in collaboration with:

university of groningen

University Medical Center Groningen

Pembrolizumab with chemotherapy for neoadjuvant and adjuvant treatment of untreated locally advanced non-metastatic triple negative breast cancer [ID1500]

Produced by	Kleijnen Systematic Reviews (KSR) Ltd in collaboration with University Medical Center (UMC) Groningen, the Netherlands (NL)
Authors	Mark Perry, Reviews Manager, KSR Ltd, United Kingdom (UK) Thea van Asselt, Health Economist, UMC Groningen, NL Lisa de Jong, Health Economist, UMC Groningen, NL Mohamed al Khayat, Health Economist, UMC Groningen, NL Maarten Postma, Health Economist, UMC Groningen, NL Rob Riemsma, Reviews Manager, KSR Ltd, UK Pawel Posadzki, Reviews Manager, KSR Ltd, UK Steven Duffy, Information Specialist, KSR Ltd, UK Charlotte Ahmadu, Health Economist, KSR Ltd, UK Nigel Armstrong, Health Economics Manager, KSR Ltd, UK Jos Kleijnen, Founder and Owner, KSR Ltd, UK Robert Wolff, Managing Director, KSR Ltd, UK
Correspondence to	Robert Wolff, Kleijnen Systematic Reviews Ltd Unit 6, Escrick Business Park Riccall Road, Escrick York, YO19 6FD United Kingdom
Date completed	30/05/2022

Source of funding: This report was commissioned by the National Institute for Health Research (NIHR) Evidence Synthesis Programme as project number 12/99/38.

Declared competing interests of the authors: None.

Commercial in confidence (CiC) data are highlighted in blue throughout the report.

Academic in confidence (AiC) data are highlighted in yellow throughout the report.

Confidential comparator prices are highlighted in green throughout the report.

Any de-personalised data are highlighted in pink throughout the report.

Rider on responsibility for report

The views expressed in this report are those of the authors and not necessarily those of the NIHR Evidence Synthesis Programme. Any errors are the responsibility of the authors.

This report should be referenced as follows:

Perry M, van Asselt T, de Jong LA, al Khayat MNMT, Postma MJ, Riemsma R, Posadzki P, Duffy S, Ahmadu C, Armstrong N, Kleijnen J, Wolff R. Pembrolizumab with chemotherapy for neoadjuvant and adjuvant treatment of untreated locally advanced non-metastatic triple negative breast cancer [ID1500]: a Single Technology Assessment. York: Kleijnen Systematic Reviews Ltd, 2022.

Contributions of authors

Mark Perry and Robert Wolff acted as project leads and systematic reviewers on this assessment, critiqued the clinical effectiveness methods and evidence and contributed to the writing of the report. Thea van Asselt acted as health economic project lead, critiqued the company's economic evaluation and contributed to the writing of the report. Lisa de Jong, Mohamed al Khayat, Maarten Postma, Charlotte Ahmadu, and Nigel Armstrong acted as health economists on this assessment, critiqued the company's economic evaluation and contributed to the writing of the report. Rob Riemsma and Pawel Posadzki acted as systematic reviewers, critiqued the clinical effectiveness methods and evidence and contributed to the writing of the report. Steven Duffy critiqued the search methods in the submission and contributed to the writing of the report. Jos Kleijnen critiqued the company's definition of the decision problem and its description of the underlying health problem and current service provision, contributed to the writing of the report and supervised the project.

Abbreviations

1L	First line
2L	Second line
3L	Third line
4L	Fourth line
AAT	Alanine aminotransferase
AC-T	Doxorubicin + cyclophosphamide followed by docetaxel
AC-TH	Doxorubicin + cyclophosphamide followed by docetaxel + trastuzumab
AE	Adverse effect/adverse event
AEOSI	Adverse event of special interest
AHRQ	Agency for Healthcare Research and Quality
AIC	Akaike information criterion
AiC	Academic in confidence
AJCC	American Joint Committee on Cancer
ASaT	All subjects as treated
ASCO	American Society of Clinical Oncology
AUC	Area under the curve
AWMSG	All Wales Medicines Strategy Group
BC	Breast cancer
BI	Budget impact
BIC	Bayesian information criterion
BNF	British National Formulary
BSA	Body surface area
CADTH	Canadian Agency for Drugs and Technologies in Health
CDSR	Cochrane Database of Systematic Reviews
CE	Cost effectiveness
CEA	Cost effectiveness analysis
CEAC	Cost effectiveness acceptability curve
CENTRAL	Cochrane Central Register of Controlled Clinical Trials
CHMP	Committee for Medicinal Products for Human Use
CI	Confidence interval
CiC	Commercial in confidence
COVID-19	Coronavirus disease 2019
CPS	Combined positive score
CS	Company submission
CSR	Clinical study report
CT	Computed tomography
CTCAE	Common Terminology Criteria for Adverse Events
DFS	Disease-free survival
DM	Distant metastasis
DSA	Deterministic sensitivity analysis
DSU	Decision Support Unit
eBC	Early breast cancer
ECOG	Eastern Cooperative Oncology Group
EFS	Event-free survival
EMA	European Medicines Agency
EORTC	European Organisation for Research and Treatment of Cancer
EQ-5D	European Quality of Life-5 Dimensions
ERG	Evidence Review Group
ESMO	European Society for Medical Oncology
FAC	Factual accuracy check
FACT-B-FBSI	Functional Assessment of Cancer Therapy Breast Symptom Index
FAS	Full analysis set
FBC	Full blood count

FDA	Food and Drug Administration
FE	Fixing errors
FEC	Fluorouracil + epirubicin + cyclophosphamide
FEC-THP	Fluorouracil + epirubicin + cyclophosphamide followed by pertuzumab + trastuzumab + taxane
FISH	Fluorescence in situ hybridization
FV	Fixing violations
GP	General practitioner
HAS	Haute Autorité de Santé
HCHS	Hospital and Community Health Services
HER2	Human epidermal growth factor receptor 2
HERC	Health Economics Research Centre
HR	Hazard ratio
HRQoL	Health-related quality of life
HSUV	Health state utility value
HTA	Health technology assessment
HUI	Health utility index
IA	Interim analysis
IC	Indirect comparison
ICER	Incremental cost effectiveness ratio
ICER	Institute for Clinical and Economic Review
IHC	Immunohistochemistry
INAHTA	Health Technology Assessment database of the International Network of Agencies for Health Technology Assessment
IO	Immune oncology
IQWiG	Institute for Quality and Efficiency in Healthcare
ISPOR	International Society for Pharmacoeconomics and Outcomes Research
ITC	Indirect treatment comparison
IV	Intravenous(ly)
KM	Kaplan-Meier
KSR	Kleijnen Systematic Reviews Ltd
LDH	Lactate dehydrogenase
LPLV	Last patient last visit
LR	Locoregional recurrence
LVEF	Left ventricular ejection fraction
LYs	Life years
LYG	Life years gained
mBC	Metastatic breast cancer
MeSH	Medical Subject Headings
MedDRA	Medical Dictionary for Regulatory Activities
MHRA	Medicines and Healthcare Products Regulatory Agency
MIMS	Monthly Index of Medical Specialities
MJ	Matters of judgement
MRI	Magnetic resonance imaging
MSD	Merck Sharp & Dohme
mTNBC	Metastatic triple-negative breast cancer
N/A	Not applicable
NCCN	National Comprehensive Cancer Network
NCI	National Cancer Institute
NHS	National Health Service
NICE	National Institute for Health and Care Excellence
NIHR	National Institute for Health Research
NL	The Netherlands
NMA	Network meta-analysis
NR	Not reported

NYHA	New York Heart Association
OS	Overall survival
PAS	Patient Access Scheme
pCR	Pathological complete response
PD	Progressed disease
PFS	Progression-free survival
PRESS	Peer Review of Electronic Search Strategies
PRISMA	Preferred reporting items for systematic reviews and meta-analyses
PRO	Patient reported outcome
PROMIS-Fatigue SF1	Patient-Reported Outcomes Measurement Information System Fatigue Short Form-1
PS	Performance status
PSA	Probabilistic sensitivity analysis
PSS	Personal Social Services
PSSRU	Personal Social Services Research Unit
PTC	Pertuzumab + trastuzumab + chemotherapy
Q3W	Every three weeks
QALY	Quality adjusted life year
QLQ-BR23	Breast Cancer-Specific Quality of Life Questionnaire
QLQ-C30	Quality of Life Questionnaire
QoL	Quality of life
QTSQ	Cancer Treatment Satisfaction Questionnaire
QTWIST	Quality-adjusted time without symptoms or toxicity
QW	Once weekly
RCT	Randomised controlled trial
RECIST	Response Evaluation Criteria in Solid Tumours
ROB2	Cochrane Risk of Bias tool version 2
RR	Relative risk; Risk ratio
SABCS	San Antonio Breast Cancer Symposium
SAE	Serious adverse effects
ScHARRHUD	School of Health and Related Research health utilities database
SD	Standard deviation
SE	Standard error
SEER	Surveillance, Epidemiology, and End Results
SF-6D	Short-Form Six-Dimension
SF-36	36-Item Short Form Survey;
SIGN	Scottish Intercollegiate Guidelines Network
SLR	Systematic literature review
SMC	Scottish Medicines Consortium
SoC	Standard of care
STA	Single technology appraisal
TA	Technology assessment
TCH	Docetaxel + carboplatin + trastuzumab
TC-HP	Docetaxel + carboplatin + trastuzumab + pertuzumab
TEAE	Treatment emergent adverse events
TNBC	Triple-negative breast cancer
UK	United Kingdom
UMC	University Medical Center
VAS	Visual analogue scale
WTP	Willingness-to-pay

Table of Contents

Table of Tables

Table of Figures

1. EXECUTIVE SUMMARY

This summary provides a brief overview of the key issues identified by the Evidence Review Group (ERG) as being potentially important for decision making. If possible, it also includes the ERG's preferred assumptions and the resulting incremental cost-effectiveness ratios (ICERs).

Section 1.1 provides an overview of the key issues. Section 1.2 presents the key model outcomes. Section 1.3 discusses the decision problem, Section 1.4 issues relate to the clinical effectiveness, and Section 1.5 issues relate to the cost effectiveness. A summary is presented in Section 1.6.

Background information on the condition, technology and evidence and information on key as well as non-key issues are in the main ERG report, see Sections 2 (decision problem), 3 (clinical effectiveness) and 4 (cost effectiveness) for more details.

All issues identified represent the ERG's view, not the opinion of the National Institute for Health and Care Excellence (NICE).

1.1 Overview of the ERG's key issues

Table 1.1: Summary of key issues

Issue #	Summary of issue	Report Section
1	Choice of population: There are major differences between the population defined in the NICE final scope and the decision problem addressed in the CS, e.g. regarding inflammatory disease, early-stage disease, participants being at high risk of recurrence and with a pre-defined ECOG PS.	2.1
2	Choice of comparator: Placebo alone is used while CS indicated that the addition of capecitabine to systemic treatment is associated with improvement in DFS, i.e. the best available comparator in the adjuvant phase might actually be capecitabine.	2.3
3	Geographical effects: Only a small subset of participants were from the UK. Subgroup analysis, based on a small dataset, suggests that geographical area is an important covariate influencing outcome, and so the observed effects may not be applicable to the UK.	3.2.1
4	TNM staging: Details on participants with stage I, II and III disease, respectively, were provided but not for the four detailed TNM grades in the inclusion criteria. As grades relate to prognosis, it is vital to know if the ratio of TNM grades is equivalent to those in the UK population.	3.2.3
5	ECOG staging: Subgroup analyses results indicated potential differences between Eastern Co-operative Oncology Group (ECOG) performance status, especially that compared to ECOG 0 participants, ECOG 1 participants did not demonstrate benefits from pembrolizumab in terms of pCR.	3.2.5.5
6	Adverse effects: Although AEs were described to be comparable between arms, the ERG notes that the risk of deaths in the pembrolizumab arm was three times that of the placebo arm. Furthermore, there was a difference in [redacted], see also Key Issue 8.	3.2.6
7	The company's model structure does not include health states for remission from LR and separate pre- and post-progression states for DM. For the ERG, this does not reflect clinical practice, i.e. the company's model does not capture costs and utilities related to these health states correctly.	4.2.2

Issue #	Summary of issue	Report Section
8	By far the largest gain in survival and QALYs is obtained in the extrapolated EFS part of the model. When using only the observed part (short time horizon), where mortality is increased in the pembrolizumab arm due to adverse events (see Key Issue 6), the ICER increases dramatically. The company has chosen to use different types of parametric distributions for the extrapolations, proper justification for this is lacking according to the ERG.	4.2.6
9	The probabilities of moving to DM (from the LR state) and death (from LR and DM state) are assumed to be constant over the entire time horizon of the model. The ERG is concerned about the lack of clinical justification for this.	4.2.6
10	The ERG considers the use of KEYNOTE-355 data as base case for the DM survival to be a potential source of bias. Although the company argues that KEYNOTE-355 is to be preferred over KEYNOTE-522 data because KEYNOTE-522 data are not sufficiently mature in the DM state, there are quite substantial differences in observed survival between these two studies.	4.2.6
11	The utility for the DM health state is relatively low compared to utilities for comparable health states in literature, which may be due to the limited number of questionnaires from patients who experienced distant metastasis in the KEYNOTE-522 trial which was used to inform this utility value. This causes doubts about the validity of the use of this utility value in the model.	4.2.8

AE = adverse effect; CS = company submission; DM = distant metastasis; ECOG = Eastern Cooperative Oncology Group; EFS = event-free survival; ERG = Evidence Review Group; ICER = incremental cost effectiveness ratio; LR = locoregional recurrence; NICE = National Institute for Health and Care Excellence; pCR = pathological complete response; PS = performance status; QALY = quality-adjusted life year; UK = United Kingdom

1.2 Overview of key model outcomes

NICE technology appraisals compare how much a new technology improves length (overall survival) and quality of life in a quality-adjusted life year (QALY). An incremental cost effectiveness ratio (ICER) is the ratio of the extra cost per QALY gained.

Overall, the technology is modelled to affect QALYs by:

- An increase in event free survival (EFS) at a relatively high utility
- A relatively lower utility in the locoregional recurrence (LR) and distant metastasis (DM) states where proportionally more chemotherapy patients reside

Overall, the technology is modelled to affect costs by:

- Its higher treatment acquisition price compared to chemotherapy alone in both the neoadjuvant and the adjuvant phase
- The higher metastatic (one-off) treatment costs for the chemotherapy arm

The inputs that have most impact on the ICERs are those related to parameters linked to EFS extrapolations followed by metastatic treatment costs. Scenarios in the company submission (CS) that have a substantial impact on the ICER are the scenarios varying the distributions for the extrapolation of EFS, and the scenario with a limited time horizon (20 years).

1.3 The decision problem: summary of the ERG's key issues

The decision problem addressed in the CS is in line with the final scope issued by NICE regarding the intervention and the outcomes addressed. However, the population and comparator were not completely aligned with the NICE remit, see Tables 1.2 and 1.3.

Table 1.2: Key issue 1: Choice of population

Report Section	**2.1**
Description of issue and why the ERG has identified it as important	While the ERG acknowledges the need to align with the marketing authorisation, it notes some major differences between the population defined in the NICE final scope and the decision problem addressed in the CS, e.g. regarding inflammatory disease, early-stage disease, participants being at high risk of recurrence and with a pre-defined ECOG PS. Of note, according the CS, *"KEYNOTE-522 is a Phase III pivotal RCT investigating the efficacy of Pembrolizumab plus chemotherapy vs chemotherapy as neoadjuvant therapy followed by pembrolizumab vs placebo as adjuvant therapy in participants with locally advanced, inflammatory, or early-stage triple-negative breast cancer at high risk of recurrence"*. The use of "or" could indicate that different permutations of these factors are possible, e.g. that participants in the trial had inflammatory and early-stage TNBC which was not locally advanced. This ambiguity adds further uncertainty to the differences described before.
What alternative approach has the ERG suggested?	Closer coherence to the NICE scope would have ensured that efficacy and safety were being specifically evaluated in the specified population
What is the expected effect on the cost effectiveness estimates?	The ambiguity around the population breadth, i.e. it is unclear whether the trial population is actually narrower or broader than the NICE scope population, makes it very difficult to estimate effects on cost effectiveness.
What additional evidence or analyses might help to resolve this key issue?	Further evidence in a subgroup more closely aligned with the NICE scope population.

