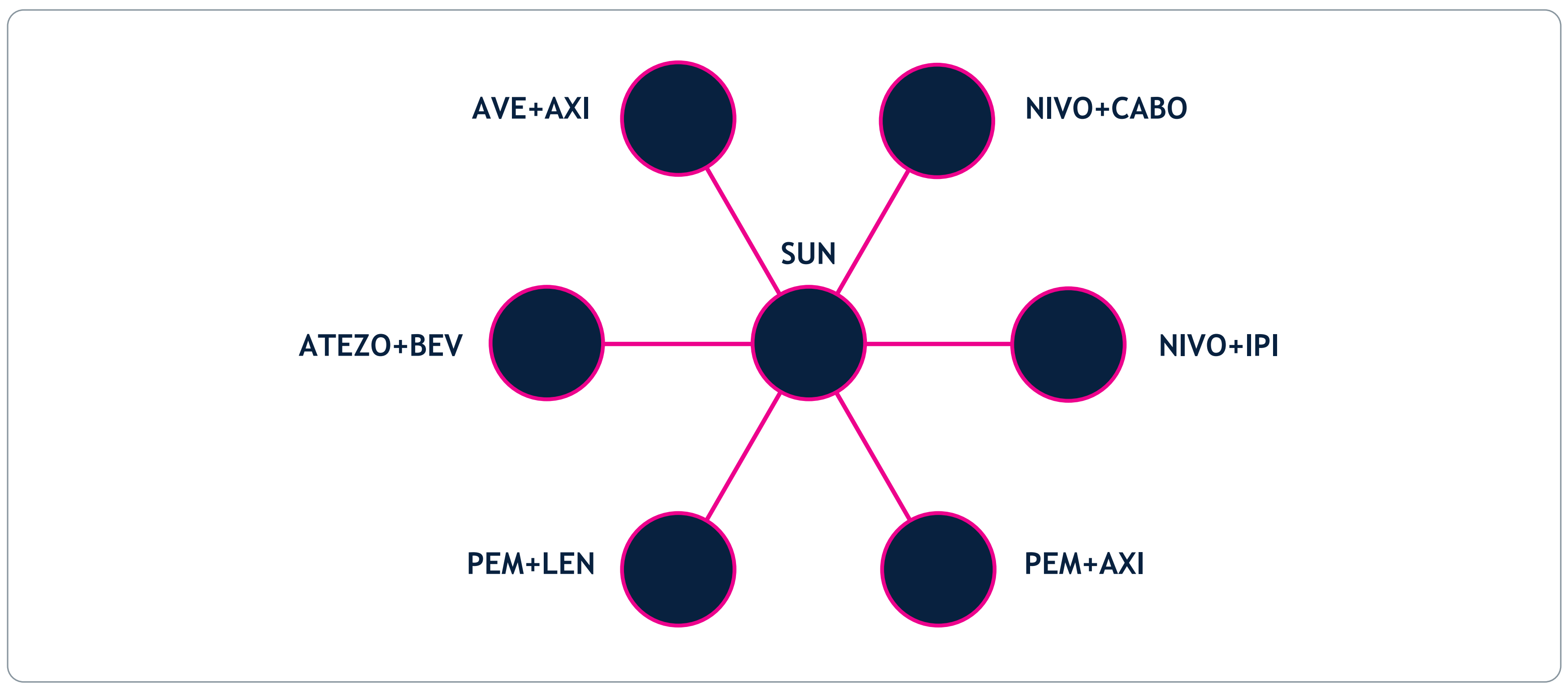
### Analyses

- An HR NMA was conducted in R using WinBUGS run via R2WinBUGS4 and Stan run via multinma[5] statistical software and shinystan[6]. Fixed effects models were fitted as only direct evidence was included in the network.
- WinBUGS generally performs MCMC by considering one parameter at a time and updating it according to the underlying joint distribution. This 'one-by-one' updating method thereby considers each dimension of the log density individually which has the advantage of only performing the computations necessary per iteration.[1]
- Stan's NUTS algorithm does not process one parameter at a time, but rather computes the entire space of all the parameters simultaneously.[2] This avoids some of the difficulties which may occur when sampling one dimension in an extremely high-dimensional problem but requires calculating the entire log density for each iteration.

