Individual patient data Network-Meta-analysis of chemotherapy and radiotherapy in locally advanced head and neck cancer.

Authors
Claire Petit, Benjamin Lacas, Jean-Pierre Pignon, Quynh Thu Le, Prof Vincent Grégoire, Prof Cai Grau, Prof Allan Hackshaw, Björn Zackrisson, Prof Mahesh KB Parmar, Ju-Whei Lee, Maria Grazia Ghi, Giuseppe Sanguineti, Stéphane Temam, Maurice Cheugoua-Zanetsie, Prof Brian O’Sullivan, Marshall R. Posner, Everett E Vokes, Juan J. Cruz Hernandez, Zbigniew Szutkowski, Eric Lartigau, Prof Volker Budach, Prof Rafal Suwiński, Prof Michael Poulsen, Shaleen Kumar, Prof Sarbani Ghosh Laskar, Prof Jean-Jacques Mazeron, Prof Branislav Jeremic, Prof John Simes, Prof Lai-Ping Zhong, Prof Jens Overgaard, Catherine Fortpied, Pedro Torres-Saavedra, Prof Jean Bourhis, Anne Aupérin, Pierre Blanchard, on behalf of the MACH-NC and MARCH Collaborative Groups*

*Members of the collaborative group are listed in Web-Appendix 1 (p 36)

Affiliations
Service de Biostatistique et d’Épidémiologie, Gustave Roussy, Oncosstat U1018, Ligue Contre le Cancer, INSERM, Université Paris-Saclay, Villejuif, France (C. Petit MD, B Lacas PhD, JP Pignon MD, M Cheugoua-Zanetsie MSc, A Aupérin MD, P Blanchard MD)
Department of Radiation Therapy, Gustave Roussy Cancer Campus, Université Paris-Sud, Université Paris-Saclay, F-94805, Villejuif, France (C. Petit MD, P Blanchard MD)
Groupe d’Oncologie Radiothérapie Tête Et Cou, Tours, France (C. Petit MD, B Lacas PhD, JP Pignon MD, Prof J Bourhis MD, A Aupérin MD, P Blanchard MD)
Department of Radiation Oncology, Stanford University School of Medicine, Stanford, CA, USA (Q T Le MD)
Radiation Oncology Department, Centre Léon Bérard, Lyon, France (Prof V Grégoire MD)
Department of Oncology, Aarhus University Hospital, Aarhus, Denmark (Prof C Grau MD, Prof J Overgaard MD)
Cancer Research UK & University College London Cancer Trials Centre, Cancer Institute, University College London Hospital, London, UK (Prof A Hackshaw MSc)
Department of Radiation Sciences, Oncology, Umeå University, Umeå, Sweden (B Zackrisson MD)
Medical Research Council Clinical Trials Unit, University College London, London, UK (Prof M KB Parmar DPhil)
Dana Farber Cancer Institute - ECOG-ACRIN Biostatistics Center, Boston, MA, USA (JW Lee PhD)
Oncology Unit 2, Veneto Institute of Oncology IOV – IRCCS, Padua, Italy (M G Ghi MD)
Department of Radiation Oncology, IRCCS Regina Elena National Cancer Institute, Rome, Italy (G Sanguineti MD)
Service de cancérologie cervico-faciale, Gustave Roussy, Université Paris-Saclay, F-94805, Villejuif, France (S Temam MD)
Department of Radiation Oncology, Princess Margaret Cancer Centre, University of Toronto, Toronto, ON, Canada (Prof B O’Sullivan MD)
Department of Medicine, Icahn School of Medicine at Mount Sinai, New York, NY, USA (M R Posner MD)
Section of Hematology-Oncology, The University of Chicago Medical Center, Chicago, Illinois, USA (E E Vokes MD)
Medical Oncology Department, University of Salamanca, Salamanca, Spain (J J Cruz Hernandez MD)
Department of Radiotherapy, Cancer Center, Marie Curie-Sklodowska Memorial Institute, Warsaw, Poland (Z Szutkowski MD)
Department of Radiotherapy, Centre Oscar Lambret, France (E Lartigau MD)
Department of Radiation Oncology, Charité Universitätsmedizin Berlin, corporate member of Freie Universität Berlin, Humboldt-Universität zu Berlin, and Berlin Institute of Health, Berlin, Germany (Prof V Budach MD)
Radiotherapy and Chemotherapy Clinic and Teaching Hospital, Maria Skłodowska-Curie National Research Institute of Oncology, Gliwice Branch, Poland (Prof R Suwiński MD)
Department of Radiotherapy Services, Mater Centre, Brisbane, QLD, Australia (Prof M Poulsen MD)
Department of Radiotherapy, Sanjay Gandhi Post Graduate Institute of Medical Sciences, Lucknow, India (S Kumar MD)
Department of Radiation Oncology, Tata Memorial Hospital, Homi Bhabha National Institute, Mumbai, India (Prof S Ghosh Laskar MD)
Département de Radiothérapie, Hôpital Pitié-Salpêtrière, Paris, France (Prof JJ Mazeron MD)
BioIRC Center for Biomedical Research, Kragujevac, Serbia (Prof B Jeremic MD)
NHMRC Clinical Trials Center, Camperdown, Australia (Prof J Simes MD)
Department of Oral and Maxillofacial-Head and Neck Oncology, Ninth People’s Hospital, College of Stomatology, Shanghai Jiao Tong University School of Medicine, Shanghai, China (Prof LP Zhong MD)
EORTC Headquarters, Brussels, Belgium (C Fortpied MSc)
NRG Oncology Statistics and Data Management Center, American College of Radiology, Philadelphia, USA (P Torres-Saavedra PhD)
Department of Radiotherapy, Centre Hospitalier Universitaire Vaudois, Lausanne, Switzerland (Prof J Bourhis MD)

**Word count:** 3579 + 362 (Research in context)/ 3500 for the manuscript text only (Microsoft Word)

**Corresponding author:**

- Dr Pierre Blanchard
  Department of Radiation Oncology
  Gustave Roussy
  114 rue Edouard Vaillant
  94 800 Villejuif, France
- Tel: 33(0)142116532
- Fax: 33(0)142115258
- E-mail: pierre.blanchard@gustaveroussy.fr
Abstract

Count: 279/300 words

Background: Randomised clinical trials (RCTs) and meta-analyses have demonstrated the survival benefit of concomitant chemoradiation (CRT) or hyperfractionated radiotherapy in the treatment of locally advanced head and neck cancer (LAHNC). However, the relative efficacy of these treatments is unknown. This study aimed to determine if one treatment was superior to the other.