CS = company submission; ECOG = Eastern Cooperative Oncology Group; ERG = Evidence Review Group; NICE = National Institute for Health and Care Excellence; PS = performance status; RCT = randomised controlled trial; TNBC = triple-negative breast cancer

Table 1.3: Key issue 2: Choice of comparator

Report Section	**2.3**
Description of issue and why the ERG has identified it as important	Choice of comparator. In the adjuvant phase of the trial, placebo alone is used as the comparator. This is based on the CS statement that active therapy is not standard treatment in the adjuvant phase according to expert opinion. However, it is stated in the CSR that the addition of capecitabine to systemic treatment is associated with improvement in DFS ([redacted]), which suggests that the best available comparator in the adjuvant phase might actually be capecitabine. Therefore, whilst it may be true that current practice does not commonly use adjuvant therapies (such as capecitabine), it is likely that the trial's use of placebo in the

Report Section	2.3
	adjuvant phase, rather than an active comparator such as capecitabine, may contribute to an increased estimate of benefit for pembrolizumab.
What alternative approach has the ERG suggested?	Including capecitabine as an active comparator in the adjuvant phase could be considered in future trials.
What is the expected effect on the cost effectiveness estimates?	The overly favourable comparator to pembrolizumab in the adjuvant phase (placebo only) may possibly enhance the overall measure of efficacy and thus augment cost effectiveness relative to what might be observed had capecitabine been part of the trial regimen.
What additional evidence or analyses might help to resolve this key issue?	Additional data collection with a subgroup using capecitabine in the adjuvant phase.
CI = confidence interval; CS = company submission; CSR = clinical study report; DFS = disease-free survival; ERG = Evidence Review Group; HR = hazard ratio	

1.4 The clinical effectiveness evidence: summary of the ERG's key issues

The ERG identified two major concerns with the evidence presented on the clinical effectiveness, related to quality of life and adverse events

Table 1.4: Key issue 3: Geographical effects

Report Section	3.2.1
Description of issue and why the ERG has identified it as important	Only a small sub-set of participants were from the UK. Crude subgroup analysis suggests that geographical area is an important covariate influencing outcome, and so the overall effects observed may not necessarily be applicable to the UK.
What alternative approach has the ERG suggested?	The ERG also specifically requested all results to be sub-grouped for 1) Europe versus rest of world and 2) UK versus rest of world. The company provided EFS data showing that the Europe subgroup had a less favourable relative effect size for pembroli-zumab (HR [redacted] compared to the rest of the world subgroup (HR [redacted] suggesting that the overall data in the trial might be providing an overly optimistic picture for European patients. The company did not provide similar data for a UK patient subgroup.
What is the expected effect on the cost effectiveness estimates?	Based on the available European data, the trial effectiveness results may be more favourable than they might be for a European-based population (such as the UK), and thus cost effectiveness may be inflated. The ERG implemented a simple fix to the efficacy in the model, assuming the HR to remain constant over time, see Section 6.1.
What additional evidence or analyses might help to resolve this key issue?	UK-specific data would help in addressing this issue. Furthermore, in order to explore the impact of regional difference in effectiveness, the model structure would need to be adapted more elaborately
CI = confidence interval; EFS = event-free survival; ERG = Evidence Review Group; HR = hazard ratio; UK = United Kingdom	

Table 1.5: Key issue 4: TNM staging

Report Section	3.2.3
Description of issue and why the ERG has identified it as important	The CS provides details of the numbers of participants in KEYNOTE-522 with Stage I, II and stage III disease, but not the four detailed TNM gradings mentioned in the inclusion criteria. It is likely that stage relates to prognosis, and so it is vital to know if the ratio of TNM stages in the trial is equivalent to ratios of TNM stages in the UK population.
What alternative approach has the ERG suggested?	In response to the request for clarification, the company provided precise data on the TNM stages for the two arms of the trial, but the company were unable to provide data on the UK prevalence of TNM stages, stating that *"data for TNM grading for TNBC patients is not available from publicly available data"*.
What is the expected effect on the cost effectiveness estimates?	Unclear
What additional evidence or analyses might help to resolve this key issue?	Information on the TNM stages in the UK population would allow a better judgement on the external validity of the trial. For example, if the trial contains a greater prevalence of lower TNM stages than the UK population, this may allow more meaningful interpretation of effect sizes.
CS = company submission; ERG = Evidence Review Group; UK = United Kingdom	

Table 1.6: Key issue 5: ECOG staging

Report Section	3.2.5.5
Description of issue and why the ERG has identified it as important	The ERG noted that subgroup analyses results indicated potential differences between ECOG PS. In particular, in contrast to the ECOG=0 subgroup, the subgroup with ECOG=1 did not demonstrate benefits from pembrolizumab in terms of pCR. The company stated that numbers were small and that therefore it was difficult to form conclusions, but the data suggest that patients with an ECOG status of 1 are unlikely to benefit from pembrolizumab (and there is a probability that the drug could even cause harm in this group, although this is uncertain).
What alternative approach has the ERG suggested?	When asked to comment on this finding, the company described the characteristics expected to be associated with an ECOG of 1. Attempts to adjust for these covariates were made by the company in post-hoc analyses, which, as expected, removed the negative effects of the highly correlated ECOG variable upon the outcome. These did not show anything other than confirm the evident correlation. The important point is that these correlating characteristics do not prevent people with ECOG 1 being less appropriate candidates for pembrolizumab, and the fact remains that if people have an ECOG score of 1 they are probably not going to experience benefits from pembrolizumab.
What is the expected effect on the cost effectiveness estimates?	For people with an ECOG score of 1, pembrolizumab is unlikely to be cost effective.
What additional evidence or analyses might help to resolve this key issue?	Further evidence with greater numbers of people with an ECOG score of 1.

Report Section	3.2.5.5
ECOG = Eastern Cooperative Oncology Group; ERG = Evidence Review Group; pCR = pathological complete response; PS = performance status	

Table 1.7: Key issue 6: Adverse effects

Report Section	3.2.6
Description of issue and why the ERG has identified it as important	Adverse effects are described as comparable between arms. However, the risk of deaths in the pembrolizumab arm was three times that of the placebo arm. Furthermore, the difference between arms in [redacted], see also Key Issue 8.
What alternative approach has the ERG suggested?	Not applicable
What is the expected effect on the cost effectiveness estimates?	Adverse effects have been included in the economic model, see Section 4.2.7, however, the ERG wanted to bring this is issue to the attention of the committee. This is linked to the Key Issue 8.
What additional evidence or analyses might help to resolve this key issue?	None
ERG = Evidence Review Group	

1.5 The cost effectiveness evidence: summary of the ERG's key issues

A full summary of the cost effectiveness evidence review conclusions can be found in Section 6.4 of this report. The company's cost effectiveness results are presented in Section 5, the ERG's summary and detailed critique in Section 4, and the ERG's amendments to the company's model and results are presented in Section 6. The key issues are discussed in Tables 1.8 to 1.12.

Table 1.8: Key issue 7: Model structure not including locoregional remission and no differentiation between pre-progression and post-progression distant metastatic patients.

Report Section	4.2.2
Description of issue and why the ERG has identified it as important	The company's model structure does not include health states for remission from locoregional recurrence and separate pre- and post-progression states for distant metastasis. The ERG believes this does

Report Section	4.2.2
	not reflect clinical practice, and therefore the company's model does not capture costs and utilities related to these health states correctly.
What alternative approach has the ERG suggested?	The ERG asked for a scenario based on the model structure of TA424, which did include remission from locoregional recurrence and separate pre- and post-progression distant metastasis health states, but this scenario was not provided by the company. The ERG was not able to adjust the model structure, as no data was available to inform remission and separate progression distant metastasis states.
What is the expected effect on the cost effectiveness estimates?	Overall impact on cost-effectiveness is uncertain as adding health states may have consequences in both directions.
What additional evidence or analyses might help to resolve this key issue?	A sensitivity analysis with an alternative model structure, including the remission and separate progression states would help to explore the impact of this issue.
ERG = Evidence Review Group; TA = technology appraisal	

Table 1.9: Key issue 8: Modelled treatment effectiveness and extrapolation for EFS state likely overestimates effectiveness of pembrolizumab

Report Section	4.2.6
Description of issue and why the ERG has identified it as important	By far the largest gain in survival and QALYs is obtained in the extrapolated EFS part of the model. When using only the observed part (short time horizon), where mortality is increased in the pembrolizumab arm due to adverse events (see Key Issue 6), the ICER increases dramatically. The company has chosen to use different types of parametric distributions for the extrapolations, proper justification for this is lacking according to the ERG.
What alternative approach has the ERG suggested?	The ERG base case uses the same type of distribution (but still individually fitted) in both arms to extrapolate EFS. This will not fully eliminate the issue that most of the QALY gain is obtained outside of the observed period.
What is the expected effect on the cost effectiveness estimates?	The ICER increases.
What additional evidence or analyses might help to resolve this key issue?	Mature comparative data on long-term EFS, and more extensive validation of the results by clinical experts.
EFS = Event-free survival; ERG = Evidence Review Group; ICER = incremental cost effectiveness ratio; QALY = quality-adjusted life year	

Table 1.10: Key issue 9: Constant transition probabilities from LR and DM states assumed without clinical justification

Report Section	4.2.6
Description of issue and why the ERG has identified it as important	The probabilities of moving to DM (from the LR state) and death (from LR and DM state) are assumed to be constant over the entire time horizon of the model. The ERG is concerned about the lack of clinical justification for this.

What alternative approach has the ERG suggested?	No alternative approach.
What is the expected effect on the cost effectiveness estimates?	Uncertain.
What additional evidence or analyses might help to resolve this key issue?	Mature data on transition probabilities over time, possibly obtained from further KEYNOTE-522 data cuts, could resolve this uncertainty.
DM = distant metastasis; ERG = Evidence Review Group; LR = locoregional recurrence	

Table 1.11: Key issue 10: The use of KEYNOTE-355 data for DM survival may not be appropriate

Report Section	**4.2.6**
Description of issue and why the ERG has identified it as important	The ERG considers the use of KEYNOTE-355 data as base case for the DM survival to be a potential source of bias. Although the company argues that KEYNOTE-355 is to be preferred over KEYNOTE-522 data because KEYNOTE-522 data are not sufficiently mature in the DM state, there are quite substantial differences in observed survival between these two studies.
What alternative approach has the ERG suggested?	The ERG prefers to use KEYNOTE-522 data to estimate transition probabilities from the DM state to death, as already presented in a company scenario. The ERG has added an additional feature to this scenario where treatment costs are adjusted accordingly.
What is the expected effect on the cost effectiveness estimates?	When only adjusting for DM survival, the ICER changes very little, but when also adjusting for treatment costs, the ICER increases.
What additional evidence or analyses might help to resolve this key issue?	Mature data on transition probabilities over time, possibly obtained from further KEYNOTE-522 data cuts, could resolve this uncertainty.
DM = distant metastasis; ERG = Evidence Review Group; ICER = incremental cost effectiveness ratio	

Table 1.12: Key issue 11: Relatively low utility in the DM health state

Report Section	**4.2.8**
Description of issue and why the ERG has identified it as important	The utility for the DM health state is relatively low compared to utilities for comparable health states in literature, which may be due to the limited number of questionnaires from patients who experienced distant metastasis in the KEYNOTE-522 trial which was used to inform this utility value. This causes doubts about the validity of the use of this utility value in the model.
What alternative approach has the ERG suggested?	To provide separate utility estimates for progressed and not-progressed patients with distant metastasis from the KEYNOTE-522. Since this was not possible due to the design of the trial and the limited number of questionnaires available in this group, the ERG considered it appropriate to conduct additional scenario analyses based on utility values for a comparable health state in patients with TNBC from literature.

Report Section	4.2.8
What is the expected effect on the cost effectiveness estimates?	The incremental QALYs are expected to decrease, resulting in an increased ICER. This was confirmed by the scenarios the company conducted in response to clarification question B19c and the scenarios conducted in the ERG model.
What additional evidence or analyses might help to resolve this key issue?	More data on the utility for patients experiencing DM in time, obtained from further KEYNOTE-522 data cuts, could resolve this uncertainty. Additionally, mature subsequent treatment data obtained from further KEYNOTE-522 data cuts may be used to estimate utilities for not-progressed and progressed patients with distant metastasis separately (line of treatment can be used as a proxy for progression status).
DM = distant metastasis; ERG = Evidence Review Group; ICER = incremental cost effectiveness ratio; QALY = quality-adjusted life year	

1.6 Summary of the ERG's view

The following tables summarise the ERG's changes to the company's base case to arrive at an ERG base case (Table 1.13). In addition, Tables 1.14 and 1.15 present the ERG scenarios.

Table 1.13: Deterministic ERG base case

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
CS base case					
Pembrolizumab + chemotherapy	[redacted]	[redacted]			
Chemotherapy	[redacted]	[redacted]	[redacted]	[redacted]	£5,940
Fixing errors 1: Enable pembrolizumab 1L treatment in DM state for IO-eligible patients in the placebo arm					
Pembrolizumab + chemotherapy	[redacted]	[redacted]			
Chemotherapy	[redacted]	[redacted]	[redacted]	[redacted]	£9,346
Fixing errors 2: Adjustment to formulas correcting for general population mortality					
Pembrolizumab + chemotherapy	[redacted]	[redacted]			
Chemotherapy	[redacted]	[redacted]	[redacted]	[redacted]	£5,976
Matters of judgement 1: Correction for efficacy of pembrolizumab adjusting for Europe versus rest of the world hazard ratio					
Pembrolizumab + chemotherapy	[redacted]	[redacted]			
Chemotherapy	[redacted]	[redacted]	[redacted]	[redacted]	£7,801
Matters of judgement 2: Use KEYNOTE-522 data to inform survival in DM state and alongside this adjust treatment costs according to the shorter survival					
Pembrolizumab + chemotherapy	[redacted]	[redacted]			
Chemotherapy	[redacted]	[redacted]	[redacted]	[redacted]	£8,976

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
Matters of judgement 3: Use lognormal distributions in EFS for both arms					
Pembrolizumab + chemotherapy	[redacted]	[redacted]			
Chemotherapy	[redacted]	[redacted]	[redacted]	[redacted]	£16,444
1L = first line; CS = company submission; DM = distant metastasis; EFS = event-free survival; ERG = Evidence Review Group; ICER = incremental cost effectiveness ratio; IO = immune oncology; QALY = quality-adjusted life year					

Table 1.14: Deterministic scenario analyses (conditional on ERG base case)

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
ERG base case					
Pembrolizumab + chemotherapy	[redacted]	[redacted]			
Chemotherapy	[redacted]	[redacted]	[redacted]	[redacted]	£43,621
Scenario 1: Limit time horizon to 5 years (similar to the observed period)					
Pembrolizumab + chemotherapy	[redacted]	[redacted]			
Chemotherapy	[redacted]	[redacted]	[redacted]	[redacted]	£397,435
Scenario 2: Set the cut-off of the piecewise model at 68 weeks instead of 50 weeks					
Pembrolizumab + chemotherapy	[redacted]	[redacted]			
Chemotherapy	[redacted]	[redacted]	[redacted]	[redacted]	£27,172
Scenario 3: Use generalized gamma distributions for EFS in both arms					
Pembrolizumab + chemotherapy	[redacted]	[redacted]			
Chemotherapy	[redacted]	[redacted]	[redacted]	[redacted]	£15,447
Scenario 4: Use lognormal distribution for pembrolizumab and generalized gamma distribution for placebo EFS					
Pembrolizumab + chemotherapy	[redacted]	[redacted]			
Chemotherapy	[redacted]	[redacted]	[redacted]	[redacted]	£53,592
Scenario 5: Adjust utility in DM health state based on KEYNOTE-355					
Pembrolizumab + chemotherapy	[redacted]	[redacted]			
Chemotherapy	[redacted]	[redacted]	[redacted]	[redacted]	£44,259
Scenario 6: Adjust utility in DM health state based on KEYNOTE-119					
Pembrolizumab + chemotherapy	[redacted]	[redacted]			
Chemotherapy	[redacted]	[redacted]	[redacted]	[redacted]	£44,362
ERG base case					
Pembrolizumab + chemotherapy	[redacted]	[redacted]			

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
Chemotherapy	[redacted]	[redacted]	[redacted]	[redacted]	£43,621
CS = company submission; EFS = event-free survival; ERG = Evidence Review Group; ICER = incremental cost effectiveness ratio; QALY = quality-adjusted life year					

Table 1.15: Probabilistic scenario analyses (conditional on ERG base case)

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)	Prob-ability
ERG base case						
Pembrolizumab + chemotherapy	[redacted]	[redacted]				
Chemotherapy	[redacted]	[redacted]	[redacted]	[redacted]	£43,621	31.9%
Scenario 1: Limit time horizon to 5 years (similar to the observed period)						
Pembrolizumab + chemotherapy	[redacted]	[redacted]				
Chemotherapy	[redacted]	[redacted]	[redacted]	[redacted]	£381,768	0.0%
Scenario 2: Set the cut-off of the piecewise model at 68 weeks instead of 50 weeks*						
Pembrolizumab + chemotherapy	[redacted]	[redacted]				
Chemotherapy	[redacted]	[redacted]	[redacted]	[redacted]	£37,272	50.8%
Scenario 3: Use generalized gamma distributions for EFS in both arms						
Pembrolizumab + chemotherapy	[redacted]	[redacted]				
Chemotherapy	[redacted]	[redacted]	[redacted]	[redacted]	£16,697	79.0%
Scenario 4: Use lognormal distribution for pembrolizumab and generalized gamma distribution for placebo EFS						
Pembrolizumab + chemotherapy	[redacted]	[redacted]				
Chemotherapy	[redacted]	[redacted]	[redacted]	[redacted]	£58,421	28.1%
Scenario 5: Adjust utility in DM health state based on KEYNOTE-355						
Pembrolizumab + chemotherapy	[redacted]	[redacted]				
Chemotherapy	[redacted]	[redacted]	[redacted]	[redacted]	£44,568	31.4%
Scenario 6: Adjust utility in DM health state based on KEYNOTE-119						
Pembrolizumab + chemotherapy	[redacted]	[redacted]				
Chemotherapy	[redacted]	[redacted]	[redacted]	[redacted]	£44,685	31.4%
CS = company submission; DM = distant metastasis; EFS = event-free survival; ERG = Evidence Review Group; ICER = incremental cost effectiveness ratio; QALY = quality-adjusted life year						
*Errors in approximately ten PSA runs. Errors were excluded from the analysis to obtain the results						

2. CRITIQUE OF COMPANY'S DEFINITION OF DECISION PROBLEM

Table 2.1: Statement of the decision problem (as presented by the company)

	Final scope issued by NICE	Decision problem addressed in the CS	Rationale if different from the final NICE scope	ERG comment
Population	Adults with previously untreated locally advanced, nonmetastatic triple-negative breast cancer (TNBC).	Adults with locally advanced, inflammatory, or early-stage TNBC at high risk of recurrence.	Wording to reflect licence wording.	There are differences in the population defined in the final scope issued by NICE and the decision problem addressed in the CS which are discussed in Section 2.1.
Intervention	Pembrolizumab in combination with standard neoadjuvant chemotherapy followed by adjuvant pembrolizumab.	Pembrolizumab in combination with standard neoadjuvant chemotherapy followed by adjuvant pembrolizumab.	N/A	The ERG has no comments.
Comparator(s)	Standard neoadjuvant/adjuvant therapy without pembrolizumab.	Carboplatin + paclitaxel followed by doxorubicin/epirubicin + cyclophosphamide (neoadjuvant phase only) followed by placebo monotherapy (adjuvant phase).	To reflect KEYNOTE-522 and clinical expert opinion which notes that after neoadjuvant chemotherapy patients do not receive additional adjuvant chemotherapy in England.	The comparator might not represent the available comparator in the adjuvant phase, as detailed in Section 2.3.
Outcomes	• Overall survival (OS) • Pathological complete response (pCR) • Event-free survival (EFS) • Adverse effects (AEs) of treatment • Health-related quality of life (HRQoL)	• OS • pCR • EFS • AEs of treatment • HRQoL	N/A	The ERG has no comments.

