ERJ open research
THE BEST IN OPEN ACCESS BASIC, TRANSLATIONAL & CLINICAL RESPIRATORY RESEARCH

Early View

Original research article

What is the optimal management of potentially resectable stage III-N2 NSCLC? Results of a fixed effects network meta-analysis and economic modelling

Matthew Evison, Ross Maconachie, Toby Mercer, Caitlin H Daly, Nicky J. Welton, Shahzeena Aslam, Doug West, Neal Navani

Please cite this article as: Evison M, Maconachie R, Mercer T, *et al*. What is the optimal management of potentially resectable stage III-N2 NSCLC? Results of a fixed effects network meta-analysis and economic modelling. *ERJ Open Res* 2023; in press (https://doi.org/10.1183/23120541.00299-2022).

What is the optimal management of potentially resectable stage III-N2 NSCLC? Results of a fixed effects network meta-analysis and economic modelling

Matthew Evison[1], Ross Maconachie[2], Toby Mercer[2], Caitlin H Daly[3], Nicky J. Welton[3], Shahzeena Aslam[4], Doug West[5], Neal Navani[6]

[1]Lung Cancer & Thoracic Surgery Directorate, Wythenshawe Hospital, Manchester University NHS Foundation Trust, Manchester, UK; [2]NICE Centre for Guidelines, National Institute for Health and Care Excellence; [3]Population Health Sciences, Bristol Medical School, University of Bristol, Canynge Hall, 39 Whatley Road, Bristol, UK; [4]Bedford and Addenbrooke's Cambridge University NHS Hospital Trusts; [5]Department of Thoracic Surgery; University Hospitals Bristol and Weston NHS Foundation Trust, Bristol, UK; [6]Department of Respiratory Medicine, University College London Hospital, London, UK.

Corresponding author: Dr Matthew Evison, m.evison@nhs.net

Author contributions: SA, DW, NN were members of the NICE Guideline Expert Committee for the 2019 update to the Diagnosis and Management of Lung Cancer. RM, TM, CD and NW led the NMA and economic analysis for the NICE Lung Cancer Guideline Committee which this paper reports and have had direct access to the data to verify the results. ME led the write up of the manuscript and all authors reviewed and agreed the final version. NICE have provided permission for manuscript submission. The views expressed in this manuscript are those of the authors and not necessarily those of NICE.

Conflicts of interest: RM currently works as Associate Director, Value, Access and Devolved Nations, Merck, Sharp and Dohme (UK) Ltd. During the time of this work his role was Technical Adviser, Centre for Guidelines, National Institute for Health and Care Excellence. MSD market treatments for lung cancer but this work was completed entirely while in employment with NICE and there are no obvious COI related to MSD's activities.

Funding: There was no funding provided for this study.

Abstract

Introduction: There is a critical need to understand the optimal treatment regimen in patients with potentially resectable stage III-N2 non-small cell lung cancer (NSCLC).

Methods: A systematic review of randomised controlled trials using a literature search including the CDSR, CENTRAL, DARE, HTA, EMBASE and MEDLINE bibliographic databases. Selected trials were used to perform a Bayesian fixed effects network meta-analysis and economic modelling of treatment regimens relevant to current day treatment options: chemotherapy plus surgery (CS), chemotherapy plus radiotherapy (CR) and chemoradiotherapy followed by surgery (CRS).

Findings: Six trials were prioritised for evidence synthesis. The fixed effects network meta-analyses demonstrated an improvement in Disease-Free Survival (DFS) for CRS vs CS and CRS vs CR of 0.34 years (95% CI 0.02-0.65) and 0.32 years (95% CI 0.05-0.58) respectively, over a five-year period. No evidence of effect were observed in overall survival although point estimates favoured CRS. The probabilities that CRS had a greater average survival time and greater probability of being alive than the reference treatment of CR 5 years were 89% and 86% respectively. Survival outcomes for CR and CS were essentially equivalent. The economic model calculated that CRS and CS had ICERs of £19,000/QALY and £78,000/QALY compared to CR. The probability that CRS generated more QALYs than CR and CS was 94%.

Interpretation: CRS provides an extended time in a disease-free state leading to improved cost-effectiveness over CR and CS in potentially resectable stage III-N2 NSCLC.

Introduction

Uncertainty exists as to the optimal management strategy for patients with potentially resectable stage III-N2 non-small cell lung cancer (NSCLC). Whilst consensus exists that optimal treatment must include both systemic treatment for distant control and local treatment for local control (e.g. surgery, radiotherapy), the optimal combination of treatments has not been established. This results in multiple treatment options being recommended within international lung cancer guidelines without consensus agreement as to the optimal strategy (1-7). These treatment combinations include chemotherapy plus surgery (CS), chemotherapy plus radiotherapy (CR), and chemotherapy, radiotherapy and surgery (CRS). Numerous randomised controlled trials (RCTs) and meta-analyses have failed to show one treatment combination to be definitively superior to another in overall survival (8-14) but there are notable findings within these studies that continue to spark debate. The Intergroup 0139 trial of CRS versus CR reported a significant increase in median progression free survival of 12.8 months for CRS versus 10.5 months for CR as well as the percentage of patients without disease progression at 5 years (22% versus 11%) but did not demonstrate a difference in overall survival (10). Concern was raised about a high mortality in patients undergoing pneumonectomy and a post-hoc unplanned analysis of only patients who had a lobectomy demonstrated higher median overall survival (33.6 versus 21.7 months) compared with statistically matched patients who received chemoradiotherapy. The weight that should be placed on this finding continues to be debated. Furthermore, a meta-analysis of CRS versus CR combined the results of the Intergroup 0139 study with a Nordic randomised controlled trial of CRS versus CR which recruited nearly 400 patients before closing early and was only published in abstract form. This meta-analysis was very close to reaching statistical significance for an improved survival with CRS (HR 0.87, CI 0.75 to 1.01, p=0.068) (12). Whilst these findings might represent evidence of benefit from CRS over CR, randomised controlled trials and meta-analyses of CS vs CR and CRS vs Cs have failed to show any evidence for the superiority of one treatment strategy over another. Given these findings, the ongoing debate as to the optimal treatment strategy and that different

multimodality treatments represent significant yet different health care costs, there is an urgent need to synthesise the published evidence and develop an economic model to define the most cost-effective treatment strategy in potentially resectable stage III-N2 NSCLC. This area was identified by the National Institute of Health and Care Excellence (NICE) for network meta-analysis (NMA) and health economic modelling as part of the 2019 update to its guideline on 'Lung Cancer: Diagnosis and Management' and this paper reports the results. The views expressed in this manuscript are those of the authors and not necessarily those of NICE.