Methods: Based on the individual patient data of meta-analyses evaluating the role of chemotherapy (MACH-NC) and of altered fractionation radiotherapy (MARCH), we performed a frequentist network meta-analysis using a 2-step random effects approach. The log-rank test, stratified by trial, was used. Overall survival (OS) was the primary endpoint. Global Cochran Q statistic was used to assess homogeneity and consistency and P-score to rank treatments (higher scores indicate more effective therapies).

Findings: There were 115 RCTs that yielded 154 comparisons (28,978 patients with 19,253 deaths and 20,579 progression events). Treatments were grouped into 16 modalities, for which 35 types of direct comparisons were available. Hyperfractionated radiotherapy with concomitant chemotherapy (HFCRT) was ranked as the best treatment for OS. The hazard ratios (HR) of HFCRT compared to platinum-based CRT was 0.82 [95% Confidence interval (CI) 0.66-1.01] for OS (P-score 97%). The superiority of HFCRT was robust to sensitivity analyses. Three other modalities of treatment had a better P-score but not a significantly better HR for OS than platinum-based CRT (P-score 78%): taxane-induction chemotherapy (TaxPF) followed by loco-regional treatment (P-score 89%), accelerated radiotherapy with concomitant chemotherapy (P-score 82%) and TaxPF followed by CRT (P-score 80%).

Interpretation: The results of this network meta-analysis suggest that further intensifying CRT, using HFCRT or TaxPF induction prior to CRT, could improve outcomes over CRT for the treatment of LAHNC, especially for HPV negative cancers.

Fundings: Institut National du Cancer (PHRC), Ligue Nationale Contre le Cancer, Fondation ARC.
Research in context

Evidence before this study

Individual patient data meta-analyses have demonstrated that concomitant chemoradiotherapy and hyperfractionated radiotherapy had the best efficacy results in the treatment of locally advanced non-metastatic head and neck cancer. A mixed treatment comparison based on the second publication of the Meta-Analysis of Chemotherapy in Head and Neck Cancer (MACH-NC) and on the first publication of the Meta-Analysis of Radiotherapy in Carcinomas of Head and neck (MARCH) compared six modalities of treatment. Altered fractionated concomitant chemoradiotherapy yielded the highest probability of survival.

For this network meta-analysis, trials included in the second update of MACH-NC, in the specific publication on induction chemotherapy with taxanes and in the first update of MARCH were included. The search has concerned PubMed, Scopus, Web of Science, Cochrane Controlled Trials meta-register, clinicaltrials.gov, and meeting proceedings, without language restriction, for published and unpublished “randomized trials” of “chemotherapy” or “radiotherapy” in “head and neck cancer”. Trials conducted up to December 31, 2016 were included. To improve homogeneity, trials conducted before January 1st, 1980 were excluded.

Added value of this study

Network meta-analyses allow comparing all treatment modalities with each other, using available direct and indirect comparisons (through common comparators). The median follow-up was 6.6 years overall (IQR 5.0–9.4). Hyperfractionated radiotherapy with concomitant chemotherapy had the highest probability of success for overall survival, progression-free survival, loco-regional control and cancer death. For distant control, loco-regional treatment with adjuvant chemotherapy had the best results. The other modalities of treatment that had good results were taxane, platin and fluorouracil-based induction chemotherapy followed by loco-regional treatment with or without concomitant chemotherapy and accelerated radiotherapy with concomitant chemotherapy.

Implications of all the available evidence

The results of the present network meta-analysis confirm that altered fractionated concomitant chemoradiotherapy is the most effective treatment and especially hyperfractionated radiotherapy with concomitant chemotherapy. Taxane-based induction chemotherapy followed by loco-regional treatment, ideally with concomitant CT is another good option in selected patients. Network meta-analyses have limitations due to the use of indirect information. These results would ideally need to be confirmed by randomised trials. Nevertheless, it could help guide clinical decision-making in locally
advanced head and neck cancer with a high risk of locoregional failure, especially HPV-negative tumours.
Introduction
During the past decades, advances in the treatment of locally advanced head and neck cancer have led to higher cure rates. The individual patient data (IPD) Meta-Analysis of Chemotherapy in Head and Neck Cancer (MACH-NC) has clearly demonstrated that the addition of concomitant chemotherapy (CT) to radiotherapy (RT) improves overall survival (OS), progression-free survival (PFS), locoregional control, and decreases cancer death\(^1\). A specific meta-analysis was conducted about induction CT in head and neck cancer\(^2\). The combination of taxane (docetaxel or paclitaxel), cisplatin, and fluorouracil (Tax-PF) was superior to cisplatin plus fluorouracil (PF). The Meta-Analysis of Radiotherapy in Carcinomas of Head and Neck (MARCH) was conducted and showed that altered fractionation radiotherapy (RT) was associated with a significant overall survival benefit compared with conventional fractionation\(^3\). However, the overall survival benefit was restricted to hyperfractionated RT. PFS was improved by altered fractionation RT, without a significant difference between type of fractionation, through an improvement in local and regional control. The results of these meta-analyses support head and neck cancer treatment recommendations, which mostly favor the use of conventional fractionation concomitant platinum-based chemoradiotherapy, alone or as adjuvant treatment after surgery, for locally advanced disease\(^4\).