Based on Table 1 and pages 10 to 12 of the CS[1]
CS = company submission; ERG = Evidence Review Group; N/A = not applicable; NICE = National Institute for Health and Care Excellence; pCR = pathological complete response; TNBC = triple-negative breast cancer

2.1 Population

The population relevant for this submission is defined in four different places:

1. The final scope issued by the National Institute for Health and Care Excellence (NICE) defined the population of interest as *"adults with previously untreated locally advanced, non-metastatic triple-negative breast cancer"* (TNBC).[3]
2. In the decision problem addressed in the company submission (CS), the population is defined as *"adults with locally advanced, inflammatory, or early stage triple-negative breast cancer at high risk of recurrence"*.[1] As noted in Table 2.1, this is *"to reflect licence wording"*.[1]
3. According to Table 4 of Appendix D of the CS, the population of interest for the systematic literature review (SLR) was *"early-stage and locally advanced non-metastatic TNBC"*.[4]
4. According to page 17 of the CS, the *"only relevant study identified by the systematic literature review"*, KEYNOTE-522 was conducted *"...in participants with locally advanced, inflammatory, or early-stage triple-negative breast cancer at high risk of recurrence"*.[1] Table 3 of the CS added that participants with Eastern Cooperative Oncology Group (ECOG) performance status (PS) of 0 or 1 were included.[1]

ERG comment: There are a number of differences between the populations assessed in 1) the NICE final scope; 2) the decision problem addressed in the CS; 3) the inclusion criteria for the SLR reported in the CS; and 4) the inclusion criteria of KEYNOTE-522, as detailed in Table 2.2. The Evidence Review Group (ERG) looked into these differences in more detail:

- Adults: Although the term "adults" was not mentioned in either the SLR inclusion criteria or the KEYNOTE-522 inclusion criteria, according to Table 5 of the CS, the range of included participants was 22 to 80 years, i.e., did not include non-adult participants.[1]
- Previously untreated: Although not reflected in the decision problem or the SLR inclusion criteria, the only study identified by the SLR reported in the CS, KEYNOTE-522, included *"untreated newly diagnosed"* patients hence this appears to be in line with the NICE final scope.[1, 3]
- Non-metastatic: According to Table 5 of the CS, 100% of participants included in KEYNOTE-522 were non-metastatic.[1] Therefore, although not clearly stated in the CS decision problem or the inclusion criteria reported for KEYNOTE-522, this is in line with the NICE final scope.[1, 3]
- Inflammatory: According to the Trial Design Overview (reported on page 3308 of the clinical study report (CSR) for KEYNOTE-522), [redacted].[2]
- Early-stage: The NICE final scope did not specify that the population of interest included "early-stage" patients.[3]
- High risk of recurrence: The company has been asked to define the term as this could be considered an important factor defining the population and thus its likely response to the intervention.[5] The company responded by stating that *"within KEYNOTE-522, 'high-risk TNBC' is synonymous with 'locally advanced TNBC', the latter defined as T1c, N1-N2; T2-T4d, N0-N2 (thus, Stage II-III) per the American Joint Committee on Cancer (AJCC) staging criteria for breast cancer"*.[6]

- ECOG PS: While the NICE final scope did not specify the population regarding ECOG PS, the trial inclusion criteria specified participants to have ECOG PS 0 or 1, i.e. the population is narrower than that defined in the NICE scope.

In the clarification letter, the company (Merck Sharp & Dohme (MSD)) has been asked to justify the discrepancies between the NICE final scope and the decision problem addressed in the CS.[5] In response, the company noted that the anticipated marketing authorisation, which the definition used in the decision problem is based on, was included in the response to the draft scope, however, *"this was marked as commercial in confidence and as such NICE were not able to make this wording public"*.[6] According to the response to the request for clarification, the Committee for Medicinal Products for Human Use (CHMP) has now adopted a positive opinion and the wording published on the website of the European Medicines Agency (EMA) is *"KEYTRUDA [pembrolizumab], in combination with chemotherapy as neoadjuvant treatment, and then continued as monotherapy as adjuvant treatment after surgery, is indicated for the treatment of adults with locally advanced, or early stage triple negative breast cancer at high risk of recurrence"*.[6, 7]

Furthermore, the company has been asked to discuss how any discrepancy in population definitions may influence how trial results should be extrapolated to clinical practice in the National Health Service (NHS) in England and Wales.[5] The response from the company is that *"MSD understands the definition applied in the KEYNOTE-522 resonates with NHS clinical practice"*.[6]

While the ERG acknowledges the need to align with the marketing authorisation, it notes some major differences between the population defined in the NICE final scope and the decision problem addressed in the CS, e.g. regarding inflammatory disease, early-stage disease, participants being at high risk of recurrence and with a pre-defined ECOG PS (see Table 2.2). These differences are noted as a key issue 1 for consideration of the committee.

Of note, according to page 17 of the CS, *"KEYNOTE-522 is a Phase III pivotal RCT* [randomised controlled trial] *investigating the efficacy of Pembrolizumab plus chemotherapy vs chemotherapy as neoadjuvant therapy followed by pembrolizumab vs placebo as adjuvant therapy in participants with locally advanced, inflammatory, or early-stage triple-negative breast cancer at high risk of recurrence"*.[1] The use of "or" could indicate that different permutations of these factors are possible, e.g. that participants in the trial had inflammatory and early-stage triple-negative breast cancer (TNBC) which was not locally advanced. This ambiguity adds further uncertainty to the differences described before.

Table 2.2: Detailed comparison of population in NICE final scope and CS decision problem

NICE final scope	CS decision problem	SLR inclusion criteria	KEYNOTE-522 inclusion criteria	ERG comment
Table on page 2 of the scope[3]	Table 1 of the CS[1]	Table 4 of Appendix D of the CS[4]	Table 3 and page 17 of the CS[1]	
Adults	Adults	-	-	As detailed above, unlikely that non-adults included
Previously untreated	-	-	Untreated newly diagnosed	"Previously untreated" not reflected in CS decision problem or SLR but in trial
Locally advanced	Locally advanced	Locally advanced	Locally advanced	Identical
-	-	-	Centrally confirmed	Inclusion criterion of trial narrower than NICE final scope
Non-metastatic	-	Non-metastatic	-	"Non-metastatic" not reflected in CS decision problem but no non-metastatic participants in the trial
-	Inflammatory	-	Inflammatory	"Inflammatory" not included in NICE final scope, see comment above
-	Early-stage	Early-stage	Early-stage	"Early stage" not included in NICE final scope
TNBC	TNBC	TNBC	TNBC	Identical
-	High risk of recurrence	-	High risk of recurrence	"High risk of recurrence" not included in NICE final scope
-	-	-	ECOG PS of 0 or 1	Population included in the trial is narrower than NICE final scope

Based on Table 1 of the CS[1]
CS = company submission; ECOG = Eastern Cooperative Oncology Group; ERG = Evidence Review Group; NICE = National Institute for Health and Care Excellence; PS = performance status; SLR = systematic literature review; TNBC = triple-negative breast cancer

2.2 Intervention

The intervention defined in the CS (*"pembrolizumab in combination with standard neoadjuvant chemotherapy followed by adjuvant pembrolizumab"*)[1] is in line with the NICE final scope definition (*"pembrolizumab in combination with standard neoadjuvant chemotherapy followed by adjuvant pembrolizumab"*)[3].

Pembrolizumab is administered in a neoadjuvant and adjuvant phase. In the neoadjuvant phase, 200 mg of pembrolizumab is given intravenously (IV) on day 1 of each 21-day cycle (Q3W) for 8 cycles, alongside:

- Cycles 1 to 4: Carboplatin area under the curve (AUC) 5-day Q3W (or AUC 1.5 weekly) + paclitaxel 80 mg/m^2 once weekly (QW)
- Cycles 5 to 8: Doxorubicin 60 mg/m^2 Q3W or epirubicin 90 mg/m^2 and cyclophosphamide 600 mg/m^2 Q3W

In the adjuvant phase, pembrolizumab is given as 200 mg IV Q3W for 9 cycles. No other therapeutic agents are given in this phase.[1]

2.3 Comparators

The description of the comparators in the NICE final scope is *"standard neoadjuvant/adjuvant therapy without pembrolizumab"*.[3] However, the company has set the comparator as *"carboplatin + paclitaxel followed by doxorubicin/epirubicin + cyclophosamide (neoadjuvant phase only) followed by placebo monotherapy (adjuvant phase)"*.[1] Crucially, this involves only placebo monotherapy for the adjuvant phase. This is justified in the CS by the fact that current United Kingdom (UK) practice does not use adjuvant treatment.[1]

ERG comment: The ERG identified the following points:

1. The best available comparator in the adjuvant phase (capecitabine) may have been overlooked.
2. The company had not adequately initially justified the exclusion of taxanes and anthracyclines
3. No justification was initially given for the choice between doxorubicin and epirubicin
4. There is a better efficacy for doxorubicin than epirubicin in the trial, and more received doxorubicin, but it is unclear if this reflects the proportion of doxorubicin use in the population

These points will now be described in detail.

1. *The best available comparator in the adjuvant phase (capecitabine) may have been overlooked.*

It is also stated in the CSR that [redacted]

Whilst it may be true that current practice does not commonly use adjuvant therapies (such as capecitabine), it is likely that the trial's use of placebo in the adjuvant phase, rather than an active comparator such as capecitabine, may contribute to an increased estimate of benefit for pembrolizumab. Thus, whilst this observed benefit may be realistic in terms of comparison to *established* practice,

therefore fulfilling the criteria outlined in the NICE final scope, it might not tell the committee how much better pembrolizumab is than the best available alternative approaches, established or not. This is key issue 2.

In the request for clarification, the ERG asked the company to explain how including capecitabine as a *"an active comparator in the adjuvant phase might have changed findings in the trial (the intervention would have been capecitabine + pembrolizumab)"*.[5] However, the question was not clearly answered as the company appeared to have misunderstood the question, assuming the question was how any *"off-study adjuvant capecitabine use"* might have affected results, which would indeed be very different, because it would assume that capecitabine might be used reactively and off-protocol, such as for people not achieving pathological complete response (pCR), and might thus lead to confounding.[6]

Related to this, another question in the request for clarification asked how any presumed differences in results between the two scenarios, the actual scenario and the scenario where capecitabine is part of the trial, might be accounted for in any sensitivity analyses.[5] Due to the reasons described before, the response did not clarify this issue.[6]

2. *The company had not adequately initially justified the exclusion of taxanes and anthracyclines*

The ERG also noted that for adjuvant treatment after surgery, NICE guidelines for early and locally advanced BC (NG101) recommends offering a regimen that contains both a taxane and an anthracycline, but that the company had not justified the exclusion of taxanes and anthracyclines in the adjuvant phase.[1] Therefore, the company was asked to justify the comparison to only placebo instead of taxane and an anthracycline as adjuvant treatment.[5] The company response was focussed on the point that because the patients had already received the drugs at the neoadjuvant phase it would not be clinically indicated for them to receive them at the adjuvant phase as well:[6]

"A taxane and anthracycline regimen for the treatment of early-stage breast cancer is generally given either before or after surgery with curative intent, but not both before and after surgery as neoadjuvant and adjuvant chemotherapy treatment, respectively. For chemotherapy, neoadjuvant vs adjuvant administration of a taxane and anthracycline regimen is considered equivalent in terms of distant recurrence, breast cancer mortality or death from any cause for breast cancer patients.[8] *The adjuvant guidelines within NG101 do not make a recommendation of what a clinician should do if a patient has already received a taxane and anthracycline in the neoadjuvant setting. As mentioned above and per common clinical practice, such a patient would not be also treated with the same adjuvant chemotherapy regimen. Furthermore, use of anthracycline is limited by a maximum exposure dose due to cardiotoxicity and adjuvant administration of a neoadjuvant chemotherapy regimen that did not result in a pathological complete response (pCR) is not recommended. A relevant clinical practice example comes from the HER2+ breast cancer space, as women who received a neoadjuvant anthracycline + taxane regimen are not treated with the same chemotherapy agents in the adjuvant setting; however, anti-HER2 treatment is given both before and after surgery independent of the surgical outcome (pCR vs not). UK Clinical experts have informed MSD that the treatments used in KEYNOTE-522 reflect the current standard of care for neoadjuvant and adjuvant treatment of TNBC where a taxane and anthracycline regimen given either before or after surgery with curative intent. From the perspective of the clinical evidence base, the early breast cancer systematic literature review conducted to support this submission did not identify any relevant publications that explored the effectiveness and safety of adjuvant taxane and/or anthracycline after administration of a neoadjuvant chemotherapy regimen (see Appendix D1.2.1). Since no relevant publications were retrieved, it was not possible to incorporate neoadjuvant chemotherapy followed by an anthracycline/taxane adjuvant treatment option via an indirect treatment comparison within the model".*

The ERG notes that there is supporting evidence. However, it did not find any papers countering the company's view that chemotherapy should only be given in one phase and not both. The ERG also looked at the SR in the NICE NG 101 guideline, which also did not provide any counterevidence to challenge the company's assertion. The ERG would have preferred to have found more objective data-based backing to confirm the fact that anthracycline/taxane chemotherapy can only be given in one phase but realises that such decisions are often made on the basis of clinical experience and consensus.

The ERG did also consider the point that the chemotherapy need not have been given at the neoadjuvant phase but could have been given at the adjuvant phase instead. The systematic review submitted by the company did support the notion that adjuvant and neoadjuvant chemotherapy are equivalent for the most important outcomes (although adjuvant chemotherapy may be better for local recurrence), thus suggesting that the placing of chemotherapy in the neoadjuvant phase was not disadvantageous. In any event, the ERG realised that any shift of chemotherapy to the adjuvant phase would not have solved the problem of pembrolizumab being compared to placebo alluded to earlier (with its implications for potentially exaggerated pembrolizumab effect sizes).[8] This is because it would simply have shifted the problem of *pembrolizumab versus placebo* in the adjuvant phase to *pembrolizumab versus placebo* in the neoadjuvant phase.

3. *No justification was initially given for the choice between doxorubicin and epirubicin*

For the comparator treatment in the second part of the neoadjuvant phase, a choice is made between doxorubicin and epirubicin, but no justification is given for this in the CS. In the request for clarification, the company was asked why this choice was made, who in the study was responsible for making the choice, and upon which criteria the choice was made.[5] The company responded that *"doxorubicin and epirubicin are commonly used neoadjuvant anthracycline regimens for TNBC. The choice of treatment was made by the investigator at the initiation of the second phase of neoadjuvant treatment and was largely dependent on local/ institutional guidance and guidelines"*.[6]

4. *There is a better efficacy for doxorubicin than epirubicin in the trial, and more received doxorubicin, but it is unclear if this reflects the proportion of doxorubicin use in the population*

Lastly, the ERG requested a comparison with NHS clinical practice in terms of the use of these treatments.[5] The company response was that *"the combination of doxorubicin or epirubicin plus cyclophosamide is available in NHS clinical practice"* but no further information was given.[6] The ERG also requested information on the implications of any difference, and a subgroup analysis of results by doxorubicin / epirubicin use.[5] The company responded with a subgroup analysis for event free-survival (EFS), where within both chemotherapy subgroups a point estimate favouring the pembrolizumab arm was observed. However, the benefit in the doxorubicin subgroup was stronger [pembrolizumab versus placebo HR 0.56. 95% CI 0.40 to 0.80] than the non-significant effect observed in the epirubicin subgroup [pembrolizumab versus placebo HR 0.78, 95% CI 0.47 to 1.31]. These results may be important because more than twice as many participants received doxorubicin (488 versus 238) in the trial. If the distribution of these drugs is different in the UK population, with a more equal distribution, or even a weighting in the opposite direction, then the distribution in the trial may be affecting external validity. More information on the UK distribution would be helpful to address this issue.

2.4 Outcomes

The NICE final scope lists the following outcome measures:[3]

- Overall survival (OS)

- Pathological complete response (pCR)
- Event-free survival (EFS)
- Adverse effects (AEs) of treatment
- Health-related quality of life (HRQoL)

These were all assessed in the KEYNOTE-522 trial.

ERG comment: These outcomes are in line with the NICE scope.

2.5 Other relevant factors

According to the CS, *"pembrolizumab plus chemotherapy followed by adjuvant pembrolizumab monotherapy does not meet the end-of-life criteria"*.[1]

The company does not envisage any equality issues with the use of pembrolizumab in combination with chemotherapy for neoadjuvant and adjuvant treatment of untreated locally advanced non-metastatic TNBC.[1]

According to the company, pembrolizumab in combination with chemotherapy followed by pembrolizumab monotherapy in the adjuvant setting is an innovative treatment option in this therapy area as the first immunotherapy agent to be appraised by NICE for use in early-stage locally advanced BC patients which are at high risk of relapse.[1]

On 26 July 2021, the Food and Drug Administration (FDA) approved pembrolizumab for high-risk, early-stage, TNBC in combination with chemotherapy as neoadjuvant treatment, and then continued as a single agent as adjuvant treatment after surgery.[9]

3. CLINICAL EFFECTIVENESS

3.1 Critique of the methods of review(s)

A SLR was conducted to identify relevant clinical evidence. Full details of the SLR search strategy, study selection process and results were reported in Appendix D.[4]

3.1.1 Searches

The following section contains a summary and critique of all searches related to clinical effectiveness presented in the CS.[1, 4] The Canadian Agency for Drugs and Technologies in Health (CADTH) evidence-based checklist for the Peer Review of Electronic Search Strategies (PRESS), was used to inform this critique.[10, 11] The CS was checked against the Single Technology Appraisal (STA) specification for company/sponsor submission of evidence.[12]

Appendix D of the CS provided details of the literature searches conducted for the SLR of clinical efficacy and safety outcomes.[4] Database searches were conducted on 27 July 2021. A summary of the resources searched is provided in Table 3.1.

Table 3.1: Resources searched for the clinical effectiveness systematic review (as reported in the CS)

Resource	Host/Source	Date Ranges	Date searched
Electronic databases			
MEDLINE and In-Process & Other Non-Indexed Citations	Ovid	1946 to July 26, 2021	27/07/21
Embase	Ovid	1974 to 2021 July 27	27/07/21
CENTRAL	EMB Reviews, Ovid	June 2021	27/07/21
Trials registries			
ClinicalTrials.gov	Internet	-	27/07/21
EU Clinical Trial Registry	Internet	-	27/07/21
Conference proceedings			
ASCO	NR	2020-2021	27/07/21
ESMO	NR	2020	27/07/21
SABCS	NR	2020	27/07/21

ASCO = American Society of Clinical Oncology; CENTRAL = Cochrane Central Register of Controlled Clinical Trials; CS = company submission; ESMO = European Society for Medical Oncology; NR = not reported; SABCS = San Antonio Breast Cancer Symposium

ERG comment:

- The CS provided full details of the literature searches for the ERG to appraise.[1, 4]
- A good range of databases, clinical trials registries and conference proceedings were searched.
- Full details of the database searches, including the database name, host platform, date range and date searched, were provided.
- Trials registers were searched, but details of the search strategies or search terms used, dates of searches, and results were not reported in the CS.[1, 4] Details of the ClinicalTrials.gov search strategy, date of search, and results, were provided in response to the ERG clarification letter.

