Long-term excess mortality and net survival among elderly DLBCL patients after frontline R-CHOP treatment

Short title: Long-term excess mortality in elderly DLBCL patients

Aurélien Belot[1*], Vincent Camus[2,3], Cédric Portugues[1], Jérome Paget[1], Loic Chartier[1], Pascale Cony-Makhoul[1], Hervé Tilly[2,3], Clémentine Joubert[1]

1. Department of Biometry, LYSARC, Hopital Lyon-Sud, Pierre-Bénite, France
2. Department of Hematology, Centre Henri Becquerel, Rouen, France
3. INSERM U1245, Centre Henri Becquerel, University of Rouen, Rouen, France

* corresponding author : aurelien.belot@lysarc.org

ORCID : https://orcid.org/0000-0003-1410-5172

Abstract

Background:

In the era of immuno-chemotherapy, data on long-term prognosis of elderly patients diagnosed with a Diffuse Large B-Cell Lymphoma (DLBCL) are scarce. In this population and on the longer term, other cause mortality is an important competing risk that needs to be accounted for.

Methods

Using clinical trial data and relative survival approaches, we estimated 10-year net survival (NS) and we described the excess mortality hazard (EMH) due (directly or indirectly) to the DLBCL, over time and according to main prognosis factors using flexible regression modelling.

Results

The 10-year NS was 65% [59; 71]. From the flexible modelling, we showed that the EMH decreases steeply after diagnosis. The variables “performance status”, “number of extra-nodal sites” and the serum “Lactate Dehydrogenase” were strongly associated with the EMH, even after adjustment on other important variables.

Conclusion

EMH is very close to zero at 10 years for the whole population, so DLBCL patients do not experience an increased mortality compared to the general population in the long-term. The number of extra-nodal sites was an important prognostic factor shortly after diagnosis; suggesting that it is correlated with an important but unmeasured prognostic factor that would lead to this selection effect over time.

Keywords

Diffuse Large B-Cell Lymphoma; Elderly; Long term outcome ; Excess mortality hazard; Net survival; Flexible modelling

Introduction

Diffuse Large B-Cell Lymphoma (DLBCL) is the most common lymphoma as it represents 30-40% of all Non-Hodgkin lymphoma cases. Patients diagnosed with DLBCL may be cured in up to 50-60% of the patients who achieve a complete remission after frontline treatment. However elderly patients may display a worse prognosis(1), even in the era of chemo-immunotherapy(2,3), with elderly males having the worst overall survival (OS) of all patients' subgroups. Data on the prognosis of DLBCL patients on the longer term and especially amongst an elderly population after frontline R-CHOP (rituximab associated with doxorubicin, cyclophosphamide, vincristine and prednisone) are scarce, even if approximately half of patients with DLBCL are older than 60 years old. A prospective long-term follow-up program on data of the LNH03-6B trial(4), in which 60 to 80 years old patients were randomised, was implemented to collect vital status among patients alive at the end of the trial. In this recent report, the 10-year PFS and OS rates were 40.4% and 49.8%, respectively, slightly higher than those in the R-CHOP arm of the LNH-98.5 trial (10-year PFS and OS: 36.5% and 43.5%, respectively), which involved a similar population(5). Nevertheless, considering the treatment outcomes for this population of patients with a number of competing risks for death and only 3.7% of relapses beyond 5 years after frontline treatment, it is important to investigate long-term DLBCL-specific mortality hazard as compared with the expected mortality hazard in the general population. Indeed, when we are interested in estimating quantities specifically related to the disease under study, one important challenge is that the overall mortality hazard (and thus the overall survival) does not fully describe how DLBCL affects patients because of the role played by other cause mortality, especially on the longer term and for elderly patients(6–8), which could therefore bias the results. If we had accurate information on the cause of death, we could use competing risks methods, such as cause-specific hazard regression to mitigate such bias(9,10). However in the longer term and especially among the elderly, the cause of death might be not available, inaccurate or difficult to measure due to multimorbidities (11–13). In this context, the so-called "relative survival approaches" circumvent this later issue by using the expected mortality hazard from the general population as a background mortality, in order to estimate an excess mortality hazard(14–16,7). This excess mortality hazard could be seen as the cancer-specific mortality hazard, and we could describe long-term association between prognostic factors and the excess mortality hazard. Additionally, we could provide net survival estimates up to 10 years after diagnosis in DLBCL cancer patients. Net survival is an epidemiological indicator which could be interpreted as the cancer survival after accounting for the other-cause mortality(15,7,8).

The aims of this work are (i) to estimate net survival according to the main prognostic factors among elderly DLBCL patients up to 10-year after frontline treatment and (ii) to gain insights into their relationship with the excess mortality hazard, including potential non-linear or time-varying effects.

Material and Method

Data

The core of the data was obtained from the LNH03-6B trial(4) (NCT00144755), which compared two R-CHOP schemes for first line DLBCL 60-80 years-old patients. The patients and the treatments assessed in this trial were previously described(5). Briefly, this trial included patients diagnosed between 2003 and 2008, and followed up to the end of 2011. Long-term follow-up was set up for patients alive at the end of the trial and who were randomized in French centres. Data for the analysis performed in this paper were extracted in June 2020.