Methods

We conducted a systematic review of randomised controlled trials comparing curative-intent multimodality treatments (CS, CR or CRS) in people with stage III-N2 NSCLC that were suitable for surgical resection. The literature search included the Cochrane Database of Systematic Reviews (CDSR), the Cochrane Central Register of Controlled Trials (CENTRAL), the Database of Abstracts of Reviews of Effects (DARE), the Health Technology Assessment (HTA) Database, the Excerpta Medical database (EMBASE) and the Medical Literature Analysis and retrieval System Online (MEDLINE) bibliographic databases and identified 4,241 studies for title and abstract screening. A similar search with economic filters found 956 titles and abstracts. A preferred reporting items for systematic reviews and meta-analyses (PRISMA) based checklist is available in Supplementary Data 1. Following further review, six trials were prioritised for evidence synthesis. Other trials were excluded from the analysis due to the irrelevance of the pairwise comparisons contained within them to current practice. This included trials in which CRS was given as chemotherapy followed by surgery followed by radiotherapy (15, 16). No economic studies were available to be included in the review (Figure 1). The included studies are listed in Table 1. Based on these data we, in consultation with the NICE Guideline Committee, concluded that the patients and interventions were reflective of those seen in current practice and that the trials were appropriate to pool.

Network Meta-analysis. Network meta-analysis (NMA) is a technique for quantitatively synthesising direct and indirect evidence of relative treatment effects. It is frequently used by NICE to aid guideline committee decision-making where more than two treatment options exist. As is common in cancer studies, we specified the two most important outcomes as overall survival (OS) and disease-free survival (DFS). Upon inspection of the Kaplan-Meier (KM) plots for these outcomes in the included trials, it was clear that the proportional hazards assumption seldom held because the survival curves frequently crossed or diverged. A network meta-analysis of published hazard ratios was therefore deemed inappropriate. Instead, we calculated and synthesised the area under each KM curve at the longest common follow-up time among studies. This is equivalent to the mean time

patients spent alive (OS) or alive and disease-free (DFS) within the restricted time period. Time spent in a health state is also an important input from a patient perspective and for health economic models. The longest common follow-up time among all studies was four years but we had five-year follow up data for five of the six studies, with the sixth study being the smallest, lowest quality and least applicable (17). We decided that the primary analysis would be conducted using the five-year follow-up data with four-year data being used in a sensitivity analysis because the extra information gained from the longer follow-up outweighed that from the small, low-quality trial. DFS and OS were jointly synthesised in a NMA to account for the correlation between these outcomes, and a separate NMA was specified for the probability of survival at five years to inform the economic model. All NMAs were conducted in a Bayesian framework; full methodology for data extraction and evidence synthesis, including the programming code is available online (18, 19). The fit of fixed and random effects NMA models was assessed and compared using the posterior mean of the residual deviance and deviance information criterion (DIC); lower values are preferred and differences of at least 3 points were considered meaningful (20). To assess the consistency assumption of NMA, i.e., no conflict between the direct and indirect evidence, the fit of an unrelated mean effects model was similarly compared to that of the selected NMA model (21). The NMA input data are shown in table 2.

Health Economics. We built a health economic model that accrued healthcare costs and Quality Adjusted Life Years (QALYs) for each intervention over a lifetime time horizon. We used the results from the NMAs to inform the first five years of the economic model. The NMAs dictated the time patients in each model arm spent in the disease-free and post-recurrence states as well as their probability of survival beyond five years. Patients surviving beyond five years were assumed to be disease-free and effectively cured of their NSCLC, and hence no further time was spent in the post-recurrence state after five years. The DFS and OS curves in the underpinning RCTs lent some support to this assumption by being well converged and plateauing at five years.

To inform the disease-free state beyond five years in the economic model, the proportion surviving at five years, along with an external estimate of mean time spent disease-free beyond five years were required. The absolute proportion surviving at five years in each model arm were calculated by adding the log-odds ratios of each treatment vs. CR from the NMA and the baseline log-odds probability of survival for those receiving CR, which was informed by the CR arm in van Meerbeeck 2007, the largest trial (9). As this was also the oldest trial and as OS has improved in this patient population over time, these data may not be reflective of current practice and so we tested this assumption in a sensitivity analysis. The mean time spent disease free beyond five years was calculated based on a post 5-year survival curve fitted to individual patient survival data in the Surveillance, Epidemiology, and End Results (SEER) database (Supplementary Figure). We matched the patient population in our trials to 2,865 similar patients with NSCLC stage IIIA-N2 conditional on having survived for 5 years post diagnosis (3,703 patients in the 4-year sensitivity analysis (22)).

Adverse events were not reported in all trials. Where possible, we obtained the number of grade 3+ adverse events and multiplied the Area Under the Curve (AUC) by the sample size in each arm to obtain the population years at risk and used these data to calculate the relevant incidence rate for CRS. We then fit a network meta-analysis model (23) to these data and used the resulting hazard ratios to calculate the average number of events experienced by patients in each arm, which were costed as an inpatient stay and had no quality of life decrement attached. Given the small differences between the interventions and the short-term nature of the events, on average, these simplifying assumptions were assessed as minor. The results favoured CRS over the other two interventions, which was unexpected, given that it is the most intensive intervention. These parameters were therefore omitted in sensitivity analyses.

As the DFS and OS curves were assumed to be fully converged by 5 years, we multiplied one minus the proportion of people alive by the proportion of disease recurrences that were deaths (fit using another NMA model (23) applied to pooled data from CS arms) to calculate the total number of

patients whose disease had recurred by 5 years. We costed these recurrences as being treated with platinum doublet chemotherapy, having no data on further lines of treatment or whether the probability that patients received further lines of treatment could reasonably be expected to differ between the arms. We did not cost downstream use of newer targeted and immunotherapies for NSCLC, firstly because it would have been impossible to determine what proportion of patients that generated the survival data used in our model would have received these treatments (due to either the age of the studies or individual ineligibility) and secondly because these treatments are often priced at society's maximum willingness to pay for one QALY and therefore do not affect the overall cost-effectiveness of the treatment pathway. Consequently, any related survival improvement in patients in current clinical practice over those in the trials that underpin our analysis is unlikely to have a big effect on the cost-effectiveness results.