Details of the search results were provided for the EU Clinical Trial Registry; details of the search strategy and date of search were not provided.[6]

- Conference proceedings were searched. The search strategies or search terms used, date of searches, and results, were not reported in the CS.[1, 4] In response to the ERG clarification letter, details of the search terms used, date of searches, URL links, and number of abstracts included in the SLR, were provided.[6]
- The database search strategies were well structured, transparent and reproducible. They included truncation, proximity operators, synonyms, and subject headings (Medical Subject Headings (MeSH) and EMTREE). There were no date limits.
- MeSH terms were used instead of EMTREE in the Embase search strategy, though the Ovid host platform does map to the correct subject heading when the search is conducted. Several MeSH and EMTREE terms were exploded when there were no terms beneath them in the tree hierarchy.
- The population facet of search terms could have been improved with more synonyms, fewer exact phrases, better use of proximity operators, and the removal of redundant terms/phrases. The combination of search terms for 'triple negative breast cancer' with search terms for 'breast cancer' using the Boolean AND was incorrect but had barely any impact on the search results.
- There were a number of redundant search lines in the intervention/comparator facet of search terms.
- The searches were limited to English language only studies and this may have introduced language bias. Best practice states that *"to reduce the risk of introducing bias, searches should not be restricted by language"*.[13] Any limits (including language) should be reported and justified according to PRISMA (Preferred reporting items for systematic reviews and meta-analyses) 2020 and PRISMA-S guidelines.[14-16]
- It would have been preferable for the database search strategies to be presented exactly as run, rather than copied into a tabular format, as item 8 of the PRISMA-S checklist recommends.[16] The Cochrane Handbook also recommends that *"...bibliographic database search strategies should be copied and pasted into an appendix exactly as run and in full, together with the search set numbers and the total number of records retrieved by each search strategy. The search strategies should not be re-typed, because this can introduce errors"*.[17]
- Study design search filters for randomised controlled trials (RCTs) designed by the Scottish Intercollegiate Guidelines Network (SIGN) were included in the search strategies, and were cited, as current practice recommends.[16]
- Separate searches for safety outcomes were not conducted. It is unlikely that efficacy searches that include study design filters for RCTs will be sensitive enough to identify safety data. Ideally, searches for safety outcomes should be carried out alongside the searches for efficacy.[18]
- The searches were conducted in July 2021. An update of the searches immediately prior to submission to NICE would have been appropriate and could have identified potentially relevant records published since July 2021.

3.1.2 Inclusion criteria

The company conducted a SLR following a pre-defined study eligibility criteria outlined in Table 3.2. Two reviewers independently screened all references retrieved from the search, critiqued in Section 3.1.1 of this report, both at title and abstract, and full text screening stages. To reach consensus, discrepancies in screening results were resolved by involving a third reviewer.

Table 3.2: Eligibility criteria used in search strategy for RCT and non-RCT evidence

	Inclusion criteria
Population	Early-stage and locally advanced non-metastatic TNBC
Interventions	**Pembrolizumab regimens:** • Pembrolizumab (200 mg Q3W x 4 cycles) + carboplatin (AUC 5 Q3W x 4 cycles or AUC 1.5 qw x 4 cycles) + paclitaxel (80 mg/ml qw x 4 cycles) • Pembrolizumab (200 mg Q3W x 4 cycles) + doxorubicin (60 mg/m2) or epirubicin (90mg/ml^2) + cyclophosphamide (600mg/ m^2 Q3W x 4 cycles) • Post-surgery: Pembrolizumab (200 mg Q3W x 9 cycles) **Preferred regimens:** • Dose-dense doxorubicin + cyclophosphamide followed by paclitaxel every three weeks • Dose-dense doxorubicin + cyclophosphamide followed by weekly paclitaxel • Docetaxel + cyclophosphamide **Other regimens:** • Dose-dense doxorubicin + cyclophosphamide • Doxorubicin + cyclophosphamide every 3 weeks (category 2B) • Cyclophosphamide + methotrexate + fluorouracil • Doxorubicin + cyclophosphamide followed by docetaxel every 3 weeks • Doxorubicin + cyclophosphamide followed by weekly paclitaxel • Epirubicin + cyclophosphamide • Docetaxel + doxorubicin + cyclophosphamide • Carboplatin + paclitaxel (80 mg/ml QW x 4 cycles) • Paclitaxel every 3 weeks followed by dose-dense doxorubicin + cyclophosphamide/ epirubicin/cyclophosphamide • Paclitaxel weekly followed by dose-dense doxorubicin + cyclophosphamide • Paclitaxel every 3 weeks followed by doxorubicin + cyclophosphamide/ epirubicin/cyclophosphamide • Paclitaxel weekly followed by doxorubicin + cyclophosphamide/ epirubicin/cyclophosphamide • Nab-paclitaxel followed by (dose-dense) doxorubicin + cyclophosphamide/ epirubicin/cyclophosphamide • Nab-paclitaxel + carboplatin followed by (dose-dense) doxorubicin + cyclophosphamide/ epirubicin/cyclophosphamide **Immunotherapy agents:** • Atezolizumab + nab-paclitaxel • Atezolizumab + nab-paclitaxel followed by atezolizumab + dose-dense doxorubicin + cyclophosphamide
Comparators	Any of the interventions listed above
Outcomes	**Efficacy outcomes:** • Pathological complete response (pCR) • Event-free survival (EFS) • Disease-free survival (DFS) • Overall survival (OS) • Landmark survival rates • Landmark EFS • Landmark DFS • Treatment duration/time to treatment discontinuation **Safety outcomes:**

	• Any adverse events • Any Grade 3 or higher adverse events • Immune-related toxicity • Treatment-emergent adverse events (any Grade, and Grade 3 or higher) • Study withdrawals **Patient-reported outcomes, including quality of life measures:** • EQ-5D • EORTC QLQ-C30 • QLQ-BR23 • FACT-B-FBSI
Time	Most recent 15 years
Study design	Phase II and III RCTs Parallel group (triple-blind/double-blind) RCT - cross over (triple-blind/double-blind) RCT - post hoc and open-label extension
Language	Only studies published in English

Based on Table 4 of CS Appendices[4]
AUC = area under the curve; CS = company submission; DFS = disease-free survival; EFS = event-free survival; EQ-5D = European Quality of Life-5 Dimensions; EORTC = European Organisation for Research and Treatment of Cancer; FACT-B-FBSI = Functional Assessment of Cancer Therapy Breast Symptom Index; OS = overall survival; pCR = pathological complete response; Q3W = every three weeks; QLQ-BR23 Breast Cancer-Specific Quality of Life Questionnaire; QLQ-C30 = Quality of Life Questionnaire; QW = once weekly; RCT = randomised controlled trial; TNBC = triple-negative breast cancer

ERG comments:

- Language restrictions: The ERG notes that an English language only restriction was applied to the clinical SLR search. The ERG considers excluding non-English language studies to be inappropriate for obtaining robust evidence on the treatment of adults with previously untreated locally advanced, non-metastatic TNBC as this does not follow-up best practice and potentially relevant studies might have been missed.
- Date restriction: Eligible articles were restricted to those published within 15 years of the SLR commencement. As the term, *"triple-negative breast cancer (TNBC)"* was first used in 2005, this date restriction appears to be appropriate for the SLR.[19]
- Study design restrictions: The study design restriction placed on eligible studies appears to only allow for randomised, controlled, prospective clinical trials above phase 1, open-label studies, and post-hoc analyses of patient sub-groups, to be included in the SLR. This would appear to be appropriate.

3.1.3 Critique of data extraction

Given that the company did not provide any information on the SLR data extraction process, the ERG asked the company to provide more information on how data extraction was conducted.[5] In response to the request for clarification, the company stated that *"two reviewers, working independently, extracted data (…) Following reconciliation between the two reviewers, a third reviewer was included to reach a consensus for any remaining discrepancies"*.[6] This response reassures the ERG that the methodology of data extraction was appropriate.

3.1.4 Quality assessment

The company conducted a quality assessment of the KEYNOTE-522 trial using the Cochrane Risk of Bias tool version 2 (ROB2) and determined the study to be of low risk of bias.[20] The quality of the KEYNOTE-522 trial has been further examined in Section 3.2.4 of this report.

ERG comment: In the request for clarification, the company was asked to provide further details on how the quality assessment process was carried out; in particular, how many reviewers were involved at each stage and how discrepancies in assessment results were resolved.[5] In the response to the request for clarification, the company stated that *"two reviewers, working independently, (...) performed the quality assessment. Following reconciliation between the two reviewers, a third reviewer was included to reach a consensus for any remaining discrepancies"*.[6] The company also provided a detailed breakdown for all ROB2 signalling questions for each paper. This response suggests that the methodology of quality evaluation was appropriate.

3.1.5 Evidence synthesis

The company considered the KEYNOTE-522 trial to be the only study identified by the clinical SLR to explore the effectiveness and safety of pembrolizumab as adjuvant therapy in adults with previously untreated locally advanced, non-metastatic TNBC, and thus did not consider a meta-analysis to be relevant to this submission.[1]

3.2 Critique of trials of the technology of interest, their analysis and interpretation (and any standard meta-analyses of these)

In the abstract/title screening phase of the CS SLR, 1383 records were excluded and 142 were retained for full text screening.[1] After an additional 4 articles were found by hand-searching, 12 final citations were included, and the other 134 articles were excluded. From the 12 final citations, 7 unique trials were identified. Of the seven identified trials, only KEYNOTE-522 reported on pembrolizumab as the intervention, as shown in Table 3.3. As such, KEYNOTE-522 was reported by the company to be the only study of relevance to this appraisal.

Table 3.3: Trials included/excluded in CS SLR

Trial	Treatment	Inclusion	CS Comments
ETNA; Gianni et al. 2018[21]	Paclitaxel *versus* Nabpaclitaxel	No	Intervention and comparators differ to KN-522
GeparSepto; Untch et al. 2016[22]	Nab-paclitaxel + epirubicin + cyclophosphamide *versus* Paclitaxel + epirubicin + cyclophosphamide	No	Intervention and comparators differ to KN-522
IMpassion031; Mittendorf et al. 2020[23]	Atezolizumab + nab-Paclitaxel *versus* Placebo + nab-Paclitaxel	No	Intervention and comparators differ or irrelevant to decision problem
KEYNOTE-522; Schmid et al. 2020[24]	Pembrolizumab + chemotherapy + anthracycline *versus* Placebo + chemotherapy + anthracycline	Yes	H2H comparison study directly informing the decision problem
NATT; Chen et al. 2013[25]	Docetaxel + cyclophosphamide + epirubicin *versus* Docetaxel + cyclophosphamide	No	Intervention and comparators differ to KN-522

Trial	Treatment	Inclusion	CS Comments
NCI 10013; Ademuyiwa et al. 2021[26]	Carboplatin + paclitaxel *versus* Atezolizumab + carboplatin + paclitaxel	No	Intervention and comparators differ or irrelevant to decision problem
Vriens et al. 2013[27]	Doxorubicin + cyclophosphamide + docetaxel *versus* Doxorubicin + cyclophosphamide + docetaxel	No	Intervention and comparators differ to KN-522
Adapted from Table 7 in CS Appendices[4] CS = company submission; SLR = systematic literature review			

ERG comment: Given the large number of 20 interventions included in the SLR, the total number of included trials appears to be quite low. Furthermore, the eligibility criteria for the SLR were vague. Therefore, the clarification letter posed four related questions to the company, as follows:[5]

1. Why was the I-Spy2 trial excluded when it involved standard neoadjuvant chemotherapy – 80 mg/m^2 IV paclitaxel, followed by doxorubicin plus IV cyclophosphamide - which is in line with the eligibility criteria?
   The company response was *"to facilitate an understanding of the relative treatment effect of interventions of interest, studies must have included at least two treatment arms of interest to be eligible for inclusion in the SLR of clinical evidence. Patients with TNBC enrolled in ISPY-2 were treated with paclitaxel with or without pembrolizumab followed by doxorubicin plus cyclophosphamide. As one of the treatment arms—pembrolizumab plus paclitaxel followed by doxorubicin plus cyclophosphamide—was not listed in the PICOS criteria, this trial was excluded from the SLR".*[6]
2. The company provided rationale for its decisions on the basis of study design and PICOS outlined and therefore the ERG is satisfied with this response.Why was the PROCEED Trial excluded from the SLR based on outcomes when it reported OS, progression-free survival (PFS), quality of life (QoL) and AEs?
   The company response was that *"the PROCEED trial (KCSG BR 11-01) enrolled patients with HER2-negative metastatic breast cancer. While subgroup results for patients with TNBC were reported for overall survival and progression-free survival in Park et al. 2019, these outcomes were not of interest to the SLR on HRQoL, and subgroup results for these patients were not reported for HRQoL measurements. Thus, this trial was excluded from the SLR of HRQoL studies".*[6]
   The ERG would respond that this approach represents an SLR protocol violation, because at no point in the protocol (Table 4 in appendix D) are eligible outcomes limited to HRQoL. Therefore, the ERG is not clear on why the paper was excluded.
3. What were the 'other' reasons for which 30 studies were excluded?
   The company response was that *"the PRISMA diagram has been updated and excluded publications table of the SLR of clinical evidence to include specific reasons for exclusion with 'Other.' Fourteen citations were excluded because full-text publications superseded them, 13 citations were excluded because they were study protocols, one citation was excluded as a duplicate, one citation was excluded because the full-text was unavailable, and one citation was excluded because it was a pooled analysis and not of interest to the SLR of clinical evidence".*[6] The ERG is satisfied with this response.
4. Why were several phase III trials excluded based on 'inappropriate study design', when phase III studies are listed as eligible in Table 4 of the CS appendices?[4]

The response to clarification was that *"additional notes are provided in Table 22 (appendix), for those references excluded due to 'study design' reasons such as non-randomized study design or prognostic/predictive/genomic/correlative study design"*.[6] The ERG is satisfied with this response.

3.2.1 Details of the included trial: the KEYNOTE-522 trial

The CS identified the KEYNOTE-522 trial as the only RCT evaluating pembrolizumab for TNBC.[1] The publications related to this trial that are cited in the CS are Schmid et al. 2020,[24] the CSR,[28] and a report of the meeting of the virtual advisory board[29].

The following information is taken from the CS.[1] The trial contains 1,174 participants, 1,173 of which are female. The mean (standard deviation, SD) age is 49.1 years (11.8) with a range of 22 to 80 range. Three quarter (75%) of participants were at stage II disease, whilst 24.9% were at stage III disease. Participants were required to be 18 years or over, with newly diagnosed TNBC of either T1c N1-2 or T2-4 N0-2, an ECOG PS 0-1 and a tissue sample for PD-L1 assessment.

Participants were randomly allocated to a treatment or placebo comparator arm, using a 2:1 ratio with stratification for nodal status, tumour size and carboplatin schedule. Participants randomised to the treatment arm (n=784) were administered pembrolizumab in combination with standard neoadjuvant chemotherapy followed by adjuvant pembrolizumab. Participants randomised to the placebo arm (n=390) were administered placebo in combination with standard neoadjuvant chemotherapy followed by placebo in the adjuvant phase. The comparator treatment was designed to reflect current practice in the UK, where no active adjuvant treatment is given, see Section 2.3.

The neoadjuvant phase lasted 24 weeks and the adjuvant phase 27 weeks, with each cycle of treatment lasting 3 weeks. Therefore, the neoadjuvant phase contained 8 treatment cycles and the adjuvant phase contained 9 treatment cycles. Table 3.4 provides extra details of the drugs used in the respective phases.

To date, outcome data have been collected at four IA points, and the IA used for the CS submission appears to be the most recent one (IA4). Median duration of follow up at IA4 is 37.8 months (range 2.7 to 48 months). Although 291 participants have discontinued treatment in the intervention arm and 106 have discontinued treatment in the placebo arm, an intention-to-treat (ITT) approach has been used and follow-up data are currently available for all participants until IA4.

Data have been collected for five patient-relevant outcomes: pCR, EFS, OS, HRQoL and AEs. Attempts to achieve allocation concealment were made by use of an interactive voice response system. Performance and detection bias were minimised by blinding of all study personnel and patients for the duration of the study. A summary of the study methodology from KEYNOTE-522 is presented in Table 3.4.

Table 3.4: Study methodology for KEYNOTE-522

Study	KEYNOTE-522
Study design	Phase III stratified double-blind randomised controlled trial
Location	The study was conducted at 177 centres in 21 countries. There were 54 sites within Europe and of these, six where in the United Kingdom. A total of 434 patients were enrolled in Europe of which 40 were from the UK. All treatments were administered in secondary care setting on an outpatient basis.
Population	Patients with untreated newly diagnosed, locally advanced, centrally confirmed TNBC and have an Eastern Cooperative Oncology Group (ECOG) performance status of 0 or 1. **Inclusion:** Male and female subjects aged 18 and older who: • Have centrally confirmed TNBC, as defined by the most recent ASCO/CAP guidelines. • Have previously untreated locally advanced non-metastatic (M0) TNBC defined as: - T1c, N1-N2 - T2, N0-N2 - T3, N0-N2 - T4a-d, N0-N2 (These TNM statuses partly equate to stage 2A, 2B and 3A) • Provide a core needle biopsy consisting of at least 2 separate tumour cores from the primary tumour at screening to the central laboratory. • Have ECOG performance status of 0 or 1 performed within 10 days of treatment initiation. • Demonstrate adequate organ function within 10 days of treatment initiation. • Have left ventricular ejection fraction of ≥50% or ≥ institution lower limit of normal (LLN). • Males and female subjects of childbearing potential must be willing to use an adequate method of contraception. **Exclusion:** Subjects were excluded from participating if they had: • history of invasive malignancy ≤5 years prior to signing informed consent except for adequately treated basal cell or squamous cell skin cancer or in situ cervical cancer. • received prior chemotherapy, targeted therapy, and radiation therapy within the past 12 months. • received prior therapy with an anti-PD-1, anti-PD-L1, or anti-PD-L2 agent or with an agent directed to another co-inhibitory T-cell receptor or has previously participated in MK-3475 clinical trials.