Expected mortality rates in France were derived from the observed mortality rates available by sex, annual age, calendar year (1975 to 2017), and Département of residence and provided by the Institut National de la Statistique et des Etudes Economiques (INSEE). For a given sex and a given Département, these observed mortality rates were smoothed for ages above 15 using a Poisson regression model that included a bidimensional smoothing spline of year and age. Mortality rates were projected for the years 2018 to 2021 using this same model. This work has been done by the biostatistical unit of the Hospices Civils de Lyon, using the `mgcv` package in R software.

Because long-term follow up was collected for patients randomised in French centers, we restricted our analysis on patients randomized in French centers. Our interest being on the first 10 years after randomisation, we limited the follow-up at 10 years so patients alive 10 years after treatment randomisation were censored at 10 years.

Statistical analysis

Relative survival approaches

The main idea of relative survival approaches is to compare the mortality hazard observed in our population of patients diagnosed with DLBCL to the expected mortality hazard in the general population with identical demographic characteristics, this latter being obtained from lifetables and detailed according to some demographic characteristics (usually at least sex and age). This comparison allows getting an estimate of excess mortality, which could be interpreted as mortality due directly or

indirectly to the disease under study, i.e. DLBCL in our case. For this interpretation to hold, two main conditions need to be met. Firstly, the mortality hazard of the studied disease should represent a negligible part of the expected mortality hazard in the general population. Secondly, the other-cause mortality hazard of the general population is equal to the other cause mortality hazard in the population analysed, within levels defined by the demographic characteristics available in the lifetable. More details on the conditions and applicability of such approaches can be found in dedicated papers (16,7,17,18).

Net survival

In relative survival approaches, one measure of interest is the net survival. Net survival is interpreted as the survival probability of cancer patients once the other causes of death have been removed. We estimated net survival using the non-parametric Pohar-Perme estimator (16,19), and according to prognostic factors. The prognostic factors considered are sex, age (<70 *vs.* ≥70), Ann-Arbor stage (I-II *vs.* III *vs.* IV), Performance status ECOG (0-1 *vs.* ≥2), number of extra nodal sites (0-1 *vs.* ≥2), serum Lactate Dehydrogenase LDH (values below *vs.* above the upper limit of the normal range ULN, i.e. LDH ≤ ULN *vs.* LDH > ULN)), and the International Prognostic Index IPI (0-2 *vs.* 3-5).

Multivariable regression model for the excess mortality hazard

We fitted a flexible regression model to estimate the excess mortality hazard (EMH) as a function of time and according to prognostic factors(20,21). We used cubic B-splines(22,23) with one knot located at the median event times for the baseline excess hazard and the time-dependent (TD) associations. For the non-linear functional form, we used quadratic splines with 1 knot located at 70 years for age and located at 2 for the number of extra nodal sites. Our model building strategy was based on 2 steps.

First, we investigated if non-linear functional forms for modeling the association between the continuous variables (age at diagnosis and number of extra nodal sites) and the EMH were required. We fitted 4 models with and without the non-linear functional form for each covariate and we retained the model with the lowest Akaike Information Criteria (AIC). The functional form used for modeling the number of extra nodal sites was a quadratic spline with one knot located at 2, while the one used for modeling age was a quadratic spline with one knot located at 70.

In a second step and from the model retained above, we used martingale residual-based tests to investigate time-dependent associations for each variable(24), and we retained time-dependent

associations at the level of 10%. For the multivariable regression modeling, we analysed the Ann-Arbor stage variable using the conventional binary categorisation (I-II *vs.* III-IV).

From the final multivariable model retained, we report the Excess Hazard Ratios (EHR) for variables with time-fixed (and linear) associations (if the variable is continuous). For time-dependent associations, we report graphically the changes of EHR according to time since diagnosis, and for non-linear functional form of a continuous variable, we depict the change of the EHR according to the variable values.

The R software was used for the analysis, with the package `relsurv`(19) for the non-parametric estimates of net survival and the package `mexhaz`(25) for the multivariable regression model.

Results

Description of the population

By focusing on the French patients, we analysed 507 patients, 56% male and a mean age of 70 years old (Table 1). Patients were mostly diagnosed with a DLBCL of stage III or IV (89%) and with a good performance status at diagnosis (77% presented an ECOG of 0 or 1). Around half of the patients had 2 or more extra nodal sites involved and 68% presented elevated LDH. We observed 230 deaths over the first 10 years of follow-up since randomisation. Using the reverse Kaplan-Meier approach(26), the median follow-up duration was 11 years; half of the patients would have had a follow-up greater than 11 years after randomisation (had they not died).

Non-parametric estimates of net survival

On the whole population, overall survival estimates at 10 years was 49.9% (95% CI: 45.4;54.9), while net survival estimates at 10 years was 64.5% (95% CI: 58.5;71.1) (Figure 1). Net survival estimates according to the main prognostic factors are displayed in Figure 2. As expected, we observed a strong association between age and net survival (older patients having a lower net survival). More surprisingly, Ann-Arbor stage did not show a strong association with net survival on our population, while ECOG and LDH had a strong impact on net survival. Finally the International Prognostic Index well discriminated patients according to net survival estimates.