Economic discounting within the first 5 years was resolved via a separate NMA, documented elsewhere (23), which apportioned events across those years. No directly applicable Health Related Quality of Life (HRQoL) values were available at the time of analysis, so we assigned well established values for advanced NSCLC (24) for pre and post progression advanced NSCLC to these health states within the model. This may underestimate the HRQoL of patients within our model. We also obtained data on temporary quality adjusted life year decrement from surgery (25) and applied this to the surgical arms of the model. Full tables of input parameters for the economic model are available online (23). The model's structure, input data and assumptions were validated by the NICE guideline committee and all analyses were performed in line with the NICE reference case (26).

Results

Network meta-analysis. The fixed effect model was preferred on the basis of model fit and due to insufficient data for the random effects model to be reliably estimated. The fixed effects network meta-analyses demonstrated an improvement in mean DFS time for CRS vs CS and CRS vs CR 0.34 years (95% CI 0.02-0.65) and 0.32 years (95% CI 0.05-0.58) respectively within the first five years after treatment; equating to approximately 4 months in each case (Table 3). There was no evidence of improvement between the interventions in terms of OS or probability of being alive at five years although point estimates favoured CRS. The probability that CRS had a greater average survival time than CR was 89% and there was an 86% chance that CRS had a greater probability of being alive at 5 years compared to CR. CS had similar point estimates and confidence intervals to CR for all three outcomes. The broad conclusions of the five-year analysis were replicated in the four-year sensitivity analysis (Figure 2). Inconsistency checks were performed using unrelated mean effects models (21) and no evidence of inconsistency was found. Overall, the NMA showed that CRS is associated with greater DFS than both CS and CR and there was no evidence that the interventions were more effective than the others for any other outcome.

Economic Model. The economic model calculated that CRS and CS had Incremental Cost Effectiveness Ratios (ICERs) of £19,000/QALY and £78,000/QALY compared to CR. Sensitivity analyses, varying the economic model's input parameters within plausible ranges did not alter these conclusions (Supplementary data 2). The probability that CRS generated more QALYs than CR was 94% and the probability that CRS generates more QALYs than CS was 85%. The one notable exception to this was setting the probability of being alive at five years equal among all three interventions (there was no evidence of improvement between the interventions in terms of OS, though point estimates favoured CRS and the probability of that CRS had a greater average survival time than CR was 89%), which increased the ICER for CRS vs CR to £41,000/QALY gained, although the probability that CRS generated more QALYs than CR was still very high at 89%. The ICERs were also much more favourable for the surgical options if using data on the baseline probability of

survival at five years from the more modern ESPATUE trial (8). The very high uncertainty in the ICER for CS vs CR, as evidenced by the wildly variable sensitivity analyses, is due to the very small and uncertain differences in QALYs between the two strategies.

Discussion

Key findings. Of the three interventions for potentially resectable stage III-N2 NSCLC examined in this analysis, chemoradiotherapy followed by surgery was demonstrated to be the most superior treatment in efficacy and cost-effectiveness. CRS was cost-effective at NICE's commonly accepted decision threshold of £20,000-£30,000 per QALY gained and extendedly dominated the cost-effectiveness of CS and CR. This dominance of CRS over CS and CR in cost-effectiveness was driven by the extended time patients spend in a disease-free state following CRS compared to the alternative treatment strategies and the improved quality of life associated with this. The trials included in this network meta-analysis incorporated all levels of disease burden under the umbrella of 'potentially resectable' stage III-N2. For example, one the largest trials (Intergroup 0139) included 76% of patients with a single N2 nodal station metastasis. The conclusions are, therefore, not restricted to patients with higher disease burden where the role of chemoradiotherapy has traditionally been placed.

Results in context of published literature. Other studies have also synthesised trial data in this area through meta-analysis (12-14, 27, 28) and did not find any statistically significant differences between interventions. However, the analyses in these studies are confined to conventional pairwise meta-analysis of hazard ratios and dichotomous outcomes. Furthermore, they did not include the same trials (i.e. pooling interventions that were not of interest or including studies that would not have met our protocol such as interventions unrelated to current practice and conference abstracts), secondly, we drew a distinction between CS and CRS as separate interventions rather than pooling them and thirdly, because the proportional hazards assumption does not hold for the vast majority of the OS and DFS Kaplan-Meier data in the included trials. Hazard ratios could, therefore, be considered inappropriate to pool and may not fully capture treatment differences that are seen in the differences between survival curves. It is quite common for survival curves to exhibit non-proportional hazards properties in trials of surgical vs non-surgical treatment because mortality can be initially higher (if the invasiveness of the surgery influences survival for some people) and

subsequently lower (e.g. if the surgery provides a cure) in the surgical arms. It was for this reason that it was felt more appropriate to pool data using the area-under-the-curve method rather than hazard ratios. While this method is well known in the field of health economics because the amount of time patients spend in a particular health state is crucial for QALY calculations, it is less common in clinical evidence synthesis. To illustrate these differences with a specific example we compare our study to that of Zhao et al, given this was also a network meta-analysis (27). Our results are likely to differ because: Zhao et al used hazard ratios, the interventions are disaggregated to the extent that the majority of the network is simply the same pairwise data as reported in the trials but with extra statistical uncertainty stemming from a shared random effects term, there are a lot of trials that included single modality therapies that would not have met our protocol and there is no analysis of DFS which is the outcome where we identified benefits of CRS.

Strengths and Limitations. A network meta-analysis requires consistency across the included studies in terms of trial setting, patient characteristics and treatment delivery. The only impact upon outcomes is therefore the type of treatment used and all patients within the selected studies would be eligible any of the treatments being studied within the NMA. The studies included in this NMA were well balanced for patient characteristics and conducted across similar multi-national western healthcare services. The NMA found no statistical evidence of inconsistency across the included studies providing strength to the findings and conclusions. Furthermore, this study is the first non-hazard ratio based meta-analysis of outcomes for radical treatments for potentially resectable stage III-N2 NSCLC. It included a wide range of network meta-analyses of treatment outcomes relevant to this population and restricted itself only to treatment options that are relevant to current practice. This is the first economic analysis in this patient population and both the statistical and economic work have benefited from the agreement of underlying assumptions and input parameters by a committee of experts and from examination at public consultation through the NICE Guidelines process. The conclusions of this study were robust to sensitivity and scenario analyses.