The IPD network meta-analysis (NMA) framework has already been applied to head and neck squamous cell cancers as a methodological proof of concept. In this analysis, the treatments were lumped in six groups, and altered fractionated concomitant chemoradiotherapy had the highest probability of survival\(^5\). Since this publication, the three abovementioned IPD meta-analyses were updated and published\(^2,3,6\). All those data allowed individualizing more detailed treatment modalities. The network is now larger in terms of treatment modalities, number of trials and number of patients, and follow-up is longer. The aim of this article is to present the results of this IPD NMA of multiple treatments for locally advanced head and neck, and determine relative and absolute differences among 16 treatment modalities.
Methods

MACH-NC and MARCH Databases and Endpoint Definitions

The MACH-NC and MARCH meta-analyses comprise IPD of randomised trials conducted up to December 31, 2016, evaluating the addition of CT to local treatment (MACH-NC) and the role of RT fractionation (MARCH) in patients with locally advanced squamous cell carcinoma of head and neck. The inclusion criteria, trial search, trial flowchart, data collection, and checking have been detailed in previous publications along with the results of the standard meta-analysis. Briefly, all trials had to include non-metastatic head and neck squamous cell cancer patients, and randomize either chemotherapy or altered fractionation radiotherapy in a way that would preclude prior knowledge of the assigned treatment.

For this network meta-analysis, we have decided to exclude trials conducted before 1980 in order to improve homogeneity between trials. The primary endpoint was OS, defined as the time from randomisation until death from any cause. Secondary endpoints were event-free survival (EFS), locoregional and distant control, cancer death and non-cancer death. EFS was defined as the time from randomisation to first failure, i.e. recurrence/progression (locoregional or distant) or death. Locoregional and distant controls were defined as the time from randomisation to the occurrence of a locoregional or distant failure, respectively, and competing risks were used. If both a locoregional failure and a distant failure occurred at the same time, patients were considered as having a distant failure only. Patients without locoregional and distant failure were censored at the date of death or last follow-up. Cancer mortality included deaths from any cause in patients with a previous failure and deaths from the treated head and neck cancer. Deaths from unknown cause without previous failure were regarded as cancer mortality if they occurred within 5 years after randomisation and as non-cancer mortality otherwise.

Statistical Methods for network meta-analysis
A specific NMA statistical analysis plan was written prior to the analysis and is available here: https://www.gustaveroussy.fr/fr/meta-analyses-protocoles-dessais-orl.

A two-step method was used. The first step was to compute hazard ratios (HR) for each trial on the basis of individual patient data using the Peto estimator for OS, EFS, cancer death and non-cancer death, and a competing risk model for locoregional and distant control. The second step was to perform the network meta-analysis using a frequentist approach. Input data for each trial comparison were the two treatments compared, the logarithm of the HR, and its variance.

To limit the number of tests for both heterogeneity and inconsistency, Rücker et al have proposed a global test, called Q test. This test is a generalization of Cochran’s test that is used to assess heterogeneity in conventional meta-analyses. The Q statistic is the sum of a statistic for heterogeneity (within designs) and a statistic for inconsistency (between designs). Inconsistency can be defined as the variability of treatment effect between direct, e.g. randomized trials, and indirect comparisons at the meta-analytic level. A random effects model was used in case of heterogeneity (P value < 0.1).

Treatments were ranked using the P-score, which measures the mean extent of certainty that a treatment is better than the competing treatments. P-score would be 100% when a treatment is certain to be the best and 0% when a treatment is certain to be the worst. Five-year absolute benefit was computed using the survival rate at 5 years for the LRT-only arms as the reference and the HR was computed using the method by Stewart and Parmar for OS and EFS.

A priori sensitivity analyses for the main efficacy endpoints were (details in appendix p 38):

1. the exclusion of the outliers in the standard meta-analysis,
2. the exclusion of trials with non-conventional chemotherapy (without platinum salts, with polychemotherapy using more than two drugs other than TaxPF or with only one drug as induction chemotherapy, with adjuvant chemotherapy),
3. the exclusion of trials based on quality criteria (less than 100 patients, follow-up less than 5 years, unknown date of randomization)
4. the exclusion of MACH-NC trials with distinctive loco-regional treatment i.e. where CT is randomized but loco-regional treatments are different in both arms (variations in RT or surgery), hence introducing a confounding factor.

Further sensitivity analyses were performed for overall survival on the cluster of patients under 70 years of age and after exclusion of trials with a majority of stage I/II tumours. This work was performed in accordance with NMA guidelines. P values less than 0.05 were considered significant for the difference between treatments. All analyses were performed using R software (version 3.6.1) and the R package netmeta.

Role of the funding source

The funders of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report. The corresponding author and the first author had full access to all the data in the study and had final responsibility for the decision to submit for publication.
Results
The network consisted of 115 trials and 28,978 patients included between January 1st, 1980 and April 30, 2012. Because of a factorial or multi-arm design or distinctive loco-regional treatment in 19 trials, these 115 trials were split into 154 trial comparisons. There were 16 different treatments:

- loco-regional treatment (LRT) alone (surgery and/or radiotherapy (RT)), which was used as the reference category;
- concomitant chemoradiotherapy with or without platin-based chemotherapy (CLRT_ or CLRT_{noP});
- induction chemotherapy (IC) followed by LRT (IC-LRT) or followed by CLRT (IC-CLRT); and 3 types of IC were considered: the association of taxane, platin and 5-Fluorouracil (TaxPF), the association of platin and 5-Fluorouracil (PF) and other type of IC;
- LRT or CLRT_{noP} followed by adjuvant chemotherapy (AC) (LRT-AC or CLRT_{noP}-AC);
- hyperfractionated RT (HFRT) alone or with concomitant chemotherapy (HFCRT), moderately accelerated RT (MART), very accelerated RT (VART) and accelerated RT with concomitant chemotherapy (ACRT).