### Outcomes

- The following performance indicators for the software packages were compared due to their relevance when comparing the two software and their respective sampling algorithms:
  - Median odds ratios (OR) and 95% credible intervals were assessed alongside treatment rankings, to ensure that the results are consistent between the software.
  - Running time, to determine the computation time required for the same number of iterations. Both NMAs were run on the same computer, using 10,000 iterations, 5,000 burin in iterations with a thinning rate of 1, and four chains.
  - Trace plots, to observe whether each model converged.
  - Qualitative outcomes regarding user experience, in particular error handling and console messages.

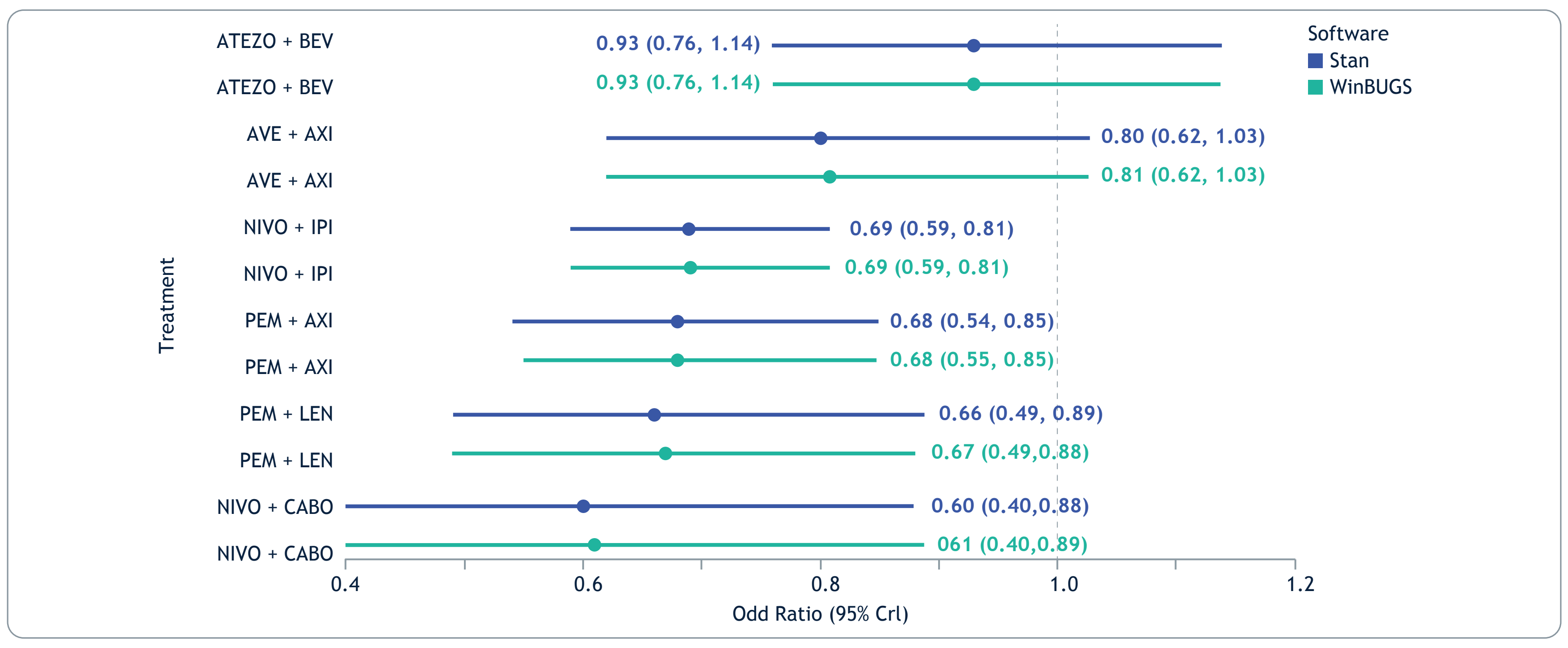
**Figure 1.** Network diagram of evidence for 1L RCC



Abbreviations: ATEZO, Atezolizumab; AVE, Avelumab; AXI, Axitinib; BEV, Bevacizumab; CABO, Cabozantinib; IPI, Ipilimumab; LEN, Lenvatinib; NIVO, Nivolumab; PEM, Pembrolizumab; SUN, Sunitinib.