However, it is important to acknowledge that the included studies were conducted over different time periods with recruitment periods extending from 1994 to 2013. Lung cancer staging has changed significantly in this time period with the introduction of Positron Emission Tomography (PET) imaging (29) and endobronchial ultrasound (30) as well as modernisation of peri-operative care, surgical techniques and radiotherapy techniques. Overall survival estimates differed somewhat between the studies, with patients typically surviving longer in the more recent trials, which may reflect these improvements in staging and treatment as well as treatment options for distant disease recurrence in the last decade. As a matter of theory, higher baseline overall survival might provide more scope for similar relative treatment effects to achieve a greater overall magnitude of benefit. It is unlikely that this would have biased our analysis in favour of CRS, however, as the study that contributed the most weight towards the positive finding for PFS, Albain 2009, was also the second-oldest in the NMA. It should be noted that while Albain 2009 is the only study with a statistically significant DFS benefit for CRS, the point estimates for DFS at five years in all the other studies in the NMA favour CRS over its comparator. Additionally, the point estimates for OS time at five years favoured CRS over its comparator in each study included in the NMA although neither these studies nor the NMA demonstrated statistical evidence of a more effective treatment for this outcome. We also make strong reference to the post-hoc analysis in the Intergroup 0139 trial showing a highly significant improvement in survival for patients undergoing lobectomy in the CRS arm matched to patients suitable for lobectomy undergoing CR. Further potential limitations of our study include that health economic data including costs and health related quality of life were not collected in the underpinning RCTs. While it is not unusual for health economic models to combine data from disparate sources it would have been preferable to have used direct evidence. Perhaps the most important limitation in terms of cost-effectiveness is the lack of evidence of effect for the proportion of patients surviving into the long-term model, which, when set equal raised the ICER substantially.

Future implications. Whilst an established treatment for stage III-N2, CRS uptake in the UK is very low (31). The findings of this study, therefore, have significant impacts on lung cancer treatment

pathways and warrant careful consideration on patient selection, efficient transition from chemoradiotherapy to surgery and minimising the risk of failure to complete all aspects of CRS. However, the most important limitation to the findings of this study in relation to future care is how rapidly the treatment paradigm for stage III-N2 lung cancer is evolving. There are recently published randomised controlled trials that have created new standards of care, but that would not have been included in our NMA even if available at the time. These include adjuvant tyrosine kinase inhibitor (TKI) treatment in patients with epidermal growth factor receptor (EGFR) mutation positive NSCLC that has been completely resected, which significantly improves disease-free survival (HR 0.17 95% CI 0.11-0.26, p<0.0001 (32)). However, this and other trials of adjuvant treatments would not have been included in the NMA as patients were selected for this treatment after complete surgical resection, based on pathological staging (including stage II and IIIA and therefore not specific to stage III-N2), following adequate recovery from surgery and after molecular profiling of the resected tumour, i.e. randomisation occurred post-operatively. This could not be applied to a NMA of upfront treatment decisions based on clinical staging. Furthermore, the recently published PACIFIC trial has demonstrated improved DFS and OS with the addition of maintenance immunotherapy following CR in patients with stage III-N2 deemed unsuitable for surgical resection (33, 34). This trial would have also been excluded from this NMA as it selects patients unsuitable for surgery. We do note there is concern about the wide variation in the definition of resectability across a global trial and that some patients with potentially resectable N2 disease could have been included but ultimately the trial could not have been included in this NMA. In real-life clinical practice though, it is possible that chemoradiotherapy followed by adjuvant immunotherapy is being recommended in patients that have 'potentially resectable' stage III-N2 NSCLC. The one very recently published trial that might have been included in this NMA and that has potential to change practice in stage III-N2 NSCLC is Checkmate 0816 that has demonstrated the addition of neoadjuvant immunotherapy to neoadjuvant chemotherapy prior to surgical resection improves DFS significantly, HR 0.63, CI, 0.43 to 0.91; p=0.005, (35). Whilst this treatment has not been compared to CRS, the hazard ratios in this

study suggest this could represent an optimal treatment regimen in stage III-N2 whilst once again noting the trial included stages IB-IIIA and was not specific to stage III-N2. The eligibility criteria in clinical practice for this new treatment regimen are yet to be established nor is it clear whether this will be based on predictive marker testing such as programmed cell death ligand-1 (PD-L1) expression, as was the case for adjuvant immunotherapy after concurrent chemoradiotherapy in the UK. All of this new data highlights the need for expert tumour board discussions in this complex and rapidly changing field of lung cancer management. However, we strongly believe this study provides important information to support treatment decisions, particularly if there are specific scenarios in the future in which patients are not eligible for adjuvant/neoadjuvant TKI/IO therapies.

Conclusions. Overall, the results of this network meta-analysis and health economic analysis provide evidence that, in patients with potentially resectable stage IIIA-N2 NSCLC, CRS provides improved disease-free survival. Living within a disease-free state is known to be associated with improved quality of life compared to a post-disease recurrence state. It is the extended period within a disease-free state, and the assumed improvement in quality of life, that drives the improved cost-effectiveness of CRS over CR and CS in our economic model. Whilst lacking evidence of effect there are also indications towards improved overall survival with CRS. The trials included in this NMA enrolled patients over 10 years ago at least and there have been practice-changing RCTs published in the last few years relevant to this area of lung cancer treatment. The results of this NMA, however, may still provide useful insights in patients deemed ineligible for newer systemic agents such as TKIs and immunotherapy.