Study	KEYNOTE-522
	• participated in an interventional clinical trial with an investigational compound or device within 4 weeks of the first dose of treatment in this current trial. • received a live vaccine within 30 days of the first dose of study treatment. • an active autoimmune disease that has required systemic treatment in past 2 years. Replacement therapy is not considered a form of systemic treatment. • diagnosis of immunodeficiency or is receiving systemic steroid therapy or any other form of immunosuppressive therapy within 7 days prior to the first dose of trial treatment. • known history of Human Immunodeficiency Virus (HIV), or known active Hepatitis B or Hepatitis C. • history of (non-infectious) pneumonitis that required steroids or current pneumonitis. • active infection requiring systemic therapy. • significant cardiovascular disease, such as: myocardial infarction, acute coronary syndrome or coronary angioplasty/stenting/bypass grafting within the last 6 months; Congestive heart failure (CHF) New York Heart Association (NYHA) Class II-IV or history of CHF NYHA class III or IV • history or current evidence of any condition, therapy, lab abnormality or other circumstance that might expose the subject to risk by participating in the trial, confound the results of the trial, or interfere with the subject's participation for the full duration of the trial. • known psychiatric or substance abuse disorders • Were pregnant or breastfeeding, or expecting to conceive children within the projected duration of the trial • known hypersensitivity to the components of the study therapy or its analogues. • known history of active TB (Bacillus Tuberculosis)
Intervention(s)	**Neo-adjuvant phase** Pembrolizumab 200 mg IV on day 1 of each 21-day cycle (Q3W) for 8 cycles plus Cycles 1- 4: Carboplatin AUC 5-day Q3W (or AUC 1.5 weekly) + paclitaxel 80mg/m^2 QW Cycles 5 to 8: Doxorubicin 60 mg/m^2 Q3W or epirubicin 90 mg/m^2 and cyclophosphamide 600 mg/m^2 Q3W **Adjuvant phase** Pembrolizumab 200 mg Q3W for 9 cycles. Total pembrolizumab cycles across neoadjuvant + adjuvant phase = 17 Q3W infusions.
Comparator(s)	**Neo-adjuvant phase** Placebo (normal saline or dextrose) IV on day 1 of each 21-day cycle (Q3W) for 8 cycles plus

Study	KEYNOTE-522
	Cycles 1- 4: Carboplatin AUC 5-day Q3W (or AUC 1.5 weekly) + paclitaxel 80mg/m^2 QW Cycles 5 to 8: Doxorubicin 60mg/m^2 Q3W or epirubicin 90mg/m^2 and cyclophosphamide 600mg/m^2 Q3W **Adjuvant phase** Placebo (normal saline or dextrose) Q3W for 9 cycles.
Additional treatments	All treatments that the investigator considered necessary for a subject's welfare could be administered at the discretion of the investigator in keeping with the community standards of medical care. Subjects were prohibited from receiving the following therapies from the time of screening until completion of all study therapy: • Immunotherapy not specified in the protocol • Chemotherapy not specified in the protocol • Investigational agents not specified in the protocol • Radiation therapy except as described in the protocol. (Post-operative radiation therapy is acceptable according to the standard of care, as applicable). • Live vaccines within 30 days prior to the first dose of trial treatment and while participating in the trial. • Glucocorticoids for any purpose other than to modulate symptoms from an AE of suspected immunologic aetiology or for use as a pre-medication for chemotherapeutic agents specified in the protocol. Inhaled steroids were allowed for management of asthma. Use of prophylactic corticosteroids to avoid allergic reactions were permitted.
Reported outcomes specified in the decision problem	• Pathological complete response (pCR) • Event-free survival (EFS) • Adverse events (AEs) • Overall survival (OS) • Health-related quality of life (HRQoL)
All other reported outcomes	• Patient reported outcomes (PRO) • Time on treatment

Based on the CS and meeting report of the Virtual Advisory Board[1, 29]
AEs = adverse events; AUC = area under the curve; CS = company submission; EFS = event-free survival; HRQoL = health-related quality of life; OS = overall survival; pCR = pathological complete response; PRO = patient reported outcomes; Q3W = every three weeks, QW = once weekly

ERG comment: Comments below have been separated into sections on duration of follow up, concomitant medications, protocol deviations and external validity. Comments relating to study methodology are covered in Section 3.2.4.

3.2.1.1 Duration of follow up

The short duration of follow-up of only 3 years precludes the assessment of mature survival data and the long-term safety profile. The ERG requested clarification of the reasons for this, along with a discussion of these limitations and the consequences for clinical decision making.[5]

In the response to the request for clarification, the company stated that *"at IA4 with median follow up at IA4 was over three years (39.1 months), the EFS HR of 0.63 (95% CI: 0.48, 0.82), with a one-sided p-value of 0.0003093 that crossed the prespecified boundary for statistical significance (0.00516941), represents a 37% reduction in the risk of disease progression precluding definitive surgery, recurrence, second primary malignancy, or death compared with placebo + NAC / placebo. The information fraction of EFS was approximately 66% [216 of the 327 events needed for the final analysis]. As noted, EFS is an endpoint listed on the FDA surrogate table for breast cancer. By the time of the IA4 Last Patient Last Visit (LPLV) there had been one year since the last exposure which occurred on 11th February 2020. Clinical experts advised MSD the pCR and EFS outcomes from KEYNOTE-522 were good and acknowledged they hoped to use the pembrolizumab combination in the future based upon the trial results. They also suggested that OS events are driven by a reduction in distant recurrences, which equates to a survival benefit in the TNBC setting based on the reduction in distant recurrences observed to date with pembrolizumab in KEYNOTE-522 and therefore, an OS benefit is expected in future analyses"*.

3.2.1.2 Concomitant medications

As shown in Table 3.4, *"all treatments that the investigator considered necessary for a subject's welfare could be administered at the discretion of the investigator in keeping with the community standards of medical care"*. The ERG asked the company to supply a table of the most frequently used concomitant medications during the KEYNOTE-522 trial and to discuss if non-protocol specified concomitant medications were used during the trial.[5]

In response to the request for clarification, the company stated that *"supportive care for the chemotherapeutic agents administered in KEYNOTE-522 could be found in the local product label for each agent. Corticosteroids (such as prednisone), insulin replacement therapy, hormonal replacements, beta blockers, thyroid replacement hormones and other medications were included in the toxicity management guidelines of immune related adverse events. As detailed in B.2.3 of the company submission the protocol specific prohibited concomitant medications. Glucocorticoids were administered to some patients, but in line with the protocol to manage immune-related adverse events, as a pre-medication for chemotherapy or for the management of asthma. Also, a proportion of patients received a vaccine (6.1%), most of which were inactivated though a small number, 3.3%, received an unspecified influenza vaccine"*.[6]

The company directed the ERG to Table 23 in the CS Appendix which was supposed to summarise the most frequently used concomitant medications used during the trial, but this table could not be found (Table 23 in the appendices detailed adverse events). Overall, however, the company responses suggested that concomitant medications were not likely to be a source of significant bias.

3.2.1.3 Protocol deviations

Section 10.2 of the KEYNOTE-522 CSR alluded to [redacted].[2] The ERG asked the company about how 'important', and 'not important' protocol deviations were classified.[5] In response to the request for clarification, the company stated that, *"protocol deviations were classified as '"important or 'not important' by a standard method assessing the potential impact of the protocol deviation on endpoints and safety"*.[6] This statement was supported by any references and more information is required.

The ERG also asked the company to discuss how COVID-19 may have affected the KEYNOTE-522 trial.[5] In response, the company stated that *"part of KEYNOTE-522 was conducted during the COVID-19 pandemic. MSD continued to follow its Standard Operating Procedures (SOPs) for study conduct, monitoring, and oversight during the pandemic. Exceptions and deviations from SOPs were documented. Study sites were advised to follow local and national guidance regarding the pandemic and to share any mitigation plans for study participant management with the Institutional Review Board/Ethics Review Committee and the sponsor. Study sites were also advised to remain in contact with study participants to monitor for safety concerns and to keep participants informed of changes to the study and other study activities. There were no changes in the planned analyses of the study due to the COVID-19 pandemic"*.[6] The ERG is satisfied that appropriate steps were taken to cater for the pandemic and that it is unlikely than the pandemic has had a negative effect on data quality.

The ERG noted that only a proportion of randomised patients in the KEYNOTE-522 trial proceeded onto receiving adjuvant therapy. The ERG requested more information on why this took place and asked for a comparison between the proportion of patients who received surgery/adjuvant therapy in the trial and the proportion of patients that would receive it in NHS clinical practice, along with a discussion of the implications of any difference.[5] In its response to the request for clarification, the company stated that *"about 98% of patients in both treatment arms underwent surgery; therefore, performance of surgery did not differentially impact start of adjuvant therapy. The primary reason for which randomized patients in either treatment arm did not proceed to adjuvant therapy was discontinuation due to adverse events… The higher incidence of discontinuation in the neoadjuvant phase in the pembrolizumab + NAC group was driven primarily by a higher discontinuation rate due to adverse events (14.3%) compared with the placebo + NAC group. Per protocol, if a participant discontinued either pembrolizumab or placebo during the neoadjuvant phase due to toxicity related to pembrolizumab/placebo, the participant was not permitted to receive it in the adjuvant phase of the study. For all other reasons for discontinuation, proportions were similar between groups"*.[6]

Table 3.5 summarises the reasons for drop-out. The company reiterated the important point that the analyses were intention-to-treat (ITT) and that *"despite fewer participants starting adjuvant treatment, KEYNOTE-522 demonstrated that the complete regimen of pembrolizumab + NAC followed by pembrolizumab monotherapy after surgery in the adjuvant phase resulted in a statistically significant and clinically meaningful improvement in both pCR and EFS in the ITT population"*.[6] This explanation reduced ERG concern about the numbers not proceeding to adjuvant therapy. The company was unable to find relevant NHS clinical data to compare the number dropping out of therapy with clinical practice, although it reported data on patients from Scotland who had been given adjuvant chemotherapy had a 20% drop-out rate, similar to that seen with pembrolizumab. However, it correctly cautioned that such data were not directly applicable because *"it included patients with all subtypes of breast cancer, while it did not include those who had neoadjuvant therapy and did not include English hospitals"*.[6] The ERG would also add that the data were from patients where adjuvant chemotherapy had been prescribed for

all, which was completely contrary to the case in this trial, where none were given adjuvant chemotherapy.

Table 3.5: Reasons for discontinuation from all treatments for participants who did not start adjuvant phase - All participants (ITT Population)

	Pembrolizumab + NAC/ Pembrolizumab	%	Placebo + NAC/ Placebo	%
Participants randomised	784		390	
Untreated participants	1	0.1	1	0.3
Treated participants	783	99.9	389	99.7
Participants who started adjuvant phase	588	75.0	331	84.9
Participants who did not start adjuvant phase	195	24.9	58	14.9
Discontinued in neoadjuvant phase	190	24.2	58	14.9
Adverse events	112	14.3	20	5.1
Clinical progression[a]	2	0.3	3	0.8
Physician decision	32	4.1	15	3.8
Progressive disease	8	1.0	7	1.8
Relapse/recurrence	7	0.9	3	0.8
Withdrawal by subject	29	3.7	10	2.6
Had surgery, but did not receive study medication	5	0.6	0	0.0
Still on treatment in neoadjuvant phase	0	0.0	0	0.0
Participants with surgery	768	98.0	381	97.7

Based on Table 5 in the response to the request for clarification[6]
Participants who did receive study medication but had surgery were included in subjects treated.

[a] Clinical progression is disease progression determined by the Investigator. "Progressive disease" is disease progression determined by imaging using RECIST 1.1 criteria.
Database cut-off date: 23 March 2021
ITT = intention-to-treat; NAC = neoadjuvant chemotherapy

In relation to the above point, more than double the number of patients on the pembrolizumab arm (compared to the comparator arm) discontinued study treatment in both the neoadjuvant phase and adjuvant phase of the KEYNOTE-522 trial. This was raised in the clarification letter, and the ERG requested that the company 1) detail and discuss study discontinuation due to adverse effects (AEs), 2) discuss the criteria used to characterise a "clinically important protocol deviation", 3) clarify if the greater number of protocol deviations with study intervention observed on the pembrolizumab arm was due to AEs, and 4) clarify if cross treatment was introduced in the KEYNOTE-522 trial as a protocol deviation.[5]

In its response to clarification, the company clarified that the discontinuation was largely due to adverse events, as detailed in the paragraph above.[6] The company defined clinically important protocol deviations as: *"those that may compromise critical data analyses, especially those pertaining to (1) primary efficacy and/or primary safety endpoints, or (2) the participant's safety"*. The company also confirmed that the protocol deviations were not related to AEs. Finally, in relation to the question about cross-treatments, the company responded with *"universal unblinding upon disease*

progression/recurrence and cross treatment on … was not allowed per protocol; however, off-study treatment with an immune-oncology agent after discontinuation of study treatment due to disease progression/recurrence was at physician's discretion. If this occurred, it was not considered to be a clinically important protocol deviation".[6]

3.2.1.4 External validity of KEYNOTE-522 trial

The ERG noted that the trial inclusion criteria specified that patients would have an ECOG PS of 0 or 1 performed within 10 days of treatment initiation. Thus, the ERG asked the company to confirm if patients in UK clinical practice with an ECOG PS ≥2 would not be expected to receive pembrolizumab as adjuvant therapy.[5] In its response to the request for clarification, the company stated that *"in previous approvals of immunotherapies in oncology a criterion is included on Blueteq forms for only patients who have an ECOG PS of 0 or 1, for example PEMB1 on the baseline funded drugs list".*[6]

The CS states that the KEYNOTE-522 trial recruited 40 participants from six UK study sites and further clarification has been requested on the exact geographical regions used and the specific effect sizes from Europe and UK.[1] This was regarded as particularly important because subgroup analysis results (see Section 3.2.5.5) indicate some potential differences between geographical regions, suggesting that overall findings in the KEYNOTE-522 study may not necessarily be applicable to a single region, and may therefore not be directly applicable to the UK. In the clarification letter, the ERG asked for all results to be sub-grouped for 1) Europe versus rest of world and 2) UK versus rest of world. The company provided EFS data showing that the Europe sub-group had a less favourable relative effect size for pembrolizumab (HR 0.73, 95% CI 0.49 to 1.08) compared to the rest of the world sub-group (HR 0.55, 95% CI 0.38 to 0.80), suggesting that the overall data might be providing an overly optimistic picture for European patients. The company did not provide similar data for a UK patient sub-group, and effectively did not respond to the direct question. Table 3.6 summarises the situation.

Table 3.6: EFS Subgroup analysis

	Effect size of pembrolizumab versus placebo for EFS
Whole cohort	HR: 0.63 (95% CI 0.48 to 0.82)
Europe versus rest of world	Europe: HR [redacted] Rest of the world: HR [redacted]
UK versus rest of world	UK: Data not provided Rest of the world: Data not provided
Based on Table 2 in the response to the request for clarification[6] and Table 13 of the CS[1] CS = company submission; CI = confidence interval; EFS = event-free survival; HR = hazard ratio; UK = United Kingdom	

The company were also asked to provide the baseline characteristics of the 40 UK patients by study arm in comparison with the overall trial ITT population's baseline characteristics. The company provided a table of the baseline characteristics of the 40 UK patients per arm. Comparison of these characteristics to the overall ITT population characteristics published in the CS showed some differences for some characteristics. Notable differences were ethnicity, with the UK data having a higher proportion of white participants (85% in UK data versus 63.5% in overall data), a higher proportion of people with ECOG PS 0 (95% versus 87%), and a greater choice of carboplatin Q3W (67.5% versus 42.8%). Tumour size (70% T1 or T2 versus 74% T1 or T2) and stage (20% stage II versus 25% stage III) were also slightly different. As implied by the company, the small number of UK participants makes such simplistic comparisons prone to sampling error but do suggest uncertainty over the question of how representative the overall data are to the UK population. It is unclear if these

potential differences in characteristics between the UK participants and the overall trial participants would affect outcomes, but they do suggest, in tandem with the EFS sub-group results previously described for Europe versus the rest of the world, that it is possible that the overall results observed in the KEYNOTE-522 trial may not necessarily be relevant to UK patients. This has been identified as key issue 3.

On being asked to discuss the generalisability of the study baseline demographic and disease characteristics to the clinical practice population in England and Wales, the company response was that *"while there is little published data on the demographics of UK patients with early stage triple negative breast cancer, we have not identified any characteristics of subjects in the trial that are not generalisable to patients in the UK. A study on patients in the Northeast London Cancer Network with TNBC (any stage) between 2005 and 2007, reported 82.8% were 69 years and under.[30] The proportion of patients under the age of 65 in KEYNOTE-522 was slightly higher, 88.8%, but this is to be expected as the trial recruited only patients with early-stage non-metastatic disease. Jack et al (2013) reported just over one in five patients were within the Black ethnicity group, which is in line with the UK KEYNOTE-522 participants. Stage at diagnosis for breast cancer data, published by the National Disease Registration Service (NDRS), is reported for all subtypes combined in England.[31] Of the 19,633 patients diagnosed with stage II and III breast cancers, 81.4% were the former, which is in line with KEYNOTE-522 ITT population and UK, 75.0%. and 80.0%, respectively. No major differences are noted between the key baseline demographic and disease characteristics in the UK versus KEYNOTE-522 ITT population, therefore we consider that the trial population is generalisable to that of UK patients."*.[6] These data appear to show that the trial sample is unlikely to spuriously favour the intervention, as might occur if the sample contained a higher proportion of people with a better prognosis than the UK patient population. However, this does not change the conclusion reached in the previous section that UK patients may not have the same reactions as patients in the rest of the world.

The ERG also asked the company to discuss the representativeness of the control arm to England and Wales and to discuss if the trial comparator is consistent with clinical practice. The company responded by stating that *"clinical experts have informed MSD the treatments used in KEYNOTE-522 reflects the current standard of care for neoadjuvant and adjuvant treatment of TNBC where both phases are used. The NICE guidelines for early and locally advanced breast cancer (NG101) recommend "people with triple-negative invasive breast cancer, consider a neoadjuvant chemotherapy regimen that contains both a platinum and an anthracycline". Local NHS cancer guidelines list carboplatin + paclitaxel followed by doxorubicin/epirubicin plus cyclophosamide (or order of chemotherapies is switched) for neoadjuvant treatment of TNBC patients"*.[6] The ERG accepts these points, with the caveat (as has been discussed) that capecitabine should have been considered as part of adjuvant therapy.

The ERG also pointed out that results by BC gene (BRCA1) mutation are missing and requested clarification whether patients would be offered pembrolizumab regardless of the BRCA mutation.[5] The company's response stated that "*determination of BRCA status was not required for KEYNOTE-522. Of the 54 (4.6%) participants with a BRCA1/2 mutation detected, 40 participants were in the pembrolizumab + NAC / pembrolizumab group and 14 participants were in the placebo + NAC / placebo group (as a reminder, randomisation ratio was 2:1). The number of participants with known BRCA status is too small to provide a meaningful assessment for pCR, EFS, or OS. Patients received pembrolizumab regardless of BRCA mutation results in KEYNOTE-522."*.[6]This response suggests that there is reasonable random mixing of this characteristic across arms (expected numbers would be 36 and 18 in the pembrolizumab and control arms as opposed to the observed 40 and 14) and the small imbalance is very unlikely to confound results. It is also clear that numbers are too small to allow any reasonable sub-group analyses.