The network is presented in Figure 1. List of trials included in each treatment comparison are given in appendix (p 2) and the main characteristics of each trial are presented in appendix (p 4-20). Median follow-up based on trials (interquartile range) was 6.6 years (5.0 to 9.4).

For overall survival, the five treatments that had the highest effect were HFCRT, IC\textsubscript{TaxPF}-LRT, ACRT, IC\textsubscript{TaxPF}-CLRT and CLRT\textsubscript{P} with respective P-scores of 97%, 89%, 82%, 80% and 78% (Table 1). The league table with full results is presented in appendix (p 21). Their respective hazard ratio (HR) with their 95% confidence interval (95%CI) compared to LRT were 0.63 (0.51-0.77), 0.69 (0.56-0.85), 0.75 (0.66-0.85), 0.75 (0.62-0.92), and 0.77 (0.72-0.83). The absolute benefits at 5 years compared to LRT alone were respectively 16.7%, 13.4%, 10.4%, 10.3%, and 9.5%. The differences were not significant when comparing the five top ranking treatments between each other. The respective HRs (95%CI) of HFCRT,
\( IC_{\text{TaxPF}}\text{-LRT}, \text{ACRT}, \text{and} IC_{\text{TaxPF}}\text{-CLRT} \) in comparison to \( \text{CLRT}_p \) were 0.82 (0.66-1.01), 0.90 (0.72-1.12), 0.97 (0.86-1.10), and 0.98 (0.81-1.19). There was significant heterogeneity (\( p=0.01 \)) but no inconsistency (\( p=0.91 \)). The forest plot of the trial comparisons that included one of HFCRT, \( IC_{\text{TaxPF}}\text{-LRT}/\text{CLRT} \) or ACRT is presented in Figure 2 and number of patients and events are presented in appendix (p 23). The other trials are described in the MACH-NC induction article\(^2\), in the MARCH articles\(^{15,3}\) and in the MACH-NC articles\(^{1,6,7}\).

Some trials had no data or events for specific secondary endpoints and were excluded from the corresponding analysis (details given in appendix p 38). The results of EFS (Table 1) are in agreement with OS. Heterogeneity was still present (\( p=0.05 \)) and no inconsistency (\( p=0.52 \)) was detected for this endpoint. The five best treatments were similar to OS, although \( IC_{\text{TaxPF}}\text{-LRT} \) and \( IC_{\text{TaxPF}}\text{-CLRT} \) swopped their ranks. HFCRT was the most effective (P-score: 97%), followed by \( IC_{\text{TaxPF}}\text{-CLRT} \) (P-score: 89%), ACRT (P-score: 82%), \( IC_{\text{TaxPF}}\text{-LRT} \) (P-score: 80%), and \( \text{CLRT}_p \) (P-score: 75%). However, none of these modalities was significantly better than another (appendix p 25). Only HFCRT had significantly better results than \( \text{CLRT}_p \), with a HR (95%CI) of 0.80 (0.65-0.98). The absolute benefits at 5 years compared to LRT were 18.6% for HFCRT, 14.9% for \( IC_{\text{TaxPF}}\text{-CLRT} \), 12.5% for ACRT, 12.2% for \( IC_{\text{TaxPF}}\text{-LRT} \), and 10.8% for \( \text{CLRT}_p \).

The results of loco-regional control (LRC) (Table 1) are also in agreement with OS and EFS. Heterogeneity was still present (\( p<0.0001 \)) and inconsistency (\( p=0.0008 \)) was detected for this endpoint. The four best treatments were the same as for EFS, with HFCRT being the most effective (P-score: 88%), followed by \( \text{CLRT}_p \) and ACRT, with respective P-scores of 84% and 79%. \( IC_{\text{TaxPF}}\text{-CLRT} \) ranked fourth but \( IC_{\text{TaxPF}}\text{-LRT} \) appeared to be less effective; their respective P-scores being 78% and 36%. The modality that ranked 5th was \( IC_{\text{PF}}\text{-CLRT} \) with a P-score of 73%. When comparing the five top ranking treatments between each other, the differences were not significant, even compared to \( \text{CLRT}_p \) (appendix p 26).

The results of distant control (Table 1) are different from the other endpoints: LRT-AC being the most effective, with a P-score of 84%, followed by \( IC_{\text{PF}}\text{-LRT}, \text{CLRT}_{\text{noP}}\text{-AC}, \text{HFRT}, \) and \( IC_{\text{TaxPF}}\text{-LRT} \) with respective
P-score of 78%, 71·3%, 70·9%, and 65%. Heterogeneity and inconsistency were significant (p<0·0001) for this endpoint. Only few hazard ratios are statistically significant (appendix p 27).

When looking at cancer-specific mortality, the results of cancer death (Table 2, appendix p 28) are in agreement with OS, EFS and LRC. There was no heterogeneity (p=0·10) nor inconsistency (p=0·80) for this endpoint. The five best treatments were HFCRT, IC\_TaxPF\_LRT, CLRT\_p, ACRT, and IC\_TaxPF\_CLRT with respective P-score of 98%, 90%, 81%, 80%, and 78%. HFCRT had significantly better results than CLRT\_p and ACRT with respective HR (95%CI) of 0·77 (0·62-0·97) and 0·77 (0·61-0·97). For non-cancer death (Table 2, appendix p 29) there was no heterogeneity (p=0·81) nor inconsistency (p=0·17). None treatment modality had a significant difference with LRT.