Table 1: Summary of trials included in the network meta-analysis; study settings, patients and interventions

STUDY	Albain 2009	Eberhardt 2015	van Meerbeeck 2007	Pless 2015	Katakami 2012	Girard 2009*
Study size	n=396	n=246	n=332	n=232	n=60	n=46
Setting	United States	Germany	Belgium/Netherlands	Switzerland/Germany/Serbia	Japan	France
Age, median (range)	60 (31 to 78)	59 (22 to 74)	61 (29 to 78)	CRT: 60 (37 to 76) CT: 59 (30 to 74)	Arm1: 57 (36 to 70) Arm2: 58 (34 to 69)	56 (NR)
Gender						
Males	63.6%	72.0%	74.1%	66.8%	66.7%	80.4%
Females	36.4%	28.0%	25.9%	33.2%	33.3%	19.6%
Performance status						
Karnofsky Score						
70 to 90	12.10%	-	-	-	-	-
90 to 100	87.90%	-	-	-	-	-
ECOG						
0	-	69%	-	69.8%	-	76.1%
1	-	30.5%	-	30.2%	-	23.9%
2	-	0.5%	-	-	-	-
Histology						
SCC	32.6%	38.5%	39.5%	33.6%	12.7%	52.2%
Adenocarcinoma	40.7	43.5%	31%	43.1%	65%	34.8%
Large cell	13.4%	22.1%	26.2%	6.5%	-	13%
mixed, other NSCLC	13.4%	22.1%	3.3%	16.8%	13.4%	-
Chemo	2x Neo 2x Adj	2x Neo	3x Neo	3x Neo	2x Neo	3x/4x/5x Neo
Dose RT (Gy)	45 to 61	45 to 71	40 to 62.5	44	40	46

*Girard 2009 was only included in exploratory sensitivity analysis

Table 2: Network meta-analysis input data; trial data for evidence synthesis (Treatment 1=CR, 2=CS and 3=CRS)

	Study	Treatment	PFS		OS		AUC Correl-ation	Survival	
			AUC	SE	AUC	SE		Probability[a]	SE
4-year data	Albain	1	1.42	0.09	2.11	0.12	0.07	0.25	0.04
		3	1.72	0.11	2.15	0.12	0.23	0.28	0.04
	Eberhardt	1	2.05	0.18	2.68	0.16	0.22	0.41	0.06
		3	2.16	0.17	2.84	0.17	0.22	0.50	0.06
	Girard	2	2.21	0.42	2.47	0.32	0.55	0.27	0.15
		3	1.65	0.34	2.14	0.32	0.44	0.24	0.11
	Katakami	2	1.47	0.24	2.60	0.23	0.14	0.31	0.09
		3	1.89	0.28	2.82	0.23	0.27	0.38	0.09
	Pless	2	1.63	0.14	2.48	0.14	0.04	0.43	0.05
		3	1.89	0.15	2.56	0.14	0.12	0.43	0.05
	van Meerbeeck	1	1.39	0.09	1.95	0.10	0.18	0.18	0.03
		2	1.36	0.10	1.79	0.11	0.24	0.21	0.03
5-year data	Albain	1	1.55	0.11	2.33	0.15	0.09	0.19	0.04
		3	1.95	0.13	2.42	0.15	0.27	0.26	0.04
	Eberhardt	1	2.41	0.23	3.09	0.21	0.25	0.41	0.06
		3	2.49	0.22	3.30	0.21	0.22	0.44	0.06
	Katakami	2	1.60	0.28	2.88	0.30	0.17	0.26	0.09
		3	2.15	0.35	3.19	0.30	0.31	0.38	0.09
	Pless	2	1.86	0.18	2.90	0.19	0.03	0.41	0.05
		3	2.13	0.19	2.94	0.18	0.13	0.35	0.05
	van Meerbeeck	1	1.52	0.12	2.11	0.12	0.23	0.14	0.03
		2	1.48	0.12	1.96	0.13	0.26	0.16	0.03

Abbreviations: AUC – area under the curve, OS – overall survival, PFS – progression free survival, SE – standard error.

[a] Probability of surviving up to 4- or 5-years.

Table 3: NMA results (CS and CRS vs CR)

		Intervention		
		Chemo-radiotherapy[a]	Chemotherapy + Surgery	Chemo-radiotherapy + Surgery
Difference in RMST (95% CrI[b])	Progression Free Life Years at 5 Years	Reference Treatment	-0.02 (-0.3, 0.26)	0.32 (0.05, 0.58)
	Post Progression Life Years at 5 Years		-0.07 (-0.43, 0.29)	-0.22 (-0.57, 0.13)
	Total Life Years at 5 Years		-0.09 (-0.38, 0.2)	0.09 (-0.19, 0.38)
Odds Ratio (95% CrI)	Being Alive at 5 Years		1.27 (0.77, 2.14)	1.25 (0.83, 1.92)

Table 4: Economic model results (absolute costs and QALYs)

Cohort ID	Name	Absolute (lifetime)		Fully incremental analysis		
		Costs	QALYs	Costs	QALYs	ICER
1	Chemoradiotherapy	£28,327	1.97682			
2	Chemotherapy and Surgery	£31,575	2.01863	£3,248	0.04181	£77,698 (vs CR)
3	Chemoradiotherapy and Surgery	£32,223	2.18170	£3,896	0.20488	£19,017 (vs CR)