Regression modelling for the Excess Mortality Hazard

Following our model building strategy, a non-linear functional form was not retained for age nor for the number of extra nodal sites (Table S1 in the appendix). Based on the martingale residual-based tests, only the number of extra nodal sites was retained for a TD effect (Figure S1 in the appendix). For the other prognostic factors, Excess hazard ratios are reported in Table 2. For one unit increase in age at diagnosis, the excess mortality hazard is increased by 6% (EHR=1.06, 95% CI 1.02;1.1), conditionally on the other covariates. There was no evidence of an associations between the EMH and either the Ann Arbor stage or sex. However, an ECOG equal or higher than 2 (compared to an ECOG of 0 or 1) was associated with a higher EMH (and so a lower net survival); the EHR was estimated to 1.91, (95% CI 1.3;2.81). It means that a patient with an ECOG equal or higher than 2 at diagnosis was exposed to an EMH 1.91 times higher compared to a patient with an ECOG value of 0 or 1 at diagnosis (conditional on all other covariates). A value of LDH above the upper limit of the normal range (compared to a value below) was also strongly associated with a higher EMH (EHR=2.73, 95% CI 1.57;4.75).

The EHR for the number of extra nodal sites varies with time since randomisation. Thus, being diagnosed with one additional extra nodal site increases the EMH on the first year after randomisation, while this effect vanishes later on (Figure 3). This can be also seen on Figure 4 where the dynamic over time of the excess mortality hazard for patients with 0, 1 or 2 extra nodal sites involved are displayed (other variables being set to specific values, as detailed in the figure title). We can see that the EMHs for patients with 2 extra nodal sites involved is much higher shortly after diagnosis and up to 2 years after randomisation. From 2 years onward, the EMH of patients with 2 extra nodal sites involved is identical to the EMH of patients with no extra nodal site involved. We can also observe on these graphs that the EMH is very close to null at 10 years (Figure 4).

Finally, to check the quality of the fit from the final model, we compared non-parametric net survival estimates with model-based net survival estimates derived on the whole population (appendix Figure S3) and for specific subgroups (appendix Figure S4). The model-based estimates of net survival nicely recover the pattern of the non-parametric net survival overall and within subgroups, thus indicating a good fit of the flexible model for the EMH.

Discussion

In this work we estimated net survival and the excess mortality hazard up to 10 years after treatment and according to known prognostic factors on a elderly population of DLBCL patients treated with frontline R-CHOP. It corresponds to important indicators for public health researchers(6) and in that regards, high quality data with important prognostic factors as obtained from randomised clinical trial are useful. We utilised a relative survival approach in order to provide "DLBCL-specific" quantities, that is the excess mortality hazard and the net survival, overall and according to prognostic factors. The excess mortality hazard, which can be interpreted as the DLBCL-specific mortality hazard, is continuously decreasing over time since randomisation, suggesting that DLBCL patients are not exposed to an increased mortality as compared to the general population in the longer term. In this population of patients aged between 60 and 80 at diagnosis, age was strongly associated with the EMH, but a non-linear functional form for age was not supported by the data. A simple linear association was retained, that is, being 1-year older when you compare patients aged 60 and 61 affects the excess mortality hazard in the same way than if you compare patients aged 78 and 79, conditionally on the other prognostic factors. As previously shown on Swedish data(27), we found that age remains a strong prognostic factor of the EMH, even after adjusting on clinical factors such as Ann-Arbor stage, ECOG or the LDH. Performance status of patients at diagnosis was also strongly associated with the EMH, while Ann-Arbor stage was not an important prognostic factor, after adjusting on the other variables such as performance status and the LDH. This surprising result for stage is explained by the the very low proportion of stage I-II in our population. We also showed that the "number of extra nodal site involved" plays a role on the excess mortality hazard shortly after diagnosis but this role disappears in the longer term. This time-varying association might be due to the fact that the variable "number of extra nodal site involved" is correlated with an unmeasured strong predictor, such as the total metabolic tumor volume (TMTV) for DLBCL patients prognosis(28). It leads to a selection effect of less frail patients, *i.e.* those with a low level of TMTV. Using TMTV as an additional predictor would be a very interesting work to conduct. Our model building strategy relied on the use of the residuals. To complement our model building strategy, we investigated if additional time-dependent effects could lead to a better prediction using the AIC. It shows that the model with one time-dependent effect for the "number of extra nodal site involved" has the lowest AIC (appendix Table S2), thus reinforcing our results. Time-dependent effects for the other covariates can also be displayed (appendix Figure S2) but such time-dependent effects are not supported by the data and were not kept in our final model.

Regarding the statistical analysis, we aimed to investigate longer-term DLBCL-specific mortality hazard instead of the overall mortality hazard, even though it has been shown to be challenging (29). We analysed the continuous variables in their original form, in order to avoid categorisation (30) and used

spline function if needed to avoid potential local biases obtained from a linear function (31). We also accounted for potential time-dependent effect for the prognostic factors, as it has been shown that mismodeling of covariates may induces residuals confounding (32). We limited our work to estimating the excess mortality over time since randomisation. Future work will investigate how intercurrent events, such as progression or the diagnosis of a second cancer affect the prognosis of DLBCL patients in the longer term using either multi-state modeling approaches(33,10) or conditional survival(34,35), while accounting for prognostic factors available.