Details of the trials excluded in sensitivity analyses are presented in appendix (p 38). For OS and EFS, the five first treatment modalities always remained consistent with HFCRT ranking first in all but one analysis (appendix p 30-31). The results of the cluster analysis in patients under 70 years of age were similar to those of the entire population analysis as well as after exclusion of trials with a majority of stage I/II tumours (appendix p 30). Heterogeneity disappeared after exclusion of outliers. For LRC and cancer death, results were also robust to sensitivity analysis. For LRC, inconsistency disappeared after exclusion of trials with non-conventional chemotherapy and the three best treatments remained unchanged. HFCRT always ranked first except in the sensitivity analysis excluding trials with distinctive loco-regional treatments (appendix p32-33). On the contrary, for distant control (appendix p 34), there was more variation in the ranking but very few comparisons were significant. Due to the small number of events, we performed an unplanned sensitivity analysis by combining treatments into seven modalities instead of 16, for distant control and non-cancer death (appendix p 35). For distant control, LRT-AC with or without concomitant CT ranked first followed by altered fractionation RT and IC-LRT, with respective P-score of 89%, 71%, and 64%; only the two first modalities had significant results compared to RT, with HR for distant metastasis of 0·23 (0·06-0·92) for LRT-AC and 0·46 (0·22-0·94) for AF-RT. For non-cancer death, there were no significant differences compared to LRT.
Discussion
The results of the present IPD network meta-analysis combining data from trials of chemotherapy (CT) and radiotherapy (RT) can be summarized as follows. Hyperfractionated RT with concomitant CT consistently ranked first for OS, EFS, LRC and cancer-specific death, and the results were robust following sensitivity analyses. The other modalities that ranked high were induction chemotherapy based on taxane, platin and fluorouracil, and accelerated RT with concomitant CT.

This work has several strengths. First, data used as input to the NMA are individual-patient data, which were checked and reanalyzed by our team, with competing risk for loco-regional and distant control. Second, the two-step frequentist NMA is a validated method\textsuperscript{10}, already used by our group\textsuperscript{16} and others\textsuperscript{17–20}. The NMA approach is also used by institutions\textsuperscript{21}. Third, the assumptions of the NMA were respected. There was no inconsistency for OS and EFS and the heterogeneity was not anymore significant after exclusion of main outliers of the standard meta-analysis without major changes in the conclusions. The transitivity assumption was theoretically respected thanks to well-defined selection criteria of studies included in the network, allowing studies to be sufficiently similar in all respects other than the treatments compared. Moreover, the difference in stage or tumour site distribution from one trial to the other is not expected to influence the results and the standard meta-analysis did not detect variation of effect according to this tumour characteristics. However, this important hypothesis cannot be formally tested. Fourth, the main results were robust to pre-defined sensitivity analyses.

This work has limitations. First, given that trials’ accrual spanned over decades, it is impossible to make sure that patients were comparable between trials. Besides, some important data, such as HPV status or smoking status, were not available. Interaction between treatment and covariates is difficult to take into account in such a large network. As age is the most important predictive factor for chemotherapy and fractionation modifications and that the benefit of concomitant chemotherapy or altered fractionation was not significant after 70 years, we have performed a sensitivity analysis without patients over 70 years that showed similar results (appendix p 30). Although the patient population
included in the NMA is large, the number of events for distant control and non-cancer death is small. The reason is that only the first event is considered and not the following ones. As a result, the analyses of these endpoints lack power even when combining treatment modalities. Moreover, the ranking of a NMA should be examined carefully, as it tends to overestimate the effect of treatment modalities with fewer trials\textsuperscript{22}. Consideration must be given to HR comparing modalities with each other. Here, HR were not significant between the top five treatments for overall survival. Concerning policy implications, few small recent trials were not included\textsuperscript{6}, nor were trials with anti-EGFRs or immunotherapy. These remarks are in lines with Hu et al. who stated that “the role of a NMA is not to provide recommendations but rather to synthesize the research in a manner that facilitates interpretation. [...] The results of network meta-analyses are a decision-supporting tool rather than a decision-making tool”\textsuperscript{23}. We have used a two-step frequentist model with IPD when one-step models are currently being developed, especially for Bayesian NMA\textsuperscript{24}. The use of Bayesian modelling could help provide credible intervals for ranking. Finally, we have not analyzed toxicity data. The data available in MACH-NC and MARCH were different with only very few toxicities in common. Thus the toxicity networks were not considered relevant. Nevertheless, it is important to put the efficacy of treatment modalities in perspective with their toxicity profile, especially since hyperfractionated radiotherapy and induction chemotherapy based on taxane, platin and fluorouracil are treatments known to be toxic.