References

1. Lim E, Baldwin D, Beckles M, Duffy J, Entwisle J, Faivre-Finn C, et al. Guidelines on the radical management of patients with lung cancer. Thorax. 2010;65 Suppl 3:iii1-27.
2. Postmus PE, Kerr KM, Oudkerk M, Senan S, Waller DA, Vansteenkiste J, et al. Early and locally advanced non-small-cell lung cancer (NSCLC): ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol. 2017;28(suppl_4):iv1-iv21.
3. Network NCC. Non-small cell lung cancer. NCCN Clinical Practice Guidelines in Oncology. 2019;Version 7.
4. Ramnath N, Dilling TJ, Harris LJ, Kim AW, Michaud GC, Balekian AA, et al. Treatment of stage III non-small cell lung cancer: Diagnosis and management of lung cancer, 3rd ed: American College of Chest Physicians evidence-based clinical practice guidelines. Chest. 2013;143(5 Suppl):e314S-e40S.
5. Excellence NIfHaC. Lung cancer: diagnosis and management. NICE Clinical Guideline. 2011;CG121.
6. Evison M, McDonald F, Batchelor T. What is the role of surgery in potentially resectable N2 non-small cell lung cancer? Thorax. 2018;73(12):1105-9.
7. Putora PM, Leskow P, McDonald F, Batchelor T, Evison M. International guidelines on stage III N2 nonsmall cell lung cancer: surgery or radiotherapy? ERJ Open Res. 2020;6(1).
8. Eberhardt WE, Pottgen C, Gauler TC, Friedel G, Veit S, Heinrich V, et al. Phase III Study of Surgery Versus Definitive Concurrent Chemoradiotherapy Boost in Patients With Resectable Stage IIIA(N2) and Selected IIIB Non-Small-Cell Lung Cancer After Induction Chemotherapy and Concurrent Chemoradiotherapy (ESPATUE). J Clin Oncol. 2015;33(35):4194-201.
9. van Meerbeeck JP, Kramer GW, Van Schil PE, Legrand C, Smit EF, Schramel F, et al. Randomized controlled trial of resection versus radiotherapy after induction chemotherapy in stage IIIA-N2 non-small-cell lung cancer. J Natl Cancer Inst. 2007;99(6):442-50.
10. Albain KS, Swann RS, Rusch VW, Turrisi AT, 3rd, Shepherd FA, Smith C, et al. Radiotherapy plus chemotherapy with or without surgical resection for stage III non-small-cell lung cancer: a phase III randomised controlled trial. Lancet. 2009;374(9687):379-86.
11. Pless M, Stupp R, Ris HB, Stahel RA, Weder W, Thierstein S, et al. Induction chemoradiation in stage IIIA/N2 non-small-cell lung cancer: a phase 3 randomised trial. Lancet. 2015;386(9998):1049-56.
12. McElnay PJ, Choong A, Jordan E, Song F, Lim E. Outcome of surgery versus radiotherapy after induction treatment in patients with N2 disease: systematic review and meta-analysis of randomised trials. Thorax. 2015;70(8):764-8.
13. Xu YP LB, Xu XL, Mao WM. Is there a survival benefit in patients with stage IIIA (N2) non-small cell lung cancer receiving neoadjuvant chemotherapy and/or radiotherapy prior to surgical resection: a systematic review and meta-analysis. Medicine. 2015;94(23):e879.
14. Pottgen C, Eberhardt W, Stamatis G, Stuschke M. Definitive radiochemotherapy versus surgery within multimodality treatment in stage III non-small cell lung cancer (NSCLC) - a cumulative meta-analysis of the randomized evidence. Oncotarget. 2017;8(25):41670-8.
15. Thomas M, Rube C, Hoffknecht P, Macha HN, Freitag L, Linder A, et al. Effect of preoperative chemoradiation in addition to preoperative chemotherapy: a randomised trial in stage III non-small-cell lung cancer. Lancet Oncol. 2008;9(7):636-48.
16. Sorensen JB RJ, Pilegaard HK et al. Surgery for NSCLC stages T1-3N2M0 having preoperative pathologically verified N2 involvement: A prospective randomized multinational phase III trial by the Nordic Thoracic Oncology Group. Journal of Clinical Oncology. 2013;31(15).

17. Girard N, Mornex F, Douillard JY, Bossard N, Quoix E, Beckendorf V, et al. Is neoadjuvant chemoradiotherapy a feasible strategy for stage IIIA-N2 non-small cell lung cancer? Mature results of the randomized IFCT-0101 phase II trial. Lung Cancer. 2010;69(1):86-93.
18. NICE. Lung cancer: diagnosis and management

NICE guideline [NG122]. NICE Clinical Guideline. 2019.
19. Daly CH, Maconachie R, Ades AE, Welton NJ. A non-parametric approach for jointly combining evidence on progression free and overall survival time in network meta-analysis. Res Synth Methods. 2021.
20. David J. Spiegelhalter NGB, Bradley P. Carlin,Angelika Van Der Linde. Bayesian measures of model complexity and fit. Journal of the Royal Statistical Society 2002;64(4):583-639.
21. Dias S, Welton NJ, Sutton AJ, Caldwell DM, Lu G, Ades AE. Evidence synthesis for decision making 4: inconsistency in networks of evidence based on randomized controlled trials. Med Decis Making. 2013;33(5):641-56.
22. Program SEaERS. Surveillance Epidemiology and End Results (SEER) Program - SEER*Stat Database: Incidence - SEER 9 Regs Research Data www.seer.cancer.govNov 2017 Sub (1973-2015) <Katrina/Rita Population Adjustment> - Linked To County Attributes - Total U.S., 1969-2016 Counties, National Cancer Institute, DCCPS, Surveillance Research Program [
23. Excellence NIfHaC. NG122 Lung Cancer: Diagnosis and managment. Evidence Review C. Appendix K https://www.nice.org.uk/guidance/ng122/documents/evidence-review-3. Last accessed 15th April 20222019 [
24. Nafees B, Stafford M, Gavriel S, Bhalla S, Watkins J. Health state utilities for non small cell lung cancer. Health Qual Life Outcomes. 2008;6:84.
25. Bendixen M, Jorgensen OD, Kronborg C, Andersen C, Licht PB. Postoperative pain and quality of life after lobectomy via video-assisted thoracoscopic surgery or anterolateral thoracotomy for early stage lung cancer: a randomised controlled trial. Lancet Oncol. 2016;17(6):836-44.
26. Excellence NIfHaC. Guide to the methods of technology appraisal: The reference case https://www.nice.org.uk/process/pmg9/chapter/the-reference-case. Last accessed 15th April 20222013 [
27. Zhao Y, Wang W, Liang H, Yang CJ, D'Amico T, Ng CSH, et al. The Optimal Treatment for Stage IIIA-N2 Non-Small Cell Lung Cancer: A Network Meta-Analysis. Ann Thorac Surg. 2019;107(6):1866-75.
28. Xu XL DL, Chen W, Zhu SM, Mao WM. Neoadjuvant chemoradiotherapy or chemotherapy followed by surgery is superior to that followed by definitive chemoradiation or radiotherapy in stage 111A (N2) nonsmall-cell lung cancer: a meta-analysis and system review. Onco Targets Ther. 2016;9:845-53.
29. Fischer B, Lassen U, Mortensen J, Larsen S, Loft A, Bertelsen A, et al. Preoperative staging of lung cancer with combined PET-CT. N Engl J Med. 2009;361(1):32-9.
30. Navani N, Nankivell M, Lawrence DR, Lock S, Makker H, Baldwin DR, et al. Lung cancer diagnosis and staging with endobronchial ultrasound-guided transbronchial needle aspiration compared with conventional approaches: an open-label, pragmatic, randomised controlled trial. Lancet Respir Med. 2015;3(4):282-9.
31. Adizie JB, Khakwani A, Beckett P, Navani N, West D, Woolhouse I, et al. Stage III Non-small Cell Lung Cancer Management in England. Clin Oncol (R Coll Radiol). 2019;31(10):688-96.
32. Wu YL, Tsuboi M, He J, John T, Grohe C, Majem M, et al. Osimertinib in Resected EGFR-Mutated Non-Small-Cell Lung Cancer. N Engl J Med. 2020;383(18):1711-23.
33. Antonia SJ, Villegas A, Daniel D, Vicente D, Murakami S, Hui R, et al. Overall Survival with Durvalumab after Chemoradiotherapy in Stage III NSCLC. N Engl J Med. 2018;379(24):2342-50.
34. Antonia SJ, Villegas A, Daniel D, Vicente D, Murakami S, Hui R, et al. Durvalumab after Chemoradiotherapy in Stage III Non-Small-Cell Lung Cancer. N Engl J Med. 2017;377(20):1919-29.
35. Forde PM, Spicer J, Lu S, Provencio M, Mitsudomi T, Awad MM, et al. Neoadjuvant Nivolumab plus Chemotherapy in Resectable Lung Cancer. N Engl J Med. 2022.