Other comments relevant to this section have already been made in Section 2.3.

3.2.2 Statistical analyses of the KEYNOTE-522 trial

The statistical analyses used for the primary endpoint, alongside the sample size calculations and methods for handling missing data are presented in Table 3.7.

Table 3.7: Summary of statistical analyses for the primary analysis in KEYNOTE-522

Treatment assignment	Approximately 1,150 subjects will be randomised (double-blind) in a 2:1 ratio between 2 treatment arms: 1. Pembrolizumab plus chemotherapy as neoadjuvant therapy and pembrolizumab as adjuvant therapy, or 2. Placebo plus chemotherapy as neoadjuvant therapy and placebo as adjuvant therapy. Stratification factors are as follows: Nodal status (Positive versus Negative) Tumour size (T1/T2 versus T3/T4) Choice of Carboplatin: Q3W versus Weekly
Analysis populations	Efficacy: Intention-to-Treat Population[32] Safety: All Subjects as Treated (ASaT)
Primary endpoints	1. Pathological complete response (pCR) rate (ypT0/Tis ypN0) 2. Event-free survival (EFS)
Statistical methods for key efficacy analyses	Treatment comparisons of the pCR rate (ypT0/Tis ypN0) will be performed using the stratified Miettinen and Nurminen method. Treatment comparisons for time-to-event endpoints such as EFS and overall survival (OS) will be evaluated using a stratified log-rank test. The hazard ratio (HR) will be estimated using a stratified Cox model.
Statistical methods for key safety analyses	The analysis of safety will follow a tiered approach. There are no Tier 1 events for this study. Point estimates and 95% confidence intervals[33] for between-treatment comparisons via the Miettinen and Nurminen method will be provided for Tier 2 safety endpoints; only point estimates by treatment group will be provided for Tier 3 safety endpoints.
Interim and final analyses	Seven efficacy interim analyses (IAs) are planned. Results will be reviewed by an external DMC. By final analysis (FA) approximately 327 EFS events are expected to have been observed (event driven). It is expected to occur at ~102 months after the first subject is randomised. Primary purpose: final EFS analysis. OS will be tested only when the null hypothesis for EFS is rejected.
Multiplicity	The overall type-I error rate over the 2 primary endpoints will be strongly controlled at 2.5% (one-sided) with 0.5% allocated to the pCR (ypT0/Tis ypN0) and 2.0% allocated to the EFS hypotheses. The graphical approach of Maurer and Bretz will be applied to re-allocate alpha among hypotheses for pCR (ypT0/Tis ypN0), EFS, and OS in subjects with locally advanced TNBC. Group sequential methods will be used to allocate alpha between the interim and final analyses for pCR(ypT0/Tis ypN0), EFS and OS in subjects with locally advanced TNBC.
Sample size and power	The FA of the study is EFS event-driven and will be conducted after approximately 327 EFS events have been observed. It may occur at ~102

	months after the first subject randomized. The planned sample size is approximately 1150 subjects 1. pCR (ypT0/Tis ypN0): the trial has an overall ~95% power to detect a true pCR rate difference of 15 percentage points (pembrolizumab + chemotherapy versus placebo + chemotherapy) at alpha = 0.5% (one-sided) with ~1,000 subjects who have or would have completed surgery after ~6 months neoadjuvant treatment at IA2. 2. EFS: the trial has an overall ~80% power at a one-sided 2.0% alpha level if the true HR is 0.71. 3. OS: the trial has an overall ~79.7% power at a one-sided 2.0% alpha level, if the true HR is 0.70
Based on Table B.2.4 of the CS[1] ASaT = all subjects as treated; CS = company submission; EFS =- event-free survival; HR = hazard ratio; IA = interim analysis; OS = overall survival; pCR = pathological complete response; Q3W = every three weeks; TNBC = triple-negative breast cancer	

ERG comment: Statistical approach appears to be rigorous and correct.

3.2.3 Baseline characteristics of the KEYNOTE-522trial

A total of 1,174 participants were allocated randomly to the two arms in a 2:1 ratio. A summary of the baseline characteristics of patients is presented in Table 3.8.

Table 3.8: Baseline characteristics of patients in the ITT population of KEYNOTE-522

	Pembrolizumab + chemotherapy / Pembrolizumab (n=784)		Placebo + chemotherapy / Placebo (n=390)		Total (n=1,174)	
	n	(%)	n	(%)	n	(%)
Sex						
Male	1	(0.1)	0	(0.0)	1	(0.1)
Female	783	(99.9)	390	(100.0)	1,173	(99.9)
Age (Years)						
< 65	700	(89.3)	342	(87.7)	1,042	(88.8)
≥65	84	(10.7)	48	(12.3)	132	(11.2)
Mean	49.2		49.1		49.1	
SD	11.8		11.9		11.8	
Median	49.0		48.0		49.0	
Range	22 to 80		24 to 79		22 to 80	
Race						
American Indian or AlaskaNative	14	(1.8)	7	(1.8)	21	(1.8)
Asian	149	(19.0)	89	(22.8)	238	(20.3)
Black or African American	38	(4.8)	15	(3.8)	53	(4.5)
Multiple	13	(1.7)	6	(1.5)	19	(1.6)
American Indian or AlaskaNative Black or African American	0	(0.0)	1	(0.3)	1	(0.1)
American Indian or AlaskaNative Black or African American White	2	(0.3)	1	(0.3)	3	(0.3)
American Indian or AlaskaNative White	7	(0.9)	2	(0.5)	9	(0.8)

	Pembrolizumab + chemotherapy / Pembrolizumab (n=784)		Placebo + chemotherapy / Placebo (n=390)		Total (n=1,174)	
	n	(%)	n	(%)	n	(%)
Black Or African AmericanWhite	3	(0.4)	2	(0.5)	5	(0.4)
White Asian	1	(0.1)	0	(0.0)	1	(0.1)
Native Hawaiian or OtherPacific Islander	1	(0.1)	0	(0.0)	1	(0.1)
White	504	(64.3)	242	(62.1)	746	(63.5)
Missing	65	(8.3)	31	(7.9)	96	(8.2)
Geographic Region						
North America	166	(21.2)	78	(20.0)	244	(20.8)
Europe	388	(49.5)	180	(46.2)	568	(48.4)
Australia	23	(2.9)	16	(4.1)	39	(3.3)
Asia	166	(21.2)	91	(23.3)	257	(21.9)
Rest of World	41	(5.2)	25	(6.4)	66	(5.6)
ECOG PS						
0	678	(86.5)	341	(87.4)	1,019	(86.8)
1	106	(13.5)	49	(12.6)	155	(13.2)
Baseline Lactate Dehydrogenase (LDH)						
≤ULN	631	(80.5)	309	(79.2)	940	(80.1)
> ULN	149	(19.0)	80	(20.5)	229	(19.5)
Missing	4	(0.5)	1	(0.3)	5	(0.4)
Menopausal Status						
Pre-menopausal	438	(55.9)	221	(56.7)	659	(56.1)
Post-menopausal	345	(44.0)	169	(43.3)	514	(43.8)
Missing	1	(0.1)	0	(0.0)	1	(0.1)
Choice of Carboplatin (Planned)						
Carboplatin (Cb) Q3W	335	(42.7)	167	(42.8)	502	(42.8)
Carboplatin (Cb) Weekly	449	(57.3)	223	(57.2)	672	(57.2)
Primary Tumour (Planned)						
Tumour Size T1/T2	580	(74.0)	290	(74.4)	870	(74.1)
Tumour Size T3/T4	204	(26.0)	100	(25.6)	304	(25.9)
Nodal Involvement (Planned)						
Nodal Status Positive	405	(51.7)	200	(51.3)	605	(51.5)
Nodal Status Negative	379	(48.3)	190	(48.7)	569	(48.5)
Metastases						
M0	784	(100.0)	390	(100.0)	1,174	(100.0)
Overall Stage						
Stage I	0	(0.0)	1	(0.3)	1	(0.1)
Stage II	590	(75.3)	291	(74.6)	881	(75.0)
Stage III	194	(24.7)	98	(25.1)	292	(24.9)

	Pembrolizumab + chemotherapy / Pembrolizumab (n=784)		Placebo + chemotherapy / Placebo (n=390)		Total (n=1,174)	
	n	(%)	n	(%)	n	(%)
PD-L1 CPS 1 Cut-off						
PD-L1 CPS ≥ 1	656	(83.7)	317	(81.3)	973	(82.9)
PD-L1 CPS < 1	128	(16.3)	69	(17.7)	197	(16.8)
Unknown	0	(0.0)	4	(1.0)	4	(0.3)
PD-L1 CPS 10 Cut-off						
PD-L1 CPS ≥10	[redacted]					
PD-L1 CPS < 10	[redacted]					
Unknown	[redacted]					
PD-L1 CPS 20 Cut-off						
PD-L1 CPS ≥20	[redacted]					
PD-L1 CPS < 20	[redacted]					
Unknown	[redacted]					
HER2 Status						
0-1+ by IHC	595	(75.9)	286	(73.3)	881	(75.0)
2+ by IHC (but FISH-)	188	(24.0)	104	(26.7)	292	(24.9)
Missing	1	(0.1)	0	(0.0)	1	(0.1)

Based on Table 5 of the CS[1]
Missing values in Race and Ethnicity are mainly because France is not permitted to report this information.
The missing value in Menopausal Status is from one male participant.
The missing value in HER2 Status is from the participant with missing IHC, but FISH-.Database Cut-off Date: 23MAR2021
CPS = combined positive score; CS = company submission; ECOG = Eastern Cooperative Oncology Group; FISH = fluorescence in situ hybridization HER2 = human epidermal growth factor receptor 2; IHC = immunohistochemistry; ITT = intention-to-treat; LDH = lactate dehydrogenase; PS = performance status; Q3W = every three weeks; ULN = upper limit of normal

ERG comment: The characteristics listed demonstrate reasonable levels of comparability between arms. Given the law of large numbers and the fact that this was a randomised trial, it can be assumed that other characteristics which were not measured would be similarly distributed. The CS provides details of the numbers of participants in KEYNOTE-522 with stage I, II and stage III disease, but not the four detailed TNM gradings mentioned in the inclusion criteria (page 19 of the CS): T1c, N1-N2; T2, N0-N2; T3, N0-N2; and T4a-d, N0-N2.[1] It is likely that stage relates to prognosis, and so it is vital to know if the ratio of stages in the trial is equivalent to ratios of stages in the UK population. The company has been asked in the clarification letter to provide more details on the numbers with TNM stages T1c, N1-N2; T2, N0-N2; T3, N0-N2; and T4a-d, N0-N2. The company provided the following Table 3.9 that highlights the numbers in each stage.

Table 3.9: Additional participant characteristics (ITT)

	Pembrolizumab + chemotherapy / Pembrolizumab	Placebo + chemotherapy / Placebo	Total
Participants in population (N)	784	390	1,174

	Pembrolizumab + chemotherapy / Pembrolizumab		Placebo + chemotherapy / Placebo		Total	
Tumour Stage and Nodal Involvement Grading						
	n	**(%)**	**n**	**(%)**	**n**	**(%)**
T1b, N1	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
T1c, N1-N2	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
T1c, N3	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
T2, N0-N2	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
T2, N3	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
T3, N0-N2	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
T4, N0-N2	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
T4a-d, N0-N2	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Based on Table 7 of the response to request for clarification[6]
Database Cut-off Date: 23MAR2021
The one patient with Stage I disease was considered a protocol deviation, as the inclusion criteria only allowed enrolment of patients with Stage II or III disease
ITT = intention-to-treat

The company has also been asked to provide tumour, node, and metastasis (TNM) grading data on the UK population of patients with TNBC, to allow evaluation of whether the proportions of participants at different stages in the trial are similar to those in the UK population.[5] The response was that *"data for TNM grading for TNBC patients is not available from publicly available data. Information published by the cancer registry is reported as stage 1, 2, 3 and 4."*.[6] This is highlighted as key issue 4.

3.2.4 Risk of bias assessment of the KEYNOTE-522 trial

A quality assessment of the KEYNOTE-522 trial was provided in the CS[1] using the Cochrane risk-of-bias tool for randomised trials (ROB-2),[34] the results of which are presented in Table 3.10. These demonstrate low risk of bias across all areas for both efficacy (EFS) and safety (AE) outcomes.

Table 3.10: Quality assessment of the KEYNOTE-522 against ROB-2 criteria

Area of potential bias	Risk of bias within the specified outcome	
	EFS	**AE**
Randomisation process	Low	Low
Deviations from the intended interventions	Low	Low
Missing outcome data	Low	Low
Measurement of the outcome	Low	Low
Selection of the reported result	Low	Low
Overall risk of bias	Low	Low

Based on Table 11 of the CS appendices[4]
AE = adverse event; CS = company submission; EFS = event-free survival

ERG comment: Neither document B of the CS nor appendices do not provide a rationale for the decisions made on the risk of bias rating.[1, 4] Furthermore, after review of the primary sources the ERG does not agree with the quality assessment in terms of the randomisation process, as detailed below.

The allocation concealment process is very briefly reported and although it is clear that treatment allocation occurred centrally using an interactive response technology system, insufficient information is given to be certain that those recruiting participants were unaware of the allocation sequence. The clarification letter requested further information, but the response did not provide any new information that had not previously been available in the CS: *"Treatment allocation/randomisation occurred centrally using an interactive voice response system / integrated web response system (IVRS/IWRS). Subjects were assigned randomly in a 2:1 ratio to pembrolizumab and placebo, respectively, after stratification. The choice of QW carboplatin or Q3W carboplatin should have been determined prior to randomisation, and carboplatin schedule was a stratification factor".*[6]

Section 10.1.2 of the KEYNOTE-522 CSR states that [redacted].[2] *The ERG requested clarification from the company about the potential effects of premature unblinding during the trial on outcomes measurement. The company response was that* [redacted]

Sponsor-approved non-emergency unblinding requests for participants who had disease progression / recurrence, knowing their study treatment would guide future treatment plans [redacted]

Inadvertent unblinding of investigator site and/or Sponsor personnel [redacted]

Emergency unblinding [redacted]

[redacted]

A summary of participants with or without an EFS event for participants with premature unblinding is provided in [Table 3.11]. [redacted] *out of* [redacted] *participants with premature unblinding already had an EFS event occurred on or prior to the date of unblinding, therefore, unblinding had no impact on the EFS data of those participants. The number of participants with premature unblinding either with an EFS event occurred after the date of unblinding, or without EFS event occurred is small* [redacted] *and generally consistent between the pembrolizumab + NAC / pembrolizumab group and the placebo + NAC / placebo group. There is no evidence to show the premature unblinding of participants without an EFS event at the time of unblinding had an impact on interpretation of the EFS results".*[6]

Table 3.11: Summary of participants with or without an EFS event. All participants with premature unblinding

	Pembrolizumab + chemotherapy/ Pembrolizumab		**Placebo + chemotherapy/ Placebo**		**Total**	
Participants in population (N)	784		390		1,174	
Scenarios						
	n	**(%)**	**n**	**(%)**	**n**	**(%)**
An EFS event occurred on or prior to the date of unblinding	[redacted]					
An EFS event occurred after the date of unblinding	[redacted]					

	Pembrolizumab + chemotherapy/ Pembrolizumab		Placebo + chemotherapy/ Placebo		Total	
No EFS event occurred	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Based on Table 9 of the response to the request for clarification[6] Database Cut-off Date: 23MAR2021 EFS = event-free survival						

It was also unclear if pathologists interpreting surgical specimens for the key outcome of pCR assessment were blinded, and the ERG requested clarification.[5] The company response was that *"all pathologists reviewing and interpreting surgical specimens for assessment of pCR were required to be blinded to treatment assignment"*.[6]

The revised ERG quality assessment, using the Cochrane ROB2 tool,[34] is presented in Table 3.12 for all three completed outcomes.

Table 3.12: ERG revised quality assessment of the KEYNOTE-522 against ROB-2 criteria

Area of potential bias	Risk of bias within the specified outcome		
	EFS	HRQoL	AE
Randomisation process	Unclear	Unclear	Unclear
Deviations from the intended interventions	Low	Low	Low
Missing outcome data	Low	Low	Low
Measurement of the outcome	Low	Low	Low
Selection of the reported result	Low	Low	Low
Overall risk of bias	**Unclear**	**Unclear**	**Unclear**
AE = adverse event; EFS = event-free survival; HRQoL = health-related quality of life			

3.2.5 Efficacy results of the KEYNOTE-522 trial

The final NICE scope lists the following outcomes that need to be covered in the TA:

- Pathological complete response (pCR)
- Event free survival (EFS)
- Adverse events (AEs)
- Overall survival (OS)
- Health related quality of life (HRQoL)

The first four of these outcomes will now be evaluated in turn. Adverse outcomes will be evaluated in Section 3.2.6.

3.2.5.1 Pathological Complete Response (pCR)

The definition for the primary pCR outcome is ypT0/Tis ypN0, meaning the absence of invasive cancer in the breast and axillary nodes. The pembrolizumab arm showed a greater magnitude of pCR events, with an absolute risk difference (95% CI) of 7.5% (1.6 to 13.4). See Table 3.13 below, and Appendix D.1.5 in the CS appendices for further information.[4]

Table 3.13: Analysis of pCR (ypT0/Tis ypN0) (All participants)

Treatment	N	Number of pCR	pCR Rate (%)	Difference in % versus placebo + chemotherapy Estimate (95% CI)[a]
Pembrolizumab + chemotherapy	784	[redacted]	[redacted]	[redacted]
Placebo + chemotherapy	390	[redacted]	[redacted]	

Based on Table 12 of the CS[1], table 12
[a] Based on Miettinen & Nurminen method stratified by nodal status (positive versus negative), tumour size (T1/T2 versus T3/T4) and choice of carboplatin (Cb) (Q3W versus Weekly).
CS = company submission; CI = confidence interval; pCR = pathological complete response; Q3W = every three weeks

ERG comment: The absolute risk difference between treatment arms for pCR (95% CI) of 7.5% (1.6 to 13.4) translates to a number needed to treat of 13.3, which would not normally be regarded as clinically significant.[35]

The CS states that definition for the primary pCR outcome is ypT0/Tis ypN0 (page 17).[1] On page 14 of the CS, this is defined as absence of invasive cancer in the breast and axillary nodes. However, it is also stated on the same page that other commonly used definitions of pCR are ypT0/Tis (absence of invasive cancer in the breast), and ypT0 ypN0 (absence of invasive and in situ cancer in the breast and axillary nodes). The company has been asked to clarify the definitions used, and its response is that *"the definition for the primary outcome of pCR is ypT0/Tis yp N0, meaning the absence of invasive cancer in the breast or all resected lymph nodes. Non-invasive breast residuals were allowed"*.[6]

The company were also asked to discuss why the definition indicative of more complete recovery (absence of invasive and in situ cancer in the breast and axillary nodes) was not used as the primary outcome pCR.[5] The response was that *"FDA guidance recognises ypT0/Tis ypN0 as an acceptable definition of pCR, and so that was selected as the definition used for pCR as the primary outcome. The alternative definition, ypT0 ypN0, was used as the definition for the secondary outcome analysis"*.[6] This confirms that the company used a less testing outcome as its primary outcome. Although potentially misleading, this is not actually a problem as the absolute risk difference (pembrolizumab – control arm) in pCR is actually more favourable to the intervention in the stricter definition: for ypT0ypN0 (the stricter definition) it is +7.6 (95% CI 1.6 to 13.6) and for ypT0/TisypN0 (the primary outcome used in the trial) it is +7.5(95% CI 1.6 to 13.4). Therefore, it could be argued that the company have slightly underestimated (rather than overestimated) its effect by using the ypT0/Tis ypN0 outcome as its primary variable.