Despite limiting the NMA to trials conducted from 1980 to 2016, readers may express concern that some trials were still conducted nearly 4 decades ago. The loco-regional treatment performed in the oldest trials is likely to be less optimal than the one performed nowadays, as surgery, anesthesia, radiotherapy techniques and supportive care have all improved over time. Imaging has also improved, and patients in older trials may have been understaged whereby even an experimental local therapy would be less effective. Additionally, the epidemiology of head and neck cancer has evolved over time, with a decrease in cancers related to tobacco and alcohol and an increase in HPV related cancers. The challenges and outcomes of these two types of cancers are quite different. Indeed, treatment for HPV-
related cancers have a better locoregional tumour control, disease-specific, and overall survival than for HPV-unrelated cancers\textsuperscript{25}. Hence, de-escalation is currently being studied for HPV-related tumours although with sobering early results\textsuperscript{26–28}. The results of our NMA suggest better outcomes with an intensification of treatment (hyperfractionated RT with concomitant CT), and this could be a strategy for p16/HPV negative tumours, although toxicity remains an important consideration since these patients may be less tolerant of intensification through this strategy due to associated co-morbidities, especially related to smoking. Although HR between the top five modalities for OS were not significant, the HR between HFCRT and conventional CLRT\textsubscript{p}, which is the accepted standard of care worldwide, was 0.82 [0.66–1.01], close to statistical significance. The corresponding HR for EFS, a validated surrogate\textsuperscript{29}, was significant (0.80 [0.65–0.98]). Moreover, the patients included in our meta-analyses have characteristics that are more consistent with HPV-negative tumours. For example, in the second publication of MARCH\textsuperscript{3,29}, with more recent studies, HPV-status was known for 17.4% of patients and was positive in only 31.0% of patients with known status. Therefore, our results would likely be applicable to patients with locally advanced HPV-negative tumours.

Hyperfractionated RT with concomitant CT (HFCRT) has been evaluated directly in seven trials included in our NMA (BiRCF\textsuperscript{31}, Duke 90040\textsuperscript{32}, EORTC 22954 (unpublished)\textsuperscript{33}, EORTC 22962 (unpublished)\textsuperscript{34}, IAR92\textsuperscript{35}, Kragujevac2\textsuperscript{36}, and SAKK10-94\textsuperscript{37}). All of these trials compared HFCRT to HFRT but one of them had a two by two design with a small number of patients (EORTC 22962\textsuperscript{34}, closed early due to slow accrual), thus HFCRT was also compared to LRT and CLRT\textsubscript{p}. None of the trials studying HFCRT was in a post-operative setting. The results for OS and EFS of these studies are reported in the upper part of Figure 2. These trials included 816 patients with only 384 patients treated in the HFCRT modality, which is a clear weakness of our analysis. A recent trial (DAHANCA 28\textsuperscript{38}) evaluated this modality of treatment in a phase I/II study of 50 patients with locally advanced HPV-negative head and neck cancer, treated with hyperfractionated, accelerated radiotherapy with concomitant weekly cisplatin and nimorazole. The 3-year actuarial LRC was 79%, and OS was 74%. Acute toxicity was high with 78% of patients requiring feeding tube. When compared with historical trials, this protocol appears to have
higher rates of late toxicity, especially with respect to feeding tube dependency and osteoradionecrosis. However, this trial was not randomised and the toxicity rate could be partly due to patient selection. It can also be argued that hyperfractionated RT is difficult to implement in the era of intensity modulation RT for head and neck cancer (none of the seven studies used this technique) but it has been done in a phase II trial with 1.25 Gy per fraction given twice a day up to 70 Gy. HFCRT is technically feasible with modern RT delivery, with an acute toxicity profile that would require adapted patient management but with acceptable long-term toxicity. It could be considered as an option for tertiary centers with a high throughput of head and neck patients.

Induction CT, especially regimens that included taxane, platin and fluorouracil, followed by loco-regional treatment and concomitant CT also yielded good results, ranking 4th for overall survival. We believe that toxic deaths that occurred before the systematic use of GCSF contributed to this ranking. In the sensitivity analysis restricted to trials mandating the use of GCSF, ICTaxPF-CLRT ranked second after HFCRT for OS, and 1st for EFS (appendix p 30-31, sensitivity analysis for outliers). Strategies with induction chemotherapy are more commonly used in clinical practice than HFCRT and this analysis partly supports this practice for advanced disease.

In conclusion, this NMA allowed evaluation of many treatment modalities, and suggests the superiority of hyperfractionated RT with concomitant CT over other treatments. This treatment, which can be difficult to implement in daily practice, could however be suitable for the treatment of HPV-negative head and neck cancers. Induction CT based on taxanes followed with ideally concomitant chemoradiotherapy is another strategy that has good results for selected patients. These treatments should ideally be further investigated in clinical trials. However, in the absence of additional randomized studies our findings can help inform current clinical decision-making.
Authors' contributions
CP, PB, and JPP with the help of the steering committee members designed and supervised the study. PB and JPP obtained funding.
PB, JB, JPP and BL searched and selected the trials. Steering committee members contributed to the identification and selection of the trials.
CP, BL, PB and JPP did the statistical analyses and wrote the draft, with revisions from the other investigators.
All authors contributed to the interpretation of the results during the investigator meeting and the revision of the manuscript. All investigators listed in Web-Appendix 1 received the manuscript for revision.

Acknowledgment
This research was funded by grants from Institut National du Cancer (Programme Hospitalier de Recherche Clinique), Ligue Nationale Contre le Cancer and Fondation ARC pour la recherche contre le cancer.
We thank the trialists and the MARCH and MACH-NC collaborative groups who agreed to share their data. The contents of this publication and methods used are solely the responsibility of the authors and do not necessarily represent the official views of the ECOG-ACRIN Cancer Research Group, and NRG Oncology.

Declaration of interest
CP reports a grant from Fondation ARC during the conduct of the study.
JPP reports grants from Ligue National Contre le Cancer, during the conduct of the study.
AA reports grants from Ligue Contre le Cancer and PHRC - Cancer National Institut, during the conduct of the study; grants from F. Hoffmann–La Roche, and from GORTEC (French Radiation and Oncology Group for Head and Neck), outside the submitted work.
EEV and QTL report personal fees outside the submitted work.
JWL reports grants from National Institutes of Health, USA, during the conduct of the study.
JJCH reports other payment from Sanofi Aventis during the conduct of the study; payment for Advisory role and conferences from Merck, Bristol Myers Squibb, Merck Sharp & Dohme España, Novartis, and Roche Pharma outside the submitted work.
All other authors declare no competing interests.
References


Figure 1 – Graphical representation of the trial network for overall survival and description of treatment modalities.
The size of the nodes is proportional to the number of patients (pts), which is given under each treatment category. The width of the lines is proportional to the number of comparisons, which are given on each line. The network included 154 comparisons from 115 trials (appendix p 2).