Figure 1: Study selection for network meta-analysis and economic modelling

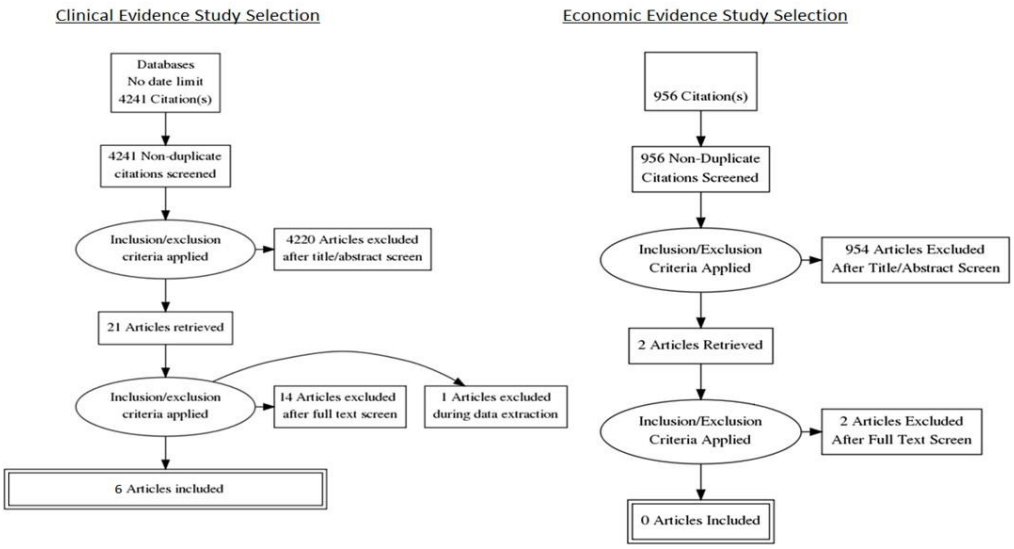
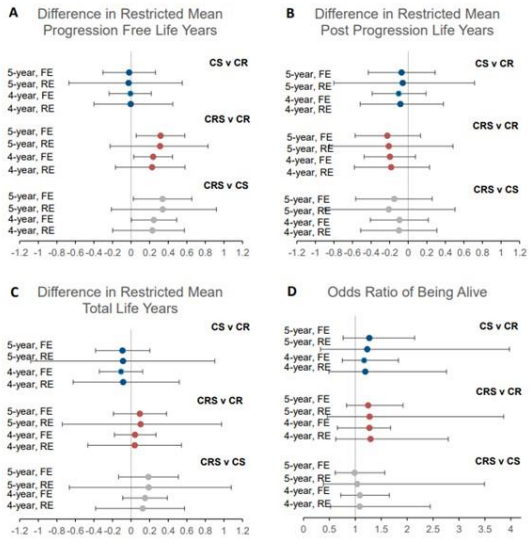


Figure 2: Difference in interventions for four key outcomes, fixed and random effects models, five and four-year data



**The only outcomes that are statistically significant are for PFS (panel A)*

Supplementary Data 1: Review Protocol

Field (based on PRISMA-P)	Content
Review question	What is the clinical and cost effectiveness of chemoradiotherapy or surgery with adjuvant treatment for the treatment for N2 stage NSCLC?
Type of review question	Intervention
Objective of the review	To provide clearer guidance regarding the treatment of N2 stage NSCLC. This question was identified during scoping meeting 2. Variation in practice has also been identified.
Eligibility criteria – population/ disease/ condition/ issue/ domain	People with stage N2 M0 NSCLC.
Eligibility criteria – intervention(s)/exposure(s)/ prognostic factor(s)	Surgery with/ without chemotherapy
Eligibility criteria – comparator(s)/control or reference (gold) standard	1. Chemoradiotherapy (radiotherapy and chemotherapy) versus 2. Tri-modality treatment

Outcomes and prioritisation	• Mortality
	  ○ Cancer-related
	  ○ Treatment-related
	  ○ All-cause
	• Quality of life (as measured by QoL instrument, for example)
	  ○ ECOG score
	  ○ EORTC score
	  ○ EQ-5D
	• Length of stay
	  ○ hospital
	  ○ ICU
	• Exercise tolerance
	• Adverse events
	  ○ Oesophagitis, pneumonitis, sepsis (grading)
	  ○ Dyspnoea
	  ○ Hypoxia and need for home oxygen
	  ○ Stroke
	  ○ Cardiovascular disease
	• Treatment-related dropout rates
	• Pain (continuous pain scales and/ or proportions of people in pain)

Eligibility criteria – study design	• RCT data. • Systematic reviews of RCTs
Other inclusion exclusion criteria	• Non English-language papers • Unpublished evidence/ conference proceedings
Proposed sensitivity/sub-group analysis, or meta-regression	No subgroup analysis identified

