Statistical modelling of competing risks with incomplete data: with applications to allogeneic stem cell transplantation

Bonneville, E.F.

List of publications

- **Bonneville, E. F.**, Resche-Rigon, M., Schetelig, J., Putter, H., de Wreede, L. C. (2022) Multiple imputation for cause-specific Cox models: Assessing methods for estimation and prediction. *Statistical Methods in Medical Research*, 31, 1860–1880. DOI: 10.1177/09622802221102623.
- **Bonneville, E. F.**, Schetelig, J., Putter, H., de Wreede, L. C. (2023) Handling missing covariate data in clinical studies in haematology. *Best Practice & Research Clinical Haematology*, 36, 101477. DOI: 10.1016/j.beha.2023.101477.
- Koster, E. A. S.[†], **Bonneville, E. F.**[†], von dem Borne, P. A., van Balen, P., Marijt, E. W. A., Tjon, J. M. L., Snijders, T. J. F., van Lammeren, D., Veelken, H., Putter, H., Falkenburg, J. H. F., Halkes, C. J. M., de Wreede, L. C. (2023) Joint models quantify associations between immune cell kinetics and allo-immunological events after allogeneic stem cell transplantation and subsequent donor lymphocyte infusion. *Frontiers in Immunology*, 14. DOI: 10.3389/fimmu.2023.1208814.
- Polverelli, N.[†], **Bonneville, E. F.**[†], de Wreede, L. C., Koster, L., Kröger, N. M., Schroeder, T., Peffault de Latour, R., Passweg, J., Sockel, K., Broers, A. E. C., Clark, A., Dreger, P., Blaise, D., Yakoub-Agha, I., Petersen, S.L., Finke, J., Chevallier, P., Helbig, G., Rabitsch, W., Sammassimo, S., Arcaini, L., Russo, D., Drozd-Sokolowska, J., Raj, K., Robin, M., Battipaglia, G., Czerw, T., Hernández-Boluda, J. C., McLornan, D. P. (2024) Impact of comorbidities and body mass index on the outcomes of allogeneic hematopoietic cell transplantation in myelofibrosis: A study on behalf of the Chronic Malignancies Working Party of EBMT. *American Journal of Hematology*, 99, 993—996. DOI: https://doi.org/10.1002/ajh.27262
- **Bonneville, E. F.**, de Wreede, L. C. and Putter, H. (2024) Why you should avoid using multiple Fine–Gray models: Insights from (attempts at) simulating proportional subdistribution hazards data. *Journal of the Royal Statistical Society Series A: Statistics in Society*, qnae056. DOI: 10.1093/jrsssa/qnae056

[†]These authors contributed equally to this work and share first authorship.

- **Bonneville, E. F.**, Beyersmann, J., Keogh, R. H., Bartlett, J. W., Morris, T. P., Polverelli, N., de Wreede, L. C., Putter, H. (2024) Multiple imputation of missing covariates when using the Fine–Gray model. *arXiv preprint* arXiv:2405.16602. DOI: 10.48550/arXiv.2405.16602.

Acknowledgements

As part of the journey that was writing this thesis, I had the privilege of collaborating with and being surrounded by a number of fantastic people, to whom I would like to express my gratitude.

First and foremost, I would like to thank my supervisors Liesbeth de Wreede and Hein Putter, for their unwavering support and mentorship during the past years. Liesbeth, the consistently high scientific standards by which you work have been an absolute inspiration. Hein, I will always admire how you are able to communicate difficult statistical concepts with ease and enthusiasm. I have learned a lot from the both of you, and will always keep a fond memory of our weekly meetings, which rarely felt complete without culturo-linguistic insights and coffee. Hartelijk dank.

This thesis would likely not have been about missing data at all had it not been for Matthieu Resche-Rigon. Thank you for motivating me to build on your earlier ideas on multiple imputation in competing risks settings. Relatedly, I would like to thank Jacco Wallinga and Marta Fiocco, for encouraging me during my master's thesis to consider a PhD position in the first place.

Furthermore, I had the pleasure of working with several statistical and clinical collaborators during the PhD trajectory. Thank you Ruth Keogh and Jan Beyersmann for welcoming me in London and Ulm respectively, and for the enjoyable collaboration that ensued also together with Tim Morris and Jonathan Bartlett. On the clinical side, I would like to thank Johannes Schetelig and Nicola Polverelli for their contributions to several articles in this thesis, as well as Constantijn Halkes and Frederik Falkenburg for the stimulating collaboration throughout the joint models project. Special thanks to Eva Koster, for being a joy to work with, and for your patience in repeatedly explaining the basics of allogeneic stem cell transplantation (alloSCT) to me.