One of the main limit of the study is that we analysed data of patients included in randomised clinical trial so these patients are probably not representative of the whole population. For example these patients may have no comorbidities even though this has been shown to be associated with the EMH in different geographical areas (36,37). Other characteristics, such as patients' socio economic position, marital status or the deprivation of the area where the patients was living were not available despite being potentially associated with prognosis for DLBCL patients in France(38). However, this lack of representativeness as compared to population-based data is counterbalanced by the availability of important prognostic factors, such as stage, performance status in a homogeneous population of patients treated all with R-CHOP.

To conclude, using high-quality data from a clinical trial coupled with a long-term follow-up program, we quantified the association between important prognostic factors and the EMH associated with DLBCL (directly or indirectly) and we showed that the EMH is almost null 5 years after the start of the treatment, even for those with a poor prognosis. Thus, DLBCL patients do not experience an increased mortality compared to the general population in the long-term

Table 1: Demographic and clinical characteristics of patients at baseline, and summary of follow-up information

	Level	Overall
Sample size		507
Sex (%)	Male	285 (56.2)
	Female	222 (43.8)
Age (mean (SD))		70.01 (5.15)
Age in category (%)	< 70	238 (46.9)
	≥ 70	269 (53.1)
Stage in category (%)	I-II	57 (11.2)
	III	84 (16.6)
	IV	366 (72.2)
Performance status in category (%)	0-1	391 (77.1)
	≥ 2	116 (22.9)
IPI in category (%)	0-2	122 (24.1)
	3-5	385 (75.9)
Number of extra nodal sites involved (mean (SD))		1.75 (1.37)
Number of extra nodal sites in category (%)	0-1	249 (49.1)
	≥ 2	258 (50.9)
LDH in category (%)	≤ ULN	161 (31.8)
	> ULN	346 (68.2)
Survival time in years (mean (SD))		5.72 (3.68)
Vital status at 10 years (%)	Alive	277 (54.6)
	Dead	230 (45.4)

SD: Standard Deviation; IPI: International Prognostic Index; LDH: lactate dehydrogenase

Table 2: Excess hazard ratios (EHR) with 95% confidence interval (95% CI) for prognostic factors with time-fixed regression coefficients, estimated from the final multivariable model

Prognostic factor	Level	Adjusted EHR (95% CI)
Age	For 1-unit increase	1.06 (1.02;1.10)
Stage	I-II	Ref
	III-IV	1.00 (0.53;1.91)
Perf. Status	0-1	Ref
	≥ 2	1.91 (1.30;2.81)
Sex	Male	Ref
	Female	0.88 (0.62;1.26)
LDH	≤ ULN	Ref
	> ULN	2.73 (1.57;4.75)

List of figures:

Figure 1: Non-parametric overall and net survival (with 95% confidence intervals) estimated on the whole cohort.

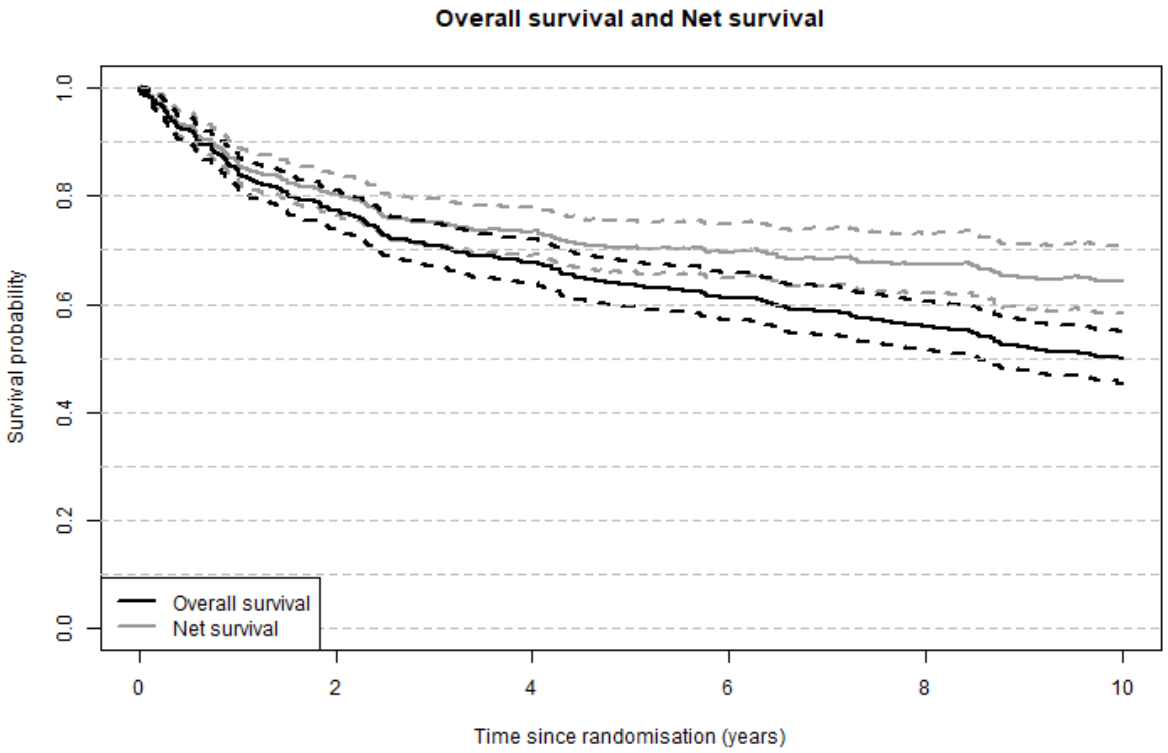


Figure 2: Non-parametric net survival estimates (solid lines) with 95% confidence intervals (dashed lines) according to each prognostic factor.