<table>
<thead>
<tr>
<th>Type of LRT</th>
<th>Type of CT</th>
<th>N os CT</th>
<th>Induction CT with TaxPF (IC1 vs others)</th>
<th>Induction CT with PF (IC1 vs others)</th>
<th>Induction CT with another regimen (IC1 vs others)</th>
<th>Concomitant platinum-based CT (CT+P)</th>
<th>Concomitant non-platinum-based CT (CT non-P)</th>
<th>Adjunct CT (AC)</th>
</tr>
</thead>
<tbody>
<tr>
<td>LRT alone surgery and/or RT</td>
<td>LRT</td>
<td>LRT</td>
<td>LRT</td>
<td>LRT</td>
<td>LRT</td>
<td>LRT</td>
<td>LRT</td>
<td>LRT</td>
</tr>
<tr>
<td>Concomitant chemoradiotherapy (CLRT) (Surgery)</td>
<td>LRT</td>
<td>LRT</td>
<td>LRT</td>
<td>LRT</td>
<td>LRT</td>
<td>LRT</td>
<td>LRT</td>
<td>LRT</td>
</tr>
<tr>
<td>Hyperfractionated RT (HFRT) (Surgery)</td>
<td>LRT</td>
<td>LRT</td>
<td>LRT</td>
<td>LRT</td>
<td>LRT</td>
<td>LRT</td>
<td>LRT</td>
<td>LRT</td>
</tr>
<tr>
<td>Moderately accelerated RT (MART) (Surgery)</td>
<td>LRT</td>
<td>LRT</td>
<td>LRT</td>
<td>LRT</td>
<td>LRT</td>
<td>LRT</td>
<td>LRT</td>
<td>LRT</td>
</tr>
<tr>
<td>Very accelerated RT (VART) (Surgery)</td>
<td>LRT</td>
<td>LRT</td>
<td>LRT</td>
<td>LRT</td>
<td>LRT</td>
<td>LRT</td>
<td>LRT</td>
<td>LRT</td>
</tr>
</tbody>
</table>

LRT=locoregional treatment, CT=chemotherapy, RT=radiotherapy, CLRT=LRT with concomitant chemoradiotherapy, IC=induction CT, AC=adjuvant CT, HFRT=hyperfractioned RT, HFCRT=HFRT with concomitant CT, MART=moderately accelerated RT, VART=very accelerated RT, ACRT=accelerated (moderately or very) RT with concomitant CT, P=platin-based CT, noP=non-platin-based CT, TaxPF=taxanes, platin and 5-Fluorouracil association, PF=platin and 5-Fluorouracil association.

*LRT=loco-regional treatment, CT=chemotherapy, RT=radiotherapy, CLRT=LRT with concomitant chemoradiotherapy, IC=induction CT, AC=adjuvant CT, HFRT=hyperfractioned RT, HFCRT=HFRT with concomitant CT, MART=moderately accelerated RT, VART=very accelerated RT, ACRT=accelerated (moderately or very) RT with concomitant CT, P=platin-based CT, noP=non-platin-based CT, TaxPF=taxanes, platin and 5-Fluorouracil association, PF=platin and 5-Fluorouracil association.*
Figure 2 – Forest plot for overall survival (on the left) and event-free survival (on the right), showing results from direct comparisons and network meta-analysis.

HR < 1 is in favor of the first treatment mentioned in the title (i.e., HFCRT for the comparison HFCRT vs HFRT). Detailed information about studies presented in this forest-plot are available in appendix (p 4-20).

For standard meta-analysis, results are presented with fixed and random effect, to study the impact of the heterogeneity on the choice of the model.

The number of event and patient for each study is available in appendix (p 23).

LRT=loco-regional treatment, CT=chemotherapy, RT=radiotherapy, CLRT=LRT with concomitant chemoradiotherapy, IC=induction CT, HFRT=hyperfractionated RT, HFCRT=HFRT with concomitant CT, ACRT=accelerated (moderately or very) RT with concomitant CT, P=platin-based CT, TaxPF=taxanes, platin and 5-Fluorouracil association, PF=platin and 5-Fluorouracil association.
Table 1 – Summary of Efficacy Endpoints