## RESULTS

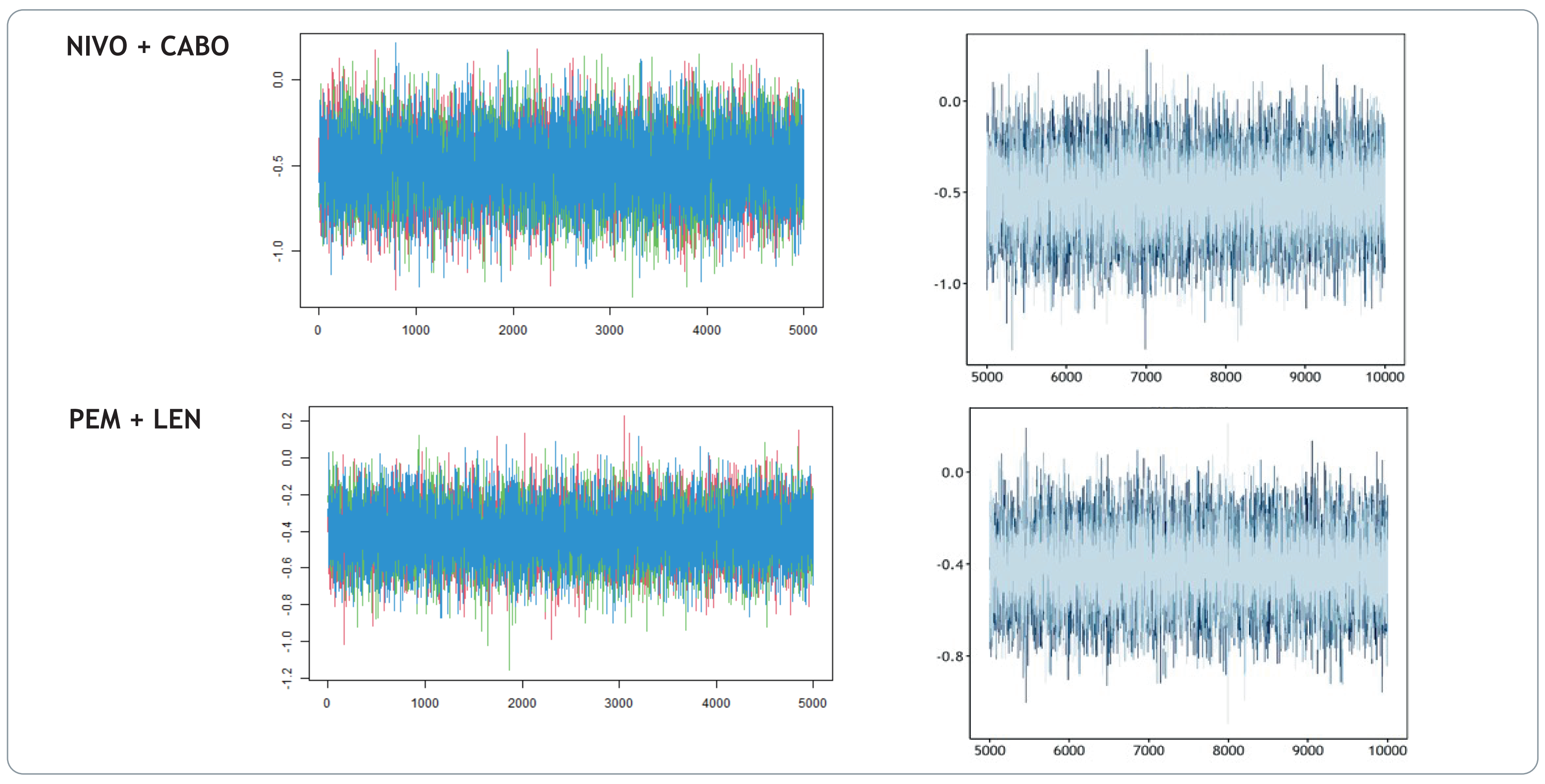
**Figure 2.** Forest plot showing the outcomes of SUN vs. comparators for analyses run in WinBUGS and Stan



