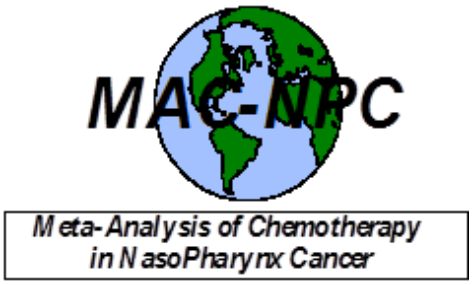

Development of restricted mean survival time difference in network meta-analysis based on data from MACNPC update.

Claire Petit, Pierre Blanchard, Jean-Pierre Pignon, Béranger Lueza

June 2017

CONTENTS

Abbreviation

AC= Adjuvant Chemotherapy

CRT= Concomitant chemoradiotherapy

CT= Chemotherapy

IC= Induction Chemotherapy

MACNPC= Meta-Analysis of Chemotherapy in NasoPharynx Cancer

MTC= Mixed Treatment Comparison

NMA= Network Meta-Analysis

rmst = Restricted Mean Survival Time

rmstD = Difference in Restricted Mean Survival Time

RT= Radiation Therapy

CONTEXT

Network meta-analysis (NMA), also known as mixed treatment comparisons (MTC), is a statistical method that deals with conditions where multiple treatments have been investigated that have not been compared altogether [1]. NMA permits evaluation of all possible pair-wise comparisons based on direct and indirect evidence, and allows ranking the different treatments according to their relative efficacies. A first NMA in nasopharynx cancer has been performed using the hazard ratio as the measure outcome of the treatment effect in each trial [2]. In this project, we will use an absolute outcome, the restricted mean survival time (RMST), that will be estimated in both arms of the trials included in MAC-NPC network meta-analysis. This project aims at repeating the NMA using the between -arms difference in RMST (rmstD) as an alternative outcome measure for the hazard ratio. The rmstD has previously been adapted to individual patient-data meta-analysis [3]–[5].

OBJECTIVE

Apply the difference in restricted mean survival time (rmstD) in a NMA and compare the results with those obtained in a NMA with hazard ratio.

ENDPOINT

The primary endpoint that will be evaluated in this NMA is the primary endpoint determined in the standard meta-analysis (MA): overall survival. The overall survival (OS) is defined as the time from randomization until death from any cause. Living patients without event will be censored at their date of last follow-up.

For this work, we choose as secondary endpoints, for exploratory analyses, progression-free survival (PFS) and locoregional control (LRC) since PFS has earlier events than OS and LRC has the fewest events. PFS is defined as the time from randomization to first progression (locoregional or distant) or death from any cause. Patients with a distant failure as a first event are censored for locoregional failure. Analyzes will be done without taking into account competing risks.

STATISTICAL ANALYSIS PLAN

Estimation of the difference in restricted mean survival time

For each trial and each outcome, we will use the absolute outcome estimated by the difference in restricted mean survival time (rmstD) and its variance, which will be restricted at a time t*. The rmstD(t*) can be defined as follows:

$$rmstD(t^*) = \int_0^{t^*} S_{Exp}(t)dt - \int_0^{t^*} S_{Cont}(t)dt = RMST_{Exp} - RMST_{Cont}$$

$$\widehat{Var}(rmstD(t^*)) = \widehat{Var}\left(RMST_{Exp}(t^*)\right) + \widehat{Var}(RMST_{Cont}(t^*))$$

This absolute outcome can be expressed as the number of life-years gained (or lost) associated with the treatment studied.

Description of the model

All models of the NMA are based on 3 hypotheses:

- the homogeneity hypothesis, as in a standard meta-analysis.
- the transitivity hypothesis, which assumes that the indirect comparison is valid through a common comparator: B ℜ A et A ℜ C ➔ B ℜ C.
- the consistency hypothesis which assumes that there is no discrepancy between direct and indirect estimates into a closed loop.

Statistical method

For the estimation of the rmstD(t*), we selected t*= 10 years for the primary analysis and t*= 5 years for a sensitivity analysis, as these were the two time points of clinical interest in the publications of MAC-NPC2. At t*=10 years, the majority of the trials included in the NMA has a follow-up long enough. We will used the pooled Kaplan-Meier method with DerSimonian-Laird random effect, because a previous study comparing the different methods to estimate the overall rmstD(t*) from IPD meta-analysis showed that this method is the best compromise in term of bias and variance [3]. If needed, an extrapolation until t* will be done with the method proposed by Brown et al [6]. A two-step method will be used, the first step is the computation of rmstD(t*) based on the individual patient data gathered by the MACNPC collaborative group. The second step is the actual network meta-analysis, using as input data for each trial the two treatments compared, the rmstD(t*) and its standard error.

Moreover, we will test the proportional hazards assumption at the trial level using the Grambsch-Therneau test [7] and at the pairwise meta-analysis level using the test proposed by Wei et al. [5]

All analyses will be performed under a frequentist approach with the R package netmeta. This package is based on graph theory methodology to model the relative treatment effects of multiple treatments under a frequentist framework. The R netmeta package will be used to provide the estimation of Q test for inconsistency [8], [9].

As previously mentioned, a DerSimonian-Laird random effects model will be used first in the NMA , even without unexplained heterogeneity, as a previous work showed that fixed effect model underestimates the variance of the overall rmstD(t*) [3].