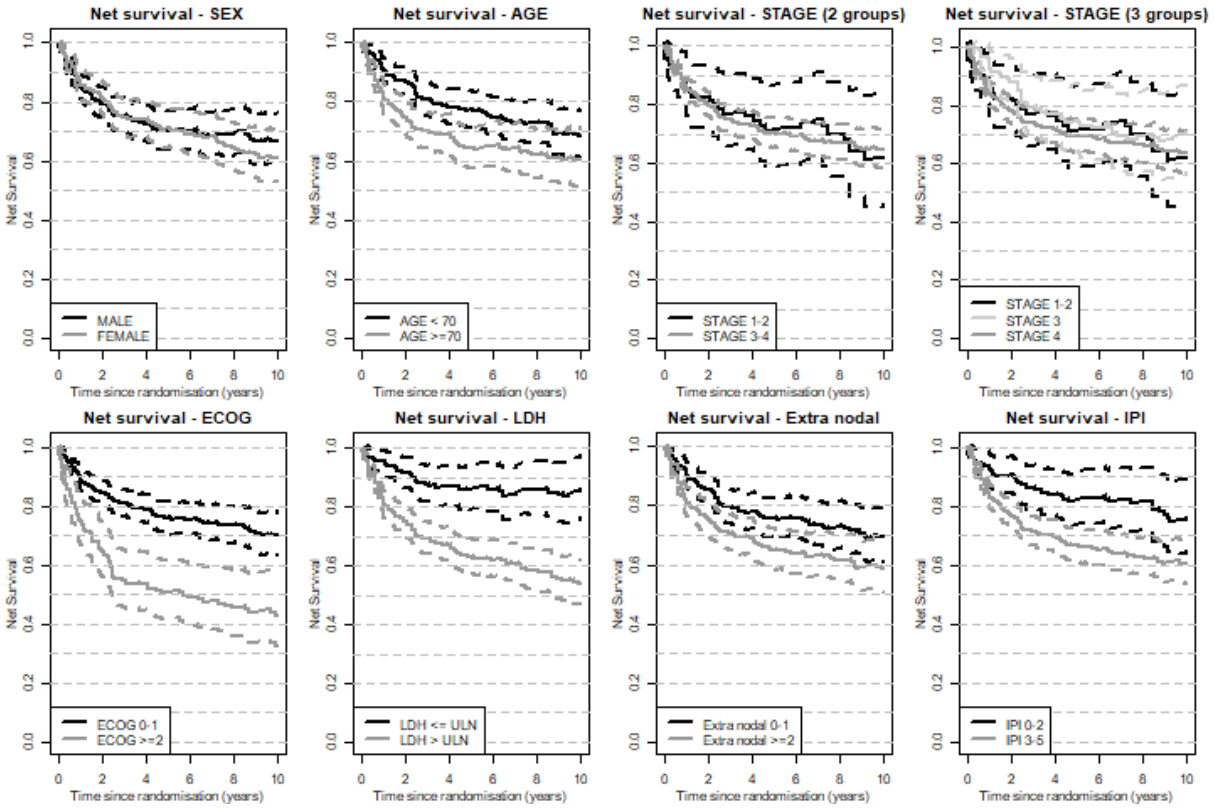


Figure 3: Excess hazard ratio for one unit increase of the number of extra nodal site involved, as estimated using a model with time-fixed effect for the number of extra-nodal sites involved (dashed line) or a time-dependent effect (solid line). The shaded area corresponds to the 95% confidence interval for the time-dependent effect.