3.2.5.2 Event-free survival (EFS)

For the outcome of event-free survival, the HR was 0.63 (95% CI 0.48 to 0.82). This was described by the CS as representing a 37% reduction in the risk of disease progression precluding definitive surgery, recurrence, second primary malignancy, or death compared with placebo + chemotherapy followed by placebo. Table 3.14 summarises the analysis of EFS, and Table 3.15 summarises the first event in EFS analyses.

Table 3.14: Analysis of event free survival (All participants)

Treatment	N	Number of events (%)	Person-months	Event rate/100 person-months	Median EFS [months] (95% CI) [a]	EFS Rate at 42 months % (95% CI)	Versus control Hazard Ratio (95% CI) [b]
Pembrolizumab arm	784	123 (15.7)	26,994.6	0.5	NR	83.5 (80.5, 86.0)	0.63 (0.48, 0.82) p-value [c] 0.0003093
Placebo arm	390	93 (23.8)	12,783.8	0.7	NR	74.9 (69.8, 79.2)	

Based on Table 13 of the CS[1]
[a] From product-limit (Kaplan-Meier) method for censored data.
[b] Based on Cox regression model with Efron's method of tie handling with treatment as a covariate stratified by nodal status (positive versus negative), tumour size (T1/T2 versus T3/T4) and choice of carboplatin (Q3W versus Weekly).
c One-sided p-value based on log-rank test stratified by nodal status (positive versus negative), tumour size (T1/T2 versus T3/T4) and choice of carboplatin (Cb) (Q3W versus Weekly).
Database Cut-off Date: 23MAR2021
CS = company submission; CI = confidence interval; EFS = event-free survival; NR = not reported; Q3W =every three weeks

Table 3.15: Summary of first event in EFS analyses

Event	Pembrolizumab arm (n=784)	Placebo arm (n=390)
	n (%)	n (%)
Any EFS Event	123 (15.7)	93 (23.8)
Secondary Primary Malignancy	6 (0.8)	4 (1.0)
Local PD Precludes Surgery	[redacted]	[redacted]
Local PD Precludes Definitive Surgery	[redacted]	[redacted]
Distant PD	[redacted]	[redacted]
Positive Margin at Last Surgery	[redacted]	[redacted]
Local Recurrence	28 (3.6)	17 (4.4)
Distant Recurrence	60 (7.7)	51 (13.1)
Death	15 (1.9)	6 (1.5)

Based on Table 15 of the CS[1]
Database Cut-off Date: 23MAR2021.
CS = company submission; EFS = event-free survival; PD = progressed disease

ERG comment: The CS refers to a 37% reduction in risk in relation to the HR of 0.63.[1] However, caution should always be taken with interpretation of the clinical importance of HRs as they cannot be interpreted in the same way as risk ratios.[36] Although the 37% reduction in hazard of recurrence is of large magnitude, this cannot be taken to imply that a similar difference in survival from recurrence will exist between the groups at longer time intervals.[36] Hence the clinical importance of this result is likely to be less clear-cut than that implied by the stated 37% reduction in "risk".

3.2.5.3 Overall Survival (OS)

The OS HR was 0.72 (95% CI 0.51 to 1.02), which was described as representing a 28% reduction in the risk of death compared with the placebo arm (Table 3.16). The median OS was not reached in either arm at month 42 and will need to be analysed in future IA as data matures.

Table 3.16: Analysis of OS (All participants)

Treatment	N	Number of events (%)	Person-months	Event rate/100 person-months (%)	Median OS [a] [months] (95% CI)	OS Rate at month 42 in %[†] (95% CI)	Versus control Hazard Ratio (95% CI) [b] p-value [c]
Pembrolizumab arm	784	80 (10.2)	28,1997.7	0.3	NR	89.2 (86.7, 91.3)	0.72 (0.51, 1.02) 0.0321377
Placebo arm	390	55 (14.1)	13,908.1	0.4	NR	84.1 (79.5, 87.7)	

Based on Table 16 of the CS[1]
[a] From product-limit (Kaplan-Meier) method for censored data.
[b] Based on Cox regression model with Efron's method of tie handling with treatment as a covariate stratified by nodal status (positive versus negative), tumour size (T1/T2 versus T3/T4) and choice of carboplatin (Q3W versus Weekly).
[c] One-sided p-value based on log-rank test stratified by nodal status (positive versus negative), tumour size (T1/T2 versus T3/T4) and choice of carboplatin (Cb) (Q3W versus Weekly).
Database Cut-off Date: 23MAR2021
CS = company submission; CI = confidence interval; NR = not reached; OS = overall survival; Q3W = every three weeks

ERG comment: The CS refers to a 28% reduction in risk in relation to the HR of 0.72.[1] However, as stated previously, caution should always be taken with interpretation of the clinical importance of HRs as they cannot be interpreted in the same way as risk ratios.[36] Hence the clinical importance of this result is unclear. This is particularly true given that there is insufficient evidence to reject the null hypothesis that the two intervention strategies have the same effects.

3.2.5.4 Quality of life

Three patient reported outcomes (PRO) questionnaires were used to assess patient HRQoL in the study for both the neoadjuvant and adjuvant phases: European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire (EORTC QLQ-C30), Breast Cancer-Specific Quality of Life Questionnaire (QLQ-BR23) and European Quality of Life-5 Dimensions (EQ-5D visual analogue scale (VAS)) PRO analyses were based on the PRO full analysis set (FAS) population, which included all randomised participants who had at least one PRO assessment available and had received at least one study treatment.

3.2.5.4.1 Neoadjuvant phase

At Week 21, the difference in mean change from baseline in EQ-5D VAS score between the pembrolizumab arm and placebo arm was [redacted] points (95% CI: -[redacted]), as shown in Table 3.17.

Table 3.17: Analysis of change from neoadjuvant baseline in EQ-5D VAS at neoadjuvant week 21 - All participants (FAS population)

Treatment	Baseline		Neoadjuvant Week 21		Change from Baseline at Week 21	
	N	Mean (SD)	N	Mean (SD)	N	LS Mean (95% CI) [a]
Pembrolizumab + chemotherapy	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Placebo + chemotherapy	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Pairwise comparison					**Difference in LS Means (95% CI)**	**p-Value**
Pembrolizumab + chemotherapy versus Placebo + chemotherapy					[redacted]	[redacted]

Based on Table 18 of the CS[1]
[a] Based on cLDA model with the PRO score as the response variable, and treatment by timepoint interaction, stratification factors (Nodal status (positive vs negative), Tumour size (T1/T2 vs T3/T4), and Choice of Carboplatin (Q3W vs Weekly)) as covariates.
For Neoadjuvant Baseline and Neoadjuvant Week 21, N is the number of participants in each treatment group with non-missing assessments at the specific time point; for change from Neoadjuvant Baseline, N is the number of participants in the analysis population in each treatment group.
CI = confidence interval; CS = company submission; EQ-5D = European Quality of Life-5 Dimensions FAS = full analysis set; Q3W = every three weeks; SD = standard deviation; VAS = visual analogue scale

3.2.5.4.2 Adjuvant phase

At Week 24 (of the adjuvant phase) the difference in mean change from baseline in EQ-5D VAS score between the pembrolizumab arm and placebo arm was [redacted] points (95% CI: [redacted], [redacted]), see Table 3.18.

Table 3.18: Analysis of change from adjuvant baseline in EQ-5D VAS at adjuvant week 24 - all participants (FAS population)

Treatment	Baseline		Adjuvant Week 24		Change from Baseline at Week 24	
	N	Mean (SD)	N	Mean (SD)	N	LS Mean (95% CI) [a]
Pembrolizumab monotherapy	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Placebo monotherapy	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Pairwise comparison					**Difference in LS Means (95% CI)**	**p-Value**
Pembrolizumab + versus Placebo					[redacted]	[redacted]

Based on Table 19 of the CS[1]
[a] Based on cLDA model with the PRO score as the response variable, and treatment by timepoint interaction, stratification factors (Nodal status (positive vs negative), Tumour size (T1/T2 vs T3/T4), and Choice of Carboplatin (Q3W vs Weekly)) as covariates.

Treatment	Baseline		Adjuvant Week 24		Change from Baseline at Week 24	
	N	Mean (SD)	N	Mean (SD)	N	LS Mean (95% CI) [a]
For Adjuvant Baseline and Adjuvant Week 24, N is the number of participants in each treatment group with non-missing assessments at the specific time point; for change from Adjuvant Baseline, N is the number of participants in the analysis population in each treatment group. CI = confidence interval; CS = company submission; EQ-5D = European Quality of Life-5 Dimensions FAS = full analysis set; Q3W = every three weeks; SD = standard deviation; VAS = visual analogue scale						

Further details of the EORTC QLQ-C30 and QLQ-BR23 results have been presented in Section 11.2.5 of the KEYNOTE-522 CSR.[28]

ERG comment: The lack of relative benefit for the pembrolizumab arm in terms of quality of life is an important finding. This may reflect the modest benefits observed for the other efficacy outcomes, alongside the significant adverse effect burden of pembrolizumab (see Section 3.2.6).

3.2.5.5 Subgroup analyses

A series of analyses were pre-specified in the KEYNOTE-522 study protocol to determine whether the treatment effect was consistent across various subgroups. The estimate of the between group treatment effect (with a nominal 95% CI) for the primary endpoints were estimated and plotted within each category of the following:

- Nodal status (positive versus negative)
- Tumour size (T1/T2 versus T3/T4)
- Choice of carboplatin (Q3W versus weekly)
- PD-L1 CPS (≥1 vs <1, ≥10 versus <10, ≥20 versus <20)
- Overall stage (Stage II versus stage III)
- Menopausal status (Pre versus post)
- Age (<65 years versus ≥ 65)
- Geographic region (Europe/Israel/North America/Australia versus Asia versus Rest of the world)
- Ethnic origin (Hispanic versus non-Hispanic)
- ECOG performance status (0 versus 1)
- HER2 status by IHC (2+ but FISH versus 0-1)
- LDH (>Upper limit of normal (ULN) versus ≤ ULN)

The treatment difference of pembrolizumab + chemotherapy compared with placebo + chemotherapy across prespecified subgroup analysis was generally consistent with the finding in the ITT population, showing directionally favourable improvement in pCR in the pembrolizumab + chemotherapy group (Figure 3.1). The same is also true for EFS (Figure 3.2). Due to the small number of events in subgroups, the results should be interpreted with caution.

ERG comment: The ERG noted that subgroup analyses results indicated potential differences between ECOG PS. In particular, in contrast to the ECOG = 0 sub-group, the sub-group with ECOG = 1 did not demonstrate benefits from pembrolizumab in terms of pCR (Figure 3.1). Thus, the ERG also asked the company to discuss the implications for decision making. The company responded by stating that *"a comparison of baseline characteristics … for all participants in KEYNOTE-522 with an ECOG PS of 1 demonstrated that, compared with the placebo + NAC / placebo group, participants in the pembrolizumab + NAC / pembrolizumab subgroup were older (median age of 53.5 years vs 47.0 years) and included greater proportions (≥5 percentage points) of the*

following parameters: participants who were post-menopausal, participants with PD-L1 positive tumors (CPS cutoff of 10), and participants with a primary tumor size of T3/T4, respectively".[6] This is noted as key issue 5.

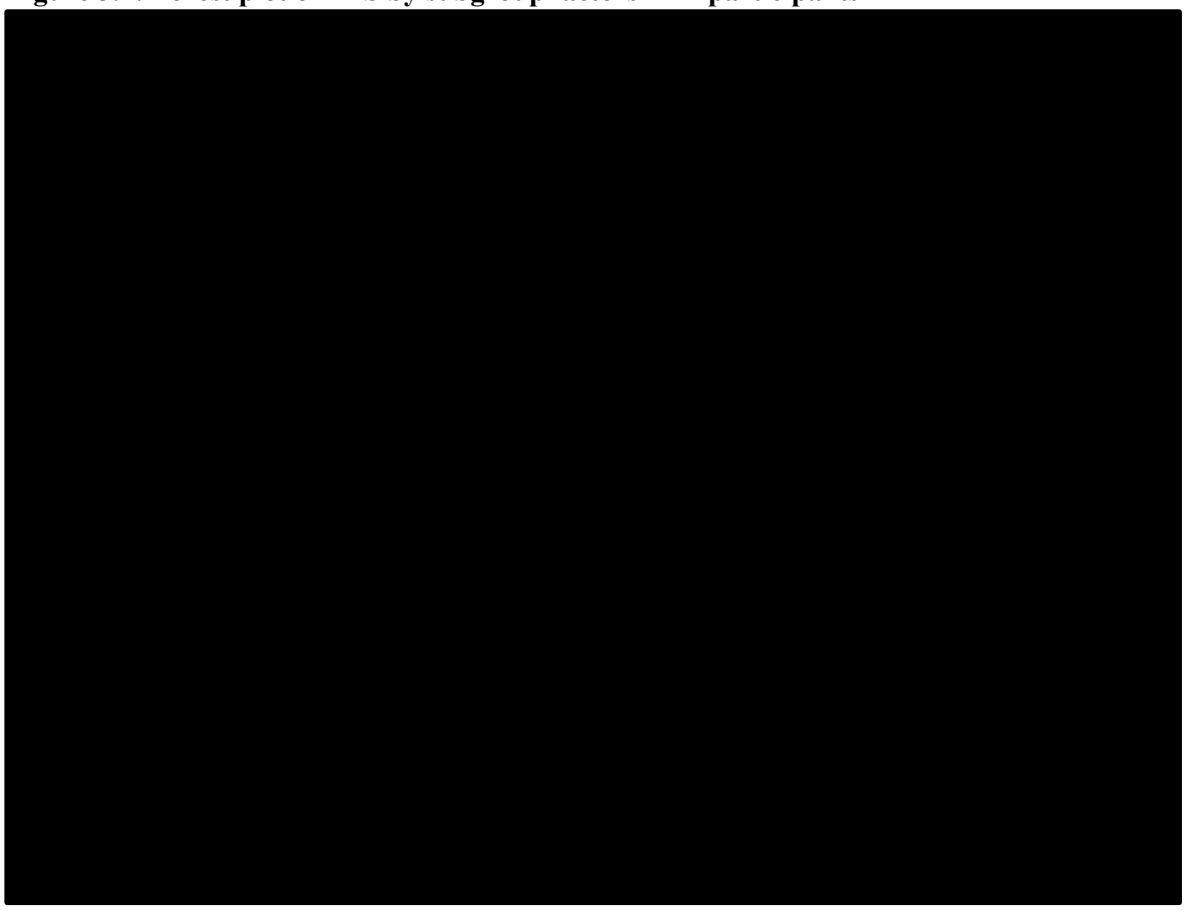
A closely related point was made in the succeeding paragraph. These statements did not provide information relevant to decision-making (in terms of the groups for which pembrolizumab might be, or might not be, useful) and merely described the characteristics expected to be associated with an ECOG of 1. Attempts to adjust for these covariates were made by the company in post-hoc analyses, which of course, removed the negative effects of the highly correlated ECOG variable upon the outcome. These did not show anything other than confirm the evident correlation. Associations of ECOG status with these characteristics are likely to be non-random effects related to the intrinsic nature of ECOG status, and so such an adjustment with these highly correlated variables was inappropriate. This can be demonstrated by considering that the relationship between ECOG status and its correlates of age or menopause status are analogous to the relationship expected between the variable of frailty and its correlates of old age and muscle weakness. One would not adjust frailty for old age and muscle weakness and then claim that frailty does not have an impact on the outcome of falls (because frailty *is* old age and muscle weakness), and in the same way it is not correct to adjust for age and menopause status and then claim that ECOG status has no effect on the outcome of pCR (because you are effectively adjusting ECOG out of the equation through multicollinearity). The important point is that these correlating characteristics do not prevent people with ECOG 1 being less appropriate candidates for pembrolizumab, and it is likely that if people have an ECOG score of 1 they are not going to experience benefits from pembrolizumab. The company stated that numbers were small and that therefore it was difficult to form conclusions, but the data suggest that patients with an ECOG status of 1 are unlikely to benefit from pembrolizumab (and there is a probability that the drug could even cause harm in this group, although this is uncertain).

Sub-group analyses also demonstrated potential differences between geographical regions. This has been commented on in detail in Section 3.2.1.

Figure 3.1: Forest plot of pCR (ypT0/Tis ypN0) by subgroup factors - All participants

Based on Figure 6 of the CS[1]
CI = confidence interval; CPS = combined positive score; CS = company submission; ECOG = Eastern Cooperative Oncology Group; FISH = fluorescence in situ hybridization; HER2 = human epidermal growth factor receptor 2; IHC = immunohistochemistry; pCR = pathological complete response; PS = performance status; Q3W = every three weeks; ULN = upper limit of normal

Figure 3.2: Forest plot of EFS by subgroup factors - All participants



Based on Figure 6 of the CS[1]
CI = confidence interval; CPS = combined positive score; CS = company submission; ECOG = Eastern Cooperative Oncology Group; FISH = fluorescence in situ hybridization; HER2 = human epidermal growth factor receptor 2; IHC = immunohistochemistry; pCR = pathological complete response; PS = performance status; Q3W = every three weeks; ULN = upper limit of normal

3.2.6 Adverse events of the KEYNOTE-522 trial

The CS reported that safety results of KEYNOTE-522 demonstrated pembrolizumab plus chemotherapy followed by pembrolizumab monotherapy had a manageable safety profile in participants with high-risk, early-stage TNBC, and that the safety profile of the pembrolizumab arm is generally consistent with the known safety profile of pembrolizumab monotherapy and a carboplatin-/anthracycline-based chemotherapy regimen.[1] No new safety concerns were identified.