The present thesis relies on various datasets of patients who have undergone an alloSCT, registered with the European Society for Bone and Marrow Transplantation (EBMT). I thank both the EBMT for permitting the use of these data, and EBMT data managers Linda Koster and Laurien Baaij for their help in preparing the datasets. I would also like to express my gratitude to all patients and centres involved in the original studies. Moreover, I would like to acknowledge all the EBMT statisticians

that I have had the opportunity of learning from during the past years: Dirk-Jan, Luuk, Jarl, Paddy, Katerina, Simona, Gloria, Giulia, Katya, Richard, Jacques-Emmanuel, Christophe, Ariane, Maud, and Myriam.

Next, I would like to show my appreciation to my colleagues from Medical Statistics (MSTAT) at the Leiden University Medical Center (LUMC): Ningning, Mitra, Irene, Susanne, Alexandros, Xu, Mirko, Junfeng, Jelle, Saskia, Mar, Erik, Bart, Roula, Stefan, Szymon, Ramin, Ghislaine, Lu, Leonie, Lies, Ewout, and Nan. Ewout, thank you for getting me interested in prediction models, and for the many spontaneous conversations in your office. Nan, it was a pleasure to work together on the validation of competing risks prediction models project.

Needless to say, completing this thesis would have been a mere fraction of the fun without being surrounded by such a brilliant group of MSTAT PhD students: Chiara, Yongxi, Anna Kaal; Riccardo, Angela, Anna Vesely, Kaya, Andrea, Mari, Damjan (les visiteurs); Toby, Jasper, Georgy, Frank (the honourable gentlemen from office 42); Marije, Doranne, Lars, and Ilaria (the OGs). Thank you all for being absolute legends.

To my paranymphs Lars and Ila. Lars, beyond the many times you have made me laugh, I thank you in particular for inspiring me to continue to pursue interests and hobbies outside of work. Ila, aside from the consistent provision of Italian snacks, I thank you for being such a dependable friend.

Outside of MSTAT, several groups of PhD students positively impacted my time at the LUMC. Thank you to the 'Hacky Hour' folks for sparking my interest in Open Science early on in my PhD: Kim, Linda, Anna Lohmann, Daniela, Tariq, and Xante. Thank you also to the lovely people of the LUMC Association for PhD Candidates (LAP) for the welcome distraction that was organising events together: Daniele, Merian, Naomi, Alice, and Nienke.

To my parents and brother. Papá, Mamá, les agradezco de todo corazón su apoyo y todos los sacrificios que han hecho por mí. Je vous aime fort. Charlie, eres un crack—gracias por siempre hacerme reír, y por ser una parte tan positiva de mi vida.

Finally, my deepest thanks go to my other half Sarah (with the long name) von Grebmer zu Wolfsthurn. Not least for actually showing me that there was an end in sight to the PhD after you completed yours, but for your boundless patience, love, and support. Thank you being the best part of my life, and for always pushing me to become a better version of myself.

Curriculum vitae

Edouard Francis Bonneville was born on July 20th 1995 in Madrid (Spain). After completing his secondary education in Colomiers (France) in 2013, he pursued a BSc in Psychology at the University of Bristol (United Kingdom). In 2016, he moved to Leiden (the Netherlands) for the MSc Statistical Science for the Life and Behavioural Sciences, which is now known under the name 'Statistics and Data Science'. As part of his master thesis, he spent three months at the RIVM Dutch National Institute for Public Health and the Environment working on a Bayesian approach for forecasting infection disease epidemics.

In 2019, he started his PhD in Biostatistics at the Department of Biomedical Data Sciences at Leiden University Medical Center (LUMC, the Netherlands) under the supervision of Dr. Liesbeth de Wreede and Prof. Dr. Hein Putter. The results of his research, which focused on statistical methodology at the intersection of competing risks and missing data, are outlined in this thesis. During his PhD programme, he visited both the Department of Medical Statistics at the London School of Hygiene & Tropical Medicine (United Kingdom), and the Institute of Statistics at Ulm University (Germany). Alongside his PhD, he was part of the LUMC Association for PhD Candidates (LAP) board for two years, during which he was treasurer and co-organised several events.

Between 2022 and 2024, he also worked part-time as a study statistician for the European Society for Bone and Marrow Transplantation (EBMT). Finally, he has served as a reviewer for the journals *Statistics in Medicine*, *Statistical Methods in Medical Research*, *BMC Medical Research Methodology*, *Biometrical Journal* and *Journal of Computational and Graphical Statistics*.