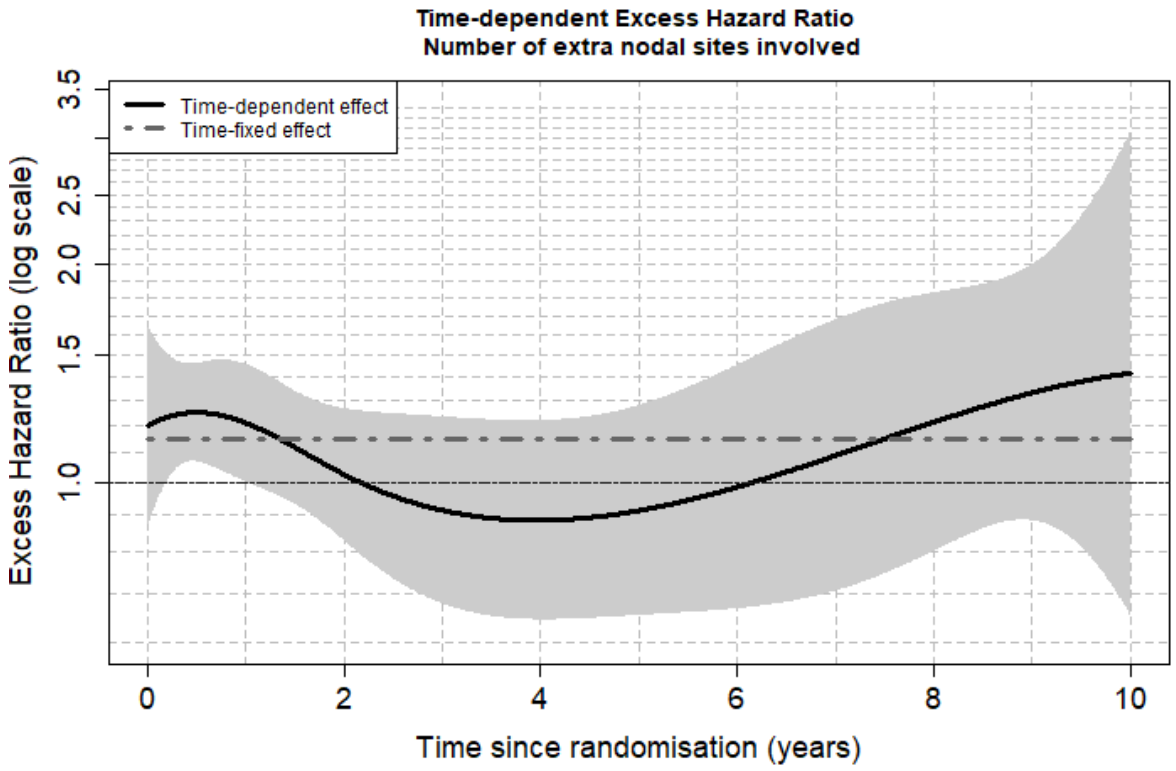
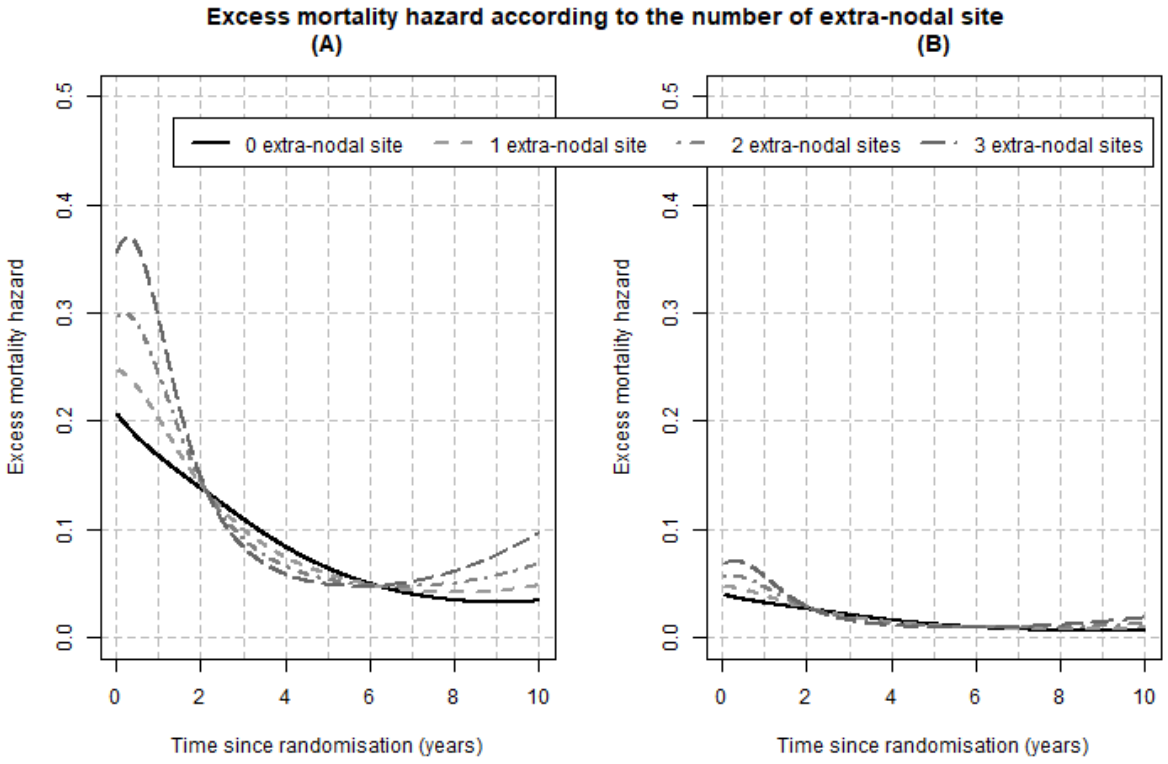


Figure 4: Excess mortality hazard according to time since diagnosis for a High risk group **(A),** male aged 70 years at diagnosis, stage III-IV, ECOG >=2, LDH > ULN and with 0, or 1, or 2, or 3 extra-nodal sites involved, and for a low risk group **(B),** male aged 70 years at diagnosis, stage I-II, ECOG 0-1, LDH ≤ ULN and with 0, or 1, or 2, or 3 extra-nodal sites involved.



References

1. Hounsome L, Eyre TA, Ireland R, Hodson A, Walewska R, Ardeshna K, et al. Diffuse large B cell lymphoma (DLBCL) in patients older than 65 years: analysis of 3 year Real World data of practice patterns and outcomes in England. Br J Cancer. 1 janv 2022;126(1):134-43.

2. Coiffier B, Thieblemont C, Van Den Neste E, Lepeu G, Plantier I, Castaigne S, et al. Long-term outcome of patients in the LNH-98.5 trial, the first randomized study comparing rituximab-CHOP to standard CHOP chemotherapy in DLBCL patients: a study by the Groupe d'Etudes des Lymphomes de l'Adulte. Blood. 23 sept 2010;116(12):2040-5.