During the combined phases, the overall incidence of AEs, drug-related AEs, Grade 3 to 5 AEs, Grade 3 to 5 drug-related AEs, deaths, deaths due to drug-related AEs, and any dose modification due to an AE were generally similar between the pembrolizumab arm and the placebo arm. However, there was a higher overall incidence of serious adverse effects (SAEs), serious drug-related AEs, and discontinuations of any drug due to an AE in the pembrolizumab arm compared with the placebo arm, reflecting the contribution of both pembrolizumab and neoadjuvant chemotherapy.

3.2.6.1 Summary of adverse events

According to the CS, comparable proportions of patients in the pembrolizumab and placebo arms experienced AEs (drug-related AEs (98.9% versus 99.7%), Grade 3 to 5 AEs

([redacted]), Grade 3 to 5 drug-related AEs (77.1% versus 73.3%), deaths ([redacted]), deaths due to drug-related AEs (0.5% versus 0.3%), and any dose modification due to an AE ([redacted]).[1]

There was a higher incidence (≥5 percentage points difference) of serious adverse events (SAEs, [redacted]%), serious drug-related AEs ([redacted]%), and discontinuations of any drug due to an AE ([redacted]%) in the pembrolizumab arm compared with the placebo arm.

Included adverse events started from the first treatment including definitive surgery and radiation therapy and up to 30 days of the last treatment including definitive surgery and radiation therapy for the non- SAEs and up to 90 days of the last treatment including definitive surgery and radiation therapy for the SAEs.

Table 3.19 summarises adverse events and effects on continuation.

Table 3.19: Adverse event summary - Combined phases (All participants)

	Pembrolizumab arm (n=789)		Placebo arm (n=389)	
	n	(%)	n	(%)
with one or more adverse events	[redacted]	[redacted]	[redacted]	[redacted]
with no adverse event	[redacted]	[redacted]	[redacted]	[redacted]
with drug-related [a] adverse events	774	(98.9)	388	(99.7)
with toxicity Grade 3-5 adverse events	[redacted]	[redacted]	[redacted]	[redacted]
with toxicity Grade 3-5 drug-related adverse events	604	(77.1)	285	(73.3)
with serious adverse events	[redacted]	[redacted]	[redacted]	[redacted]
with serious drug-related adverse events	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
who died	7	(0.9)	1	(0.3)

	Pembrolizumab arm (n=789)		Placebo arm (n=389)	
	n	(%)	n	(%)
who died due to a drug-related adverse event	4	(0.5)	1	(0.3)
[redacted]				
[redacted]				
[redacted]				
[redacted]				
[redacted]				
[redacted]				
[redacted]				
[redacted]				
[redacted]				
[redacted]*				
[redacted]				
[redacted]				
[redacted]				
[redacted]				
[redacted]				
[redacted]				
[redacted]				
[redacted]*				
[redacted]				
[redacted]				
[redacted]				
[redacted]				
[redacted]				
[redacted]				
[redacted]*				
[redacted]				
[redacted]				
[redacted]				
[redacted]				

Based on Table 24 of the CS[1]

[a] Determined by the investigator to be related to the drug.

	Pembrolizumab arm (n=789)		Placebo arm (n=389)	
	n	(%)	n	(%)
[b] Defined as an action taken of dose reduced, drug interrupted, or drug withdrawn. Grades are based on NCI CTCAE version 4.0.[37] MedDRA preferred terms "Neoplasm Progression", "Malignant Neoplasm Progression" and "Disease progression" not related to the drug are excluded. Database Cut-off Date: 23MAR2021 CS = company submission; CTCAE = Common Terminology Criteria for Adverse Events; MedDRA = Medical Dictionary for Regulatory Activities; NCI = National Cancer Institute				

The most frequently reported AEs (incidence ≥30%) in either arm were [redacted].

AEs (incidence ≥15%) with a greater risk difference for pembrolizumab arm (where the lower bound of the 95% CI for the treatment difference was >0) during the combined phases were [redacted]. These events were primarily Grade 1 or 2. There were no AEs (incidence ≥15%) with a greater risk difference for the placebo arm identified during the combined phases. In both treatment arms, most AEs occurred in the first 3 months of initiating study intervention; the exposure-adjusted event rate decreased at 3 to 6 months and continued to decrease beyond 12 months (Table 3.20).

Table 3.20: Participants with AEs by decreasing incidence (incidence ≥10% in at least one arm; ASaT population)

	Pembrolizumab arm		Placebo arm	
	n	(%)	n	(%)
Participants in population	783		389	
with one or more adverse events	[redacted]	[redacted]	[redacted]	[redacted]
with no adverse events	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

	Pembrolizumab arm		Placebo arm	
	n	(%)	n	(%)
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Based on Table 25 of the CS[1]
Every participant is counted a single time for each applicable specific adverse event.

	Pembrolizumab arm		Placebo arm	
	n	(%)	n	(%)
MedDRA preferred terms "Neoplasm progression", "Malignant neoplasm progression" and "Disease progression" not related to the drug are excluded. Database Cut-off Date: 23MAR2021 CS = company submission; MedDRA = Medical Dictionary for Regulatory Activities;				

ERG comment: The CS reported that the risk of deaths were comparable between arms.[1] However, the risk of deaths in the pembrolizumab arm were three times that of the placebo arm, please see Section 3.2.6.6 for further comments related to this issue.

3.2.6.2 Drug related AEs

The overall incidences of drug-related AEs during the combined phases were similar between the pembrolizumab (98.9%) and placebo (99.7%) arms (Table 3.21).

The incidences of the most frequently reported drug-related AEs (incidence ≥30%) during the combined phases were generally similar between the two treatment groups (Table 3.21) and included:

- Pembrolizumab arm: nausea, alopecia, anaemia, neutropenia, fatigue, and diarrhoea.
- Placebo arm: nausea, alopecia, anaemia, neutropenia, and fatigue.

Table 3.21: Participants with drug related AEs by decreasing incidence (incidence ≥5% in one or more treatment arms; ASaT population)

	Pembrolizumab arm		Placebo arm	
	n	(%)	n	(%)
Participants in population	783		389	
with one or more adverse events	774	(98.9)	388	(99.7)
with no adverse events	9	(1.1)	1	(0.3)
Nausea	495	(63.2)	245	(63)
Alopecia	471	(60.2)	220	(56.6)
Anaemia	429	(54.8)	215	(55.3)
Neutropenia	367	(46.9)	185	(47.6)
Fatigue	330	(42.1)	151	(38.8)
Diarrhoea	238	(30.4)	98	(25.2)
Alanine aminotransferase increased	204	(26.1)	98	(25.2)
Asthenia	198	(25.3)	102	(26.2)
Neutrophil count decreased	185	(23.6)	112	(28.8)
Vomiting	200	(25.5)	86	(22.1)
Constipation	188	(24)	85	(21.9)
Rash	196	(25)	66	(17)
Neuropathy peripheral	154	(19.7)	84	(21.6)
Aspartate aminotransferase increased	157	(20.1)	63	(16.2)
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

	Pembrolizumab arm		Placebo arm	
	n	(%)	n	(%)
[redacted]				

Based on Table 26 of the CS[1]
Every participant is counted a single time for each applicable specific adverse event.
Database Cut-off Date: 23MAR2021
CS = company submission

3.2.6.3 Grade 3 to 5 AEs

The overall incidence of Grade 3 to 5 AEs during the combined phases was generally similar between the 2 treatment groups arms (Table 3.22). There were no specific trends noted in the pembrolizumab

arm that suggest any new safety concerns. The types and frequencies of the most common Grade 3 to 5 AEs (incidence ≥5%) during the combined phases were generally similar between the 2 treatment arms. The only risk difference of Grade 3 to 5 AEs (incidence ≥5%) during the combined phases that favoured either treatment group was [redacted], which had a greater risk in the pembrolizumab arm (where the lower bound of the 95% CI for the treatment difference was >0).

Table 3.22: Participants with Grade 3-5 AEs by decreasing incidence (incidence ≥5% in one or more treatment arms; ASaT population)

	Pembrolizumab arm		Placebo arm	
	n	(%)	n	(%)
Participants in population	783		389	
with one or more adverse events	[redacted]	[redacted]	[redacted]	[redacted]
with no adverse events	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Based on Table 27 of the CS[1]
Every participant is counted a single time for each applicable specific adverse event.
Grades are based on NCI CTCAE version 4.0.
MedDRA preferred terms "Neoplasm progression", "Malignant neoplasm progression" and "Disease progression" not related to the drug are excluded.
Database Cut-off Date: 23MAR2021
CS = company submission; CTCAE = Common Terminology Criteria for Adverse Events; MedDRA = Medical Dictionary for Regulatory Activities; NCI = National Cancer Institute

ERG comments: No comments

3.2.6.4 Drug related Grade 3-5 AEs

The overall incidences of drug-related Grade 3 to 5 AEs as determined by the investigator during the combined phases were generally similar between the pembrolizumab (77.1%) and placebo arms (73.3%). The incidences of the most frequently reported drug-related Grade 3 to 5 AEs (incidence ≥5%) during the combined phases were generally similar between treatment groups (Table 3.23).

Table 3.23: Participants with drug related Grade 3-5 AEs by decreasing incidence (incidence ≥5% in one or more treatment arms; ASaT population)

	Pembrolizumab arm		Placebo arm	
	n	(%)	n	(%)
Participants in population	783		389	
with one or more adverse events	604	(77.1)	285	(73.3)
with no adverse events	179	(22.9)	104	(26.7)
Neutropenia	270	(34.5)	130	(33.4)
Neutrophil count decreased	146	(18.6)	90	(23.1)

	Pembrolizumab arm		Placebo arm	
	n	(%)	n	(%)
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Anaemia	141	(18)	58	(14.9)
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Alanine aminotransferase increased	43	(5.5)	9	(2.3)
Based on Table 28 of the CS[1] Every participant is counted a single time for each applicable specific adverse event. Grades are based on NCI CTCAE version 4.0. Database Cut-off Date: 23MAR2021 CS = company submission; CTCAE = Common Terminology Criteria for Adverse Events; NCI = National Cancer Institute				

ERG comments: No comments

3.2.6.5 Serious adverse effects

The overall incidence of SAEs was higher in the pembrolizumab arm compared with the placebo arm. The SAEs observed for participants in the pembrolizumab arm were reported by the company [1]to be generally consistent with the known safety profiles of pembrolizumab monotherapy and a carboplatin-/anthracycline-based chemotherapy regimen (Table 3.24).

Table 3.24: Participants with serious AEs up to 90 days after last dose by decreasing incidence (incidence ≥1% in one or more treatment arms; ASaT population)

	Pembrolizumab arm		Placebo arm	
	n	(%)	n	(%)
Participants in population	783		389	
with one or more adverse events	[redacted]	[redacted]	[redacted]	[redacted]
with no adverse events	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Based on Table 29 of the CS[1] Every participant is counted a single time for each applicable specific adverse event.				

	Pembrolizumab arm		Placebo arm	
	n	(%)	n	(%)
MedDRA preferred terms "Neoplasm progression", "Malignant neoplasm progression" and "Disease progression" not related to the drug are excluded. Database Cut-off Date: 23MAR2021 CS = company submission; MedDRA = Medical Dictionary for Regulatory Activities				

ERG comments: The difference between arms in SAEs is large and requires consideration in the overall evaluation of the study drug.

3.2.6.6 Deaths due to Adverse Events

Deaths due to AEs during the combined phases occurred in [redacted] ([redacted]%) participants in the pembrolizumab arm and [redacted] ([redacted]%) participant in the placebo arm. There were 4 deaths in the pembrolizumab arm considered drug related. Deaths due to AE in 3 participants were considered related to pembrolizumab (pneumonitis in 1 participant in the neoadjuvant phase, pulmonary embolism in 1 participant in the adjuvant phase, and autoimmune encephalitis in 1 participant in the adjuvant phase). One participant in the neoadjuvant phase experienced 3 AEs resulting in death: sepsis and multiple organ dysfunction syndrome, which were considered related to chemotherapy, and myocardial infarction, which was not considered to be drug related. In the placebo arm, the 1 reported death due to an AE (septic shock) occurred during the neoadjuvant phase and was considered related to chemotherapy by the investigator. No new safety signals were identified upon review of these fatal events.

ERG comments: For pembrolizumab versus placebo, the relative risk of death is 3, which requires consideration in the overall evaluation of the study drug. The probability of a difference this large arising by chance is 0.01. This, together with comments in Section 3.2.6.5, has been noted as key issue 6.

3.2.6.7 Adverse events of special interest (AEOSI)

The overall incidence of AEOSI during the combined phases was higher in the pembrolizumab arm (43.6%) compared with the placebo arm (21.9%).

There were 2 deaths due to an AEOSI (pneumonitis and autoimmune encephalitis) in the pembrolizumab arm, which were considered related to pembrolizumab by the investigator. The most frequently reported AEOSIs (incidence ≥5%) by category, during the combined phases were hypothyroidism, infusion reactions, severe skin reactions, and hyperthyroidism in the pembrolizumab arm and hypothyroidism and infusion reactions in the placebo arm. The incidence of hypothyroidism in the pembrolizumab arm was higher than anticipated based on the known safety profile of pembrolizumab monotherapy and higher than the placebo arm (Table 3.26).

Table 3.25: Participants with AEOSI by category (incidence >0% in one or more treatment arms; ASaT population

	Pembrolizumab arm		Placebo arm	
	n	(%)	n	(%)
Participants in population	783		389	
with one or more adverse events	341	(43.6)	85	(21.9)
with no adverse events	442	(56.4)	304	(78.1)
Infusion Reactions	141	(18)	45	(11.6)

	Pembrolizumab arm		**Placebo arm**	
	n	**(%)**	**n**	**(%)**
Hypothyroidism	118	(15.1)	22	(5.7)
Severe Skin Reactions	45	(5.7)	4	(1)
Hyperthyroidism	41	(5.2)	7	(1.8)
Adrenal Insufficiency	20	(2.6)	0	(0)
Pneumonitis	17	(2.2)	6	(1.5)
Thyroiditis	16	(2)	5	(1.3)
Hypophysitis	15	(1.9)	1	(0.3)
Colitis	13	(1.7)	3	(0.8)
Hepatitis	11	(1.4)	3	(0.8)
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Based on Table 30 of the CS[1] Every participant is counted a single time for each applicable specific adverse event. A participant with multiple adverse events within a bolded term is counted a single time for that bolded term. "Infusion related reaction" includes infusion related reactions due to pembrolizumab and chemotherapy, for example, Paclitaxel. Database Cut-off Date: 23MAR2021 CS = company submission				

ERG comments: No comments.

3.2.7 Included studies: Supporting evidence

Not applicable.

3.2.8 Ongoing studies

The next database cut off (IA5) is calendar-driven and will take place in [redacted].

ERG comment: In the clarification letter, the ERG has requested to know when data from IA5 can be made available.[5] The response from the company is *"as dual-primary endpoints pCR (at IA1) and EFS (at IA4) achieved statistical significance, the study continues to follow OS in a blinded manner. Per the protocol, the next interim analysis (IA5) will occur ~60 months after the first participant was*

randomized, 1 year after IA4. If OS achieves statistical significance, the external DMC will inform MSD and updated efficacy results may be available in Q3 2022. If OS doesn't achieve statistical significance, the study will continue in a blinded manner".[6]

3.3 Critique of trials identified and included in the indirect comparison and/or multiple treatment comparison

No indirect comparison (IC) and/or multiple treatment comparison was carried out.

3.4 Critique of the indirect comparison and/or multiple treatment comparison

According to Section B.2.9 of the CS, *"clinical expert advice sought confirmed that the KEYNOTE-522 study design and choice of comparators is appropriate and generalisable of the treatment pathway in the UK setting".*[1] The ERG asked the company to provide supporting references and please provide a report describing the clinical expert advice solicitation.[5] The response from the company is that *"the report from the advisory board is provided as a separate confidential reference for consideration".*[6]

3.5 Additional work on clinical effectiveness undertaken by the ERG

Not applicable.

3.6 Conclusions of the clinical effectiveness section

The CS and response to clarification provided full details for the ERG to appraise the literature searches conducted to identify studies about the clinical efficacy and safety outcomes of pembrolizumab + chemotherapy and competing interventions for the neoadjuvant treatment of locally advanced non-metastatic TNBC. The searches were conducted in July 2021. Searches were transparent and reproducible, and comprehensive search strategies were used. A good range of databases and grey literature resources were searched. Strategies included an extensive list of search terms for the population and comparators, and validated search filters for study design. The ERG was concerned about the language bias of restricting searches to English language only.

The evidence from the CS suggests that pembrolizumab given alongside standard neoadjuvant therapy, followed by pembrolizumab given alone in the adjuvant phase, is more clinically effective than placebo given alongside standard neoadjuvant therapy, followed by placebo given alone in the adjuvant phase.[1] The intervention arm demonstrated a benefit in event-free survival (HR 0.63, 95% CI 0.48 to 0.82), a small but significant benefit in pCR (absolute risk difference of 7.5% (95% CI: 1.6% to 13.4%), equating to a number-needed-to-treat of around 13) and a trend for a benefit in OS (HR 0.72, 95% CI 0.51 to 1.02). However, benefits in terms of quality of life were not observed, suggesting that the net positive balance between clinical benefits and harms of pembrolizumab were insufficient to have a positive impact on patients' quality of life.

Although the adverse events of the intervention are reported by the CS to be consistent with expectations, 43.6% of participants in the pembrolizumab arm experienced SAEs, compared to 28.5% of participants in the placebo arm, and three times the proportion of participants died in the pembrolizumab arm (0.9%) compared to the placebo arm (0.3%).[1] The moderate benefits of pembrolizumab therefore need to be considered in the light of its potential harms.

An important issue for consideration is the choice of comparator in the trial. It is likely that the use of placebo in the adjuvant phase, rather than an active comparator such as capecitabine (which is associated with an improvement in DFS) may have contributed to an increased estimate of benefit for

pembrolizumab. Whilst this observed benefit may be realistic in terms of comparison to established practice, it may be over-optimistic in evaluating pembrolizumab in relation to the best available alternative therapies.

4. COST EFFECTIVENESS

4.1 ERG comment on company's review of cost effectiveness evidence

One set of systematic literature searches was performed to identify CE studies, and cost and healthcare resource use studies (CS Appendices G and I), and a separate search was conducted to identify HRQoL studies