<table>
<thead>
<tr>
<th>Treatment data</th>
<th>Overall survival</th>
<th>Event-free survival</th>
<th>Loco-regional control</th>
<th>Distant control</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>115 trials</td>
<td>112 trials</td>
<td>110 trials</td>
<td>100 trials</td>
</tr>
<tr>
<td></td>
<td>154 comparisons</td>
<td>151 comparisons</td>
<td>150 comparisons</td>
<td>137 comparisons</td>
</tr>
<tr>
<td></td>
<td>28,978 patients</td>
<td>28,315 patients</td>
<td>27,309 patients</td>
<td>25,042 patients</td>
</tr>
<tr>
<td></td>
<td>19,253 events</td>
<td>20,579 events</td>
<td>10,882 events</td>
<td>3,065 events</td>
</tr>
<tr>
<td>P value global</td>
<td>0.07</td>
<td>0.11</td>
<td>&lt;0.0001</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>P value heterogeneity</td>
<td>0.01</td>
<td>0.05</td>
<td>&lt;0.0001</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>P value inconsistency</td>
<td>0.91</td>
<td>0.52</td>
<td>0.0008</td>
<td>&lt;0.0001</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>HR (95% CI)</th>
<th>p-score (%)</th>
<th>HR (95% CI)</th>
<th>p-score (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ref</td>
<td>HR (95% CI)</td>
<td>ref</td>
<td>HR (95% CI)</td>
</tr>
<tr>
<td>LRT</td>
<td>ref</td>
<td>21</td>
<td>ref</td>
<td>12</td>
</tr>
<tr>
<td>HFCRT</td>
<td>0.63 (0.51-0.77)</td>
<td>97</td>
<td>0.60 (0.49-0.73)</td>
<td>97</td>
</tr>
<tr>
<td>ICtaxPF-LRT</td>
<td>0.69 (0.56-0.85)</td>
<td>89</td>
<td>0.71 (0.59-0.87)</td>
<td>80</td>
</tr>
<tr>
<td>ACRT</td>
<td>0.75 (0.66-0.85)</td>
<td>82</td>
<td>0.71 (0.63-0.80)</td>
<td>82</td>
</tr>
<tr>
<td>ICtaxPF-CLRT</td>
<td>0.75 (0.62-0.92)</td>
<td>80</td>
<td>0.66 (0.55-0.80)</td>
<td>89</td>
</tr>
<tr>
<td>CLRTp</td>
<td>0.77 (0.72-0.83)</td>
<td>78</td>
<td>0.74 (0.70-0.79)</td>
<td>75</td>
</tr>
<tr>
<td>HFRT</td>
<td>0.85 (0.76-0.95)</td>
<td>61</td>
<td>0.84 (0.76-0.93)</td>
<td>54.5</td>
</tr>
<tr>
<td>CLRTnoP</td>
<td>0.89 (0.81-0.98)</td>
<td>50</td>
<td>0.88 (0.81-0.97)</td>
<td>42.7</td>
</tr>
<tr>
<td>ICtax-LRT</td>
<td>0.90 (0.82-0.99)</td>
<td>47</td>
<td>0.93 (0.85-1.02)</td>
<td>30</td>
</tr>
<tr>
<td>VART</td>
<td>0.90 (0.81-1.01)</td>
<td>46.5</td>
<td>0.88 (0.79-0.98)</td>
<td>42.8</td>
</tr>
<tr>
<td>ICtax-CLRT</td>
<td>0.90 (0.72-1.13)</td>
<td>45.5</td>
<td>0.83 (0.66-1.03)</td>
<td>54.8</td>
</tr>
<tr>
<td>MART</td>
<td>0.94 (0.87-1.01)</td>
<td>37</td>
<td>0.89 (0.83-0.96)</td>
<td>40</td>
</tr>
<tr>
<td>LRT-AC</td>
<td>1.03 (0.90-1.17)</td>
<td>18</td>
<td>0.99 (0.86-1.13)</td>
<td>17</td>
</tr>
<tr>
<td>CLRTnoP-AC</td>
<td>1.07 (0.84-1.36)</td>
<td>16</td>
<td>0.95 (0.75-1.20)</td>
<td>28</td>
</tr>
<tr>
<td>ICother-CLRT</td>
<td>1.15 (0.73-1.82)</td>
<td>15.8</td>
<td>/</td>
<td>/</td>
</tr>
<tr>
<td>ICother-LRT</td>
<td>1.04 (0.93-1.16)</td>
<td>15.2</td>
<td>1.05 (0.94-1.17)</td>
<td>6</td>
</tr>
</tbody>
</table>

Results are in bold if they are statistically significant and the three modalities of treatment with the highest p-score are highlighted in grey.

HR=hazard ratio, CI=Confidence Interval, LRT=loco-regional treatment, CT=chemotherapy, RT=radiotherapy, CLRT=LRT with concomitant chemoradiotherapy, IC=induction CT, AC=adjuvant CT, HFCRT=HFRT with concomitant CT, MART=moderately accelerated RT, VART=very accelerated RT, ACRT=accelerated RT (moderately or very) with concomitant CT, P=platin-based CT, noP=not platin-based CT, TaxPF=taxanes, platin and 5-Fluorouracil association, PF=platin and 5-Fluorouracil association.
Table 2 – Summary of Cancer death and Non-cancer death Endpoints

<table>
<thead>
<tr>
<th>Treatment data</th>
<th>Cancer death</th>
<th>Non-cancer death</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>73 trials</td>
<td>70 trials</td>
</tr>
<tr>
<td></td>
<td>104 comparisons</td>
<td>96 comparisons</td>
</tr>
<tr>
<td></td>
<td>21,753 patients</td>
<td>21,533 patients</td>
</tr>
<tr>
<td></td>
<td>11,039 events</td>
<td>3,645 events</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Cancer death (95% CI)</th>
<th>Non-cancer death (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>P value global</strong></td>
<td>0.25</td>
<td>0.57</td>
</tr>
<tr>
<td><strong>P value heterogeneity</strong></td>
<td>0.10</td>
<td>0.81</td>
</tr>
<tr>
<td><strong>P value inconsistency</strong></td>
<td>0.80</td>
<td>0.17</td>
</tr>
</tbody>
</table>

HR = hazard ratio, CI = Confidence Interval, LRT = loco-regional treatment, CT = chemotherapy, RT = radiotherapy, CLRT = LRT with concomitant chemoradiotherapy, IC = induction CT, AC = adjuvant CT, HFRT = hyperfractionated RT, HFCRT = HFRT with concomitant CT, MART = moderately accelerated RT, VART = very accelerated RT, ACRT = accelerated RT (moderately or very) with concomitant CT, P = platin-based CT, noP = not platin-based CT, TaxPF = taxanes, platin and 5-Fluorouracil association, PF = platin and 5-Fluorouracil association

Results are in bold if they are statistically significant and the three modalities of treatment with the highest p-score are highlighted in grey. See methods for the definition of cancer and non-cancer death, events used for hazard ratio of cancer and non-cancer deaths respectively.