3. Peyrade F, Jardin F, Thieblemont C, Thyss A, Emile JF, Castaigne S, et al. Attenuated immunochemotherapy regimen (R-miniCHOP) in elderly patients older than 80 years with diffuse large B-cell lymphoma: a multicentre, single-arm, phase 2 trial. Lancet Oncol. mai 2011;12(5):460-8.

4. Delarue R, Tilly H, Mounier N, Petrella T, Salles G, Thieblemont C, et al. Dose-dense rituximab-CHOP compared with standard rituximab-CHOP in elderly patients with diffuse large B-cell lymphoma (the LNH03-6B study): a randomised phase 3 trial. Lancet Oncol. mai 2013;14(6):525-33.

5. Camus V, Belot A, Obéric L, Sibon D, Ghesquieres H, Thieblemont C, et al. Outcomes of elderly diffuse large B-cell lymphoma patients treated with R-CHOP: 10-year follow-up of the LNH03-6B trial. Blood Adv. 23 juin 2022;bloodadvances.2022007609.

6. Mariotto AB, Noone AM, Howlader N, Cho H, Keel GE, Garshell J, et al. Cancer survival: an overview of measures, uses, and interpretation. J Natl Cancer Inst Monogr. nov 2014;2014(49):145-86.

7. Belot A, Ndiaye A, Luque-Fernandez MA, Kipourou DK, Maringe C, Rubio FJ, et al. Summarizing and communicating on survival data according to the audience: a tutorial on different measures illustrated with population-based cancer registry data [Internet]. Clinical Epidemiology. 2019 [cité 4 févr 2020]. Disponible sur: https://www.dovepress.com/summarising-and-communicating-on-survival-data-according-to-the-audien-peer-reviewed-fulltext-article-CLEP

8. Eloranta S, Smedby KE, Dickman PW, Andersson TM. Cancer survival statistics for patients and healthcare professionals – a tutorial of real-world data analysis. J Intern Med. 2021;289(1):12-28.

9. Belot A, Abrahamowicz M, Remontet L, Giorgi R. Flexible modeling of competing risks in survival analysis. Stat Med. 2010;29(23):2453-68.

10. Geskus RB. Data analysis with competing risks and intermediate states. Boca Raton: Taylor & Francis; 2016. 247 p. (Chapman & Hall/CRC biostatistics series).

11. Percy C, Stanek E, Gloeckler L. Accuracy of cancer death certificates and its effect on cancer mortality statistics. Am J Public Health. mars 1981;71(3):242-50.

12. Fall K, Strömberg F, Rosell J, Andrèn O, Varenhorst E, South-East Region Prostate Cancer Group. Reliability of death certificates in prostate cancer patients. Scand J Urol Nephrol. 2008;42(4):352-7.

13. Johnson CJ, Hahn CG, Fink AK, German RR. Variability in cancer death certificate accuracy by characteristics of death certifiers. Am J Forensic Med Pathol. juin 2012;33(2):137-42.

14. Estève J, Benhamou E, Croasdale M, Raymond L. Relative survival and the estimation of net survival: Elements for further discussion. Stat Med. mai 1990;9(5):529-38.

15. Perme MP, Stare J, Estève J. On Estimation in Relative Survival. Biometrics. 2012;68(1):113-20.

16. Pohar Perme M, Estève J, Rachet B. Analysing population-based cancer survival - settling the controversies. BMC Cancer. 03 2016;16(1):933.

17. Kipourou DK, Perme MP, Rachet B, Belot A. Direct modeling of the crude probability of cancer death and the number of life years lost due to cancer without the need of cause of death: a pseudo-observation approach in the relative survival setting. Biostatistics. 6 mai 2020;kxaa017.

18. Belot A, Pohar-Perme M. Social Disparities in Cancer Survival: Methodological Considerations. In: Launoy G, Zadnik V, Coleman MP, éditeurs. Social Environment and Cancer in Europe [Internet]. Cham: Springer International Publishing; 2021 [cité 8 déc 2021]. p. 39-54. Disponible sur: https://link.springer.com/10.1007/978-3-030-69329-9_5

19. Perme MP, Pavlic K. Nonparametric Relative Survival Analysis with the R Package relsurv. J Stat Softw. 30 nov 2018;87(1):1-27.

20. Remontet L, Bossard N, Belot A, Estève J. An overall strategy based on regression models to estimate relative survival and model the effects of prognostic factors in cancer survival studies. Stat Med. 2007;26(10):2214-28.

21. Charvat H, Remontet L, Bossard N, Roche L, Dejardin O, Rachet B, et al. A multilevel excess hazard model to estimate net survival on hierarchical data allowing for non-linear and non-proportional effects of covariates. Stat Med. 2016;35(18):3066-84.

22. Gauthier J, Wu QV, Gooley TA. Cubic splines to model relationships between continuous variables and outcomes: a guide for clinicians. Bone Marrow Transplant. 1 oct 2019;

23. Perperoglou A, Sauerbrei W, Abrahamowicz M, Schmid M. A review of spline function procedures in R. BMC Med Res Methodol [Internet]. 6 mars 2019 [cité 10 mars 2020];19. Disponible sur: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6402144/

24. Danieli C, Bossard N, Roche L, Belot A, Uhry Z, Charvat H, et al. Performance of two formal tests based on martingales residuals to check the proportional hazard assumption and the functional form of the prognostic factors in flexible parametric excess hazard models. Biostatistics. 1 juill 2017;18(3):505-20.