The reporting of the results will include a description of the networks (primary and for each sensitivity analysis), effect sizes from direct evidence, indirect evidence, and the network meta-analysis (at least for the primary analysis, through a comparative rmstD(t*) plot for each comparison of interest), a ranking of the treatment that includes the uncertainty of the ranking estimates. Ranking will be performed by P-score.

The forced consistency in multi-arms trials will be handled specifically. Heterogeneity will be quantified using the I^2, which represents the proportion of total variation in study estimates that is due to heterogeneity [10]. Inconsistency in the network can be assessed globally or for each closed loops. In order to limit the number of tests for both heterogeneity and inconsistency, Rücker et al have proposed a global test, called Q test [11]. This test is a generalization of Cochran's test that is used to assess heterogeneity in conventional meta-analysis. The Q statistic is the sum of a Q statistic for heterogeneity (within designs) and a Q statistic for inconsistency (between designs).

Sensitivity analyses will be used in case of significant network inconsistency ($p<0.1$). The *netmeta* package allows identifying in which closed loop the inconsistency is located [9]. The trials responsible for inconsistency will be determined by comparing direct and indirect estimates and trial forest plots within the inconsistent closed loop. The effect of trial removal on the network overall consistency and estimation will be investigated.

Rücker and Schwarzer have proposed a frequentist analogue to SUCRA (surface under the cumulative ranking curve, method used to rank treatment within the Bayesian framework) called P-Score that works without resampling, and measures the mean extent of certainty that a treatment is better than the competing treatments. P-Score would be 1 when a treatment is certain to be the best and 0 when a treatment is certain to be the worst [12].

This work will be performed in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) Extension Statement for Reporting of Systematic Reviews Incorporating Network Meta-analyses of Health Care Interventions [13]. P-values <0.05 will be considered significant. All analyses will be performed using the R software version 3.2.1.

Sensitivity analyses

Clinical sensitivity analyses will be performed in coherence with the standard meta-analysis, either by excluding a certain category of patients or a certain category of trials.

Network meta-analysis will be repeated with t* = 5 years for the estimation of rmstD(t*) and compared to HR censored at 5 years. Indeed, at t*=5 years, all trials have a follow-up long enough and do not necessitate survival extrapolation.

Another sensitivity analysis will be done after the exclusion of trials with non-proportional hazards.

REFERENCES

[1] P. Blanchard *et al.*, "Mixed treatment comparison meta-analysis of altered fractionated radiotherapy and chemotherapy in head and neck cancer," *J. Clin. Epidemiol.*, vol. 64, no. 9, pp. 985–992, Sep. 2011.

[2] L. Ribassin-Majed *et al.*, "What Is the Best Treatment of Locally Advanced Nasopharyngeal Carcinoma? An Individual Patient Data Network Meta-Analysis," *J. Clin. Oncol.*, vol. 35, no. 5, pp. 498–505, Feb. 2017.

[3] B. Lueza, F. Rotolo, J. Bonastre, J.-P. Pignon, and S. Michiels, "Bias and precision of methods for estimating the difference in restricted mean survival time from an individual patient data meta-analysis," *BMC Med. Res. Methodol.*, vol. 16, no. 1, Dec. 2016.

[4] B. Lueza, A. Mauguen, J.-P. Pignon, O. Rivero-Arias, J. Bonastre, and MAR-LC Collaborative Group, "Difference in Restricted Mean Survival Time for Cost-Effectiveness Analysis Using Individual Patient Data Meta-Analysis: Evidence from a Case Study," *PLOS ONE*, vol. 11, no. 3, p. e0150032, Mar. 2016.

[5] Y. Wei, P. Royston, J. F. Tierney, and M. K. B. Parmar, "Meta-analysis of time-to-event outcomes from randomized trials using restricted mean survival time: application to individual participant data," *Stat. Med.*, vol. 34, no. 21, pp. 2881–2898, Sep. 2015.

[6] J. Brown Byron W., M. Hollander, and R. M. Korwar, "Nonparametric Tests of Independence for Censored Data with Application to Heart Transplant Studies," *Reliab. Biometry Stat. Anal. Lifelength*, pp. 327–54, 1974.

[7] P. M. Grambsch and T. M. Therneau, "Proportional hazards tests and diagnostics based on weighted residuals," *Biometrika*, vol. 81, no. 3, pp. 515–526, Aug. 1994.

[8] U. Krahn, H. Binder, and J. König, "A graphical tool for locating inconsistency in network meta-analyses," *BMC Med. Res. Methodol.*, vol. 13, no. 1, p. 35, Mar. 2013.

[9] G. Rücker, G. Schwarzer, U. Krahn, and J. König, *netmeta: Network meta-analysis with R*. 2014.

[10] J. P. T. Higgins and S. G. Thompson, "Quantifying heterogeneity in a meta-analysis," *Stat. Med.*, vol. 21, no. 11, pp. 1539–1558, Jun. 2002.

[11] G. Rücker, "Network meta-analysis, electrical networks and graph theory," *Res. Synth. Methods*, vol. 3, no. 4, pp. 312–324, Dec. 2012.

[12] G. Rücker and G. Schwarzer, "Ranking treatments in frequentist network meta-analysis works without resampling methods," *BMC Med. Res. Methodol.*, vol. 15, no. 1, p. 58, Dec. 2015.

[13] B. Hutton *et al.*, "The PRISMA Extension Statement for Reporting of Systematic Reviews Incorporating Network Meta-analyses of Health Care Interventions: Checklist and Explanations," *Ann. Intern. Med.*, vol. 162, no. 11, p. 777, Jun. 2015.