Selection process – duplicate screening/selection/analysis	10% of the abstracts were reviewed by two reviewers, with any disagreements resolved by discussion or, if necessary, a third independent reviewer. If meaningful disagreements were found between the different reviewers, a further 10% of the abstracts were reviewed by two reviewers, with this process continued until agreement is achieved between the two reviewers. From this point, the remaining abstracts will be screened by a single reviewer. This review made use of the priority screening functionality with the EPPI-reviewer systematic reviewing software. See Appendix B for more details.
Data management (software)	See appendix B.
Information sources – databases and dates	No date limit. See appendix C. Main Searches: • Cochrane Database of Systematic Reviews – CDSR • Cochrane Central Register of Controlled Trials – CENTRAL • Database of Abstracts of Reviews of Effects – DARE • Health Technology Assessment Database – HTA • EMBASE (Ovid) • MEDLINE (Ovid) • MEDLINE In-Process (Ovid) Citation searching will be carried out in addition on analyst/committee selected papers. The search will not be date limited because this is a new review question.
Identify if an update	Update. Original Question (linked): What is the most effective treatment for patients with resectable non-small cell lung cancer? **Recommendations that may be affected:** 1.4.27 Patients with stage I or II NSCLC who are medically inoperable but suitable for radical radiotherapy should be offered the CHART regimen. **[2005]**
Author contacts	Guideline update
Highlight if amendment to previous protocol	For details please see section 4.5 of Developing NICE guidelines: the manual
Search strategy – for one database	For details please see appendix C
Data collection process – forms/ duplicate	A standardised evidence table format will be used, and published as appendix G (clinical evidence tables) or H (economic evidence tables) of the full guideline.
Data items – define all variables to be collected	For details please see evidence tables in appendix G (clinical evidence tables) or H (economic evidence tables) of the full guideline.

Methods for assessing bias at outcome/study level	Standard study checklists were used to critically appraise individual studies. For details please see section 6.2 of Developing NICE guidelines: the manual The risk of bias across all available evidence was evaluated for each outcome using an adaptation of the 'Grading of Recommendations Assessment, Development and Evaluation (GRADE) toolbox' developed by the international GRADE working group http://www.gradeworkinggroup.org/ For further detail see Appendix B.
Criteria for quantitative synthesis (where suitable)	For details please see section 6.4 of Developing NICE guidelines: the manual
Methods for analysis – combining studies and exploring (in)consistency	For details please see the methods chapter of the full guideline. See appendix B.
Meta-bias assessment – publication bias, selective reporting bias	For details please see section 6.2 of Developing NICE guidelines: the manual. See appendix B.
Assessment of confidence in cumulative evidence	For details please see sections 6.4 and 9.1 of Developing NICE guidelines: the manual See appendix B.
Rationale/ context – Current management	For details please see the introduction to the evidence review in the full guideline.
Describe contributions of authors and guarantor	A multidisciplinary committee developed the guideline. The committee was convened by NICE Guideline Updates Team and chaired by Gary McVeigh in line with section 3 of Developing NICE guidelines: the manual. Staff from NICE Guideline Updates Team undertook systematic literature searches, appraised the evidence, conducted meta-analysis and cost-effectiveness analysis where appropriate, and drafted the guideline in collaboration with the committee. For details please see the methods chapter of the full guideline.
Sources of funding/support	The NICE Guideline Updates Team is an internal team within NICE.
Name of sponsor	The NICE Guideline Updates Team is an internal team within NICE.
Roles of sponsor	The NICE Guideline Updates Team is an internal team within NICE.
PROSPERO registration number	N/A

Supplementary Data 2: Economic model and scenario analyses results (ICERs) exploring plausible variations in the model's input parameters

Scenario	CRS vs CR	CS vs CR	CRS vs CS	Notes
Base Case (5y, FE, disc)	£19,829	£74,925	£4,151	
Base Case PSA	£19,017	£77,698	£3,973	Based on the mean of 5,000 iterations
5Y Random Effects	£20,082	£158,757	£4,064	Random rather than fixed effects NMAs used for first 5 years
No adverse events	£21,268	£68,004	£7,968	Adverse events = 0 for all treatments
Adverse events from NMA	£19,009	£72,704	£3,729	Based on NMA (see appendix J) rather than pairwise data
No treatment disutility	£18,877	£60,509	£4,163	Surgical patients suffer no post-surgery utility decrement
No long term utility decrement	£19,689	£72,305	£4,156	Standard age related utility decrements not applied
Exponential survival curve	£20,129	£81,291	£4,142	Survival in patients alive at 5 years modelled using an Exponential curve
Long term PFS costs = 100%	£21,787	£84,893	£3,829	Costs for patients surviving 5 years progression free = those within the first 5 years
Long term PFS costs = 50%	£20,563	£78,663	£4,030	Costs for patients surviving 5 years progression free half those within the first 5 years
% undergoing surgery MA = all trials	£22,148	£80,950	£5,521	% patients dropping out of surgery after chemotherapy derived from all trials in NMA
% undergoing surgery = 100%	£26,417	£100,174	£6,088	% patients dropping out of surgery after chemotherapy = 0%
Discount rate = 0%	£16,093	£33,397	£4,250	No economic discounting
4y Fixed Effects NMA	£20,205	£128,347	£6,185	NMAs are from 4 year outcomes rather than 5 year. Survival continues from 4 years
Progs that are deaths set equal	£21,178	£78,732	£4,800	% of progressions that are in fact deaths set equal among treatments
PFS Utility = 0.72	£21,214	£80,927	£4,429	Progression free utility set to lowest value from literature review
PFS Utility = 0.83	£18,770	£70,411	£3,937	Progression free utility set to highest value from literature review
Max util, Max decr between PFS and PPS	£19,595	£74,711	£4,091	PFS utility and utility decrement from progression set to highest available values
Min util, Min decr between PFS and PPS	£20,250	£75,906	£4,248	PFS utility and utility decrement from progression set to lowest available values
OR of survival set equal	£41,105	dominated	£3,805	OR of survival = 1 for CS and CRS vs CR

Cost of Surgery = CC 6+	£30,062	£123,274	£3,537	Assume cost of surgery = to most complex in class
Cost of Surgery = CC 0-2	£15,433	£54,155	£4,414	Assume cost of surgery = to least complex in class
Cost of Progressed State Halved	£27,201	£85,067	£10,734	Monthly cost of the post progression state halved
Eberhardt baseline for NMAs	£12,281	dominated	£716	Baseline population CR data from Eberhardt 2015

Supplementary Figure: diagram of economic model