25. Charvat H, Belot A. mexhaz: An R Package for Fitting Flexible Hazard-Based Regression Models for Overall and Excess Mortality with a Random Effect. J Stat Softw. 12 juill 2021;98:1-36.

26. Schemper M, Smith TL. A note on quantifying follow-up in studies of failure time. Control Clin Trials. août 1996;17(4):343-6.

27. Hedström G, Hagberg O, Jerkeman M, Enblad G, on behalf of the Swedish Lymphoma Study Group. The impact of age on survival of diffuse large B-cell lymphoma – a population-based study. Acta Oncol. 3 juill 2015;54(6):916-23.

28. Vercellino L, Cottereau AS, Casasnovas O, Tilly H, Feugier P, Chartier L, et al. High total metabolic tumor volume at baseline predicts survival independent of response to therapy. Blood. 16 avr 2020;135(16):1396-405.

29. Schaffar R, Belot A, Rachet B, Woods L. On the use of flexible excess hazard regression models for describing long-term breast cancer survival: a case-study using population-based cancer registry data. BMC Cancer. 28 janv 2019;19(1):107.

30. Royston P, Altman DG, Sauerbrei W. Dichotomizing continuous predictors in multiple regression: a bad idea. Stat Med. 15 janv 2006;25(1):127-41.

31. Abrahamowicz M, Berger R du, Graver SA. Flexible Modeling of the Effects of Serum Cholesterol on Coronary Heart Disease Mortality. Am J Epidemiol. 15 avr 1997;145(8):714-29.

32. Brenner H, Blettner M. Controlling for continuous confounders in epidemiologic research. Epidemiol Camb Mass. juill 1997;8(4):429-34.

33. Putter H, Fiocco M, Geskus RB. Tutorial in biostatistics: competing risks and multi-state models. Stat Med. 2007;26(11):2389-430.

34. Bouvier AM, Remontet L, Hédelin G, Launoy G, Jooste V, Grosclaude P, et al. Conditional relative survival of cancer patients and conditional probability of death: A French National Database analysis. Cancer. 1 oct 2009;115(19):4616-24.

35. Hieke S, Kleber M, König C, Engelhardt M, Schumacher M. Conditional Survival: A Useful Concept to Provide Information on How Prognosis Evolves over Time. Clin Cancer Res. 1 avr 2015;21(7):1530-6.

36. Smith MJ, Belot A, Quartagno M, Luque Fernandez MA, Bonaventure A, Gachau S, et al. Excess Mortality by Multimorbidity, Socioeconomic, and Healthcare Factors, amongst Patients Diagnosed with Diffuse Large B-Cell or Follicular Lymphoma in England. Cancers. 19 nov 2021;13(22):5805.

37. Lee SF, Evens AM, Ng AK, Luque-Fernandez MA. Socioeconomic inequalities in treatment and relative survival among patients with diffuse large B-cell lymphoma: a Hong Kong population-based study. Sci Rep. déc 2021;11(1):17950.

38. Le Guyader-Peyrou S, Orazio S, Dejardin O, Maynadié M, Troussard X, Monnereau A. Factors related to the relative survival of patients with diffuse large B-cell lymphoma in a population-based study in France: does socio-economic status have a role? Haematologica. mars 2017;102(3):584-92.

Acknowledgements

The authors thank the patients and their families, the LYSARC, and all of the investigators in the LYSA centers. We thank in particular Richard Delarue who was the Principal Investigator of the original LNH03-6B trial. We also thank Mr Fabrice Inglès, clinical project coordinator at LYSARC, Lyon, France for the coordination of the long-term follow-up program, and Dr Adeline Bernier for some advice on a draft of the paper.

Authors' contributions:

ABE, CJO, VCA and HTI designed the research. ABE analyzed the data and wrote the first draft of the paper. All authors contributed to the writing of the manuscript, interpreted the data, and commented on. All authors approved the final version of the manuscript.

Ethics approval and consent to participate

For the original data collection of the LNH03-6B clinical trial (number NCT00144755), local or national ethics committees approved the study protocol according to the laws of each country. The study was performed in accordance with the Declaration of Helsinki. Patients provided written informed consent before inclusion. For the prospective long-term follow-up (LTFU) program, ethics approval were provided by the CCTIRS on 10/03/2011 and CNIL on 22/07/2011.

Consent for publication

Not applicable.

Data availability

De-identified participant data can be provided by (and after approval from) the LYSARC. Requests should be sent to the corresponding author.

Competing interests

The authors declare no conflicts of interest relevant to this study.

Relevant financial activities outside the submitted work:

VC: Honouraria: Octapharma, Abbvie, Roche, Incyte, Gilead-Kite, Bristol-Myers Squibb, Sanofi, Novartis; Travel Grants: Pfizer, Roche, Gilead-Kite, Novartis; Research Grants: Iqone Health care.

HT: Honouraria: Celgene, Roche, Karyopharm, Astra-Zeneca, Bristol-Myers Squibb, Grants: Celgene

Funding information

The author(s) received no specific funding for this work.