### Quantitative outcomes

- Figure 2 shows the point estimates and CrIs of the two analyses. Differences in the median OR were minor for all comparisons (<0.002). The estimates for nivolumab + ipilimumab versus sunitinib were 0.690 using Stan (CrI 0.589-0.809) versus 0.690 using WinBUGS (CrI 0.588-0.809). Treatment rankings and the interpretation of the results also did not differ between the software.
- Running time was improved with use of Stan with a minor difference in the order of seconds.
- A selection of trace plots are shown in Figure 3, showing convergence was reached with both software.

**Figure 3.** Trace plots of WinBUGS (left) and Stan (right) for selected treatments



### Qualitative outcomes

- Users reported that Stan was perceived as easier to debug because of the additional information provided in error messages. Console warnings were seen as more actionable.

## DISCUSSION

- The deterministic outcomes of OR, CrI and treatment rankings were very similar, confirming that Stan and WinBUGS lead to similar results even though they used different algorithms.
- There were no notable differences in convergence and running time. For simple networks and fixed effects models, WinBUGS and Stan may perform equally well.
- An advantage in using Stan were the more user-friendly error messages, as Stan made diagnosing NMA models exceedingly user-friendly with the shinyStan functionality. This likely follows due to the age differences between the software. WinBUGS has not been updated since 2007 due to development shifting to OpenBUGS,[6] while Stan is still in active development.

## CONCLUSIONS

- This case study provides further insights about the differences between WinBUGS and Stan in a Bayesian HR-based NMA with a medium-sized star-shaped network.
- Findings suggest that comparable results and interpretations can be derived when conducting an HR based NMA with Stan or WinBUGS in a straightforward evidence network, despite differences in the algorithms used.
- Considering the similarity of the analysis results obtained from WinBUGS and Stan, and the advantages in running time and user experience regarding for Stan, Stan may be considered the more convenient software for conducting Bayesian HR-based NMA.

**REFERENCES**

1. Lunn, D.J., Thomas, A., Best, N. et al. WinBUGS - A Bayesian modelling framework: Concepts, structure, and extensibility. Statistics and Computing 10, 325–337 (2000).
2. Stan Modeling Language Users Guide and Reference Manual [program], 2022.
3. Bosma NA, Warkentin MT, Gan CL, et al. Efficacy and Safety of First-line Systemic Therapy for Metastatic Renal Cell Carcinoma: A Systematic Review and Network Meta-analysis. *European Urology Open Science* 2022;37:14-26.
4. Sturtz S, Ligges U, Gelman A. R2WinBUGS: A Package for Running WinBUGS from *R. Journal of Statistical Software* 2005;12(3):1-16. doi: 10.18637/jss.v012.i03.
5. Phillippo, D. M. (2022). multinma: Bayesian Network Meta-Analysis of Individual and Aggregate Data. https://doi.org/10.5281/zenodo.3904454 (last assessed October 2022).
6. shinystan Vignettes [program], 2022.
7. Lunn, D., Spiegelhalter, D., Thomas, A., & Best, N. (2009). The BUGS project: Evolution, critique and future directions. Statistics in medicine, 28(25), 3049-3067.

**ACKNOWLEDGEMENTS**

We want to thank Mark Lambton, Niels Dunnewind and Matt Knowles for their software support, Elisabeth Fenwick and Claire Simons for their review of the analysis, and Sonja Kroep for her contributions conceptualizing this project.

Presented at ISPOR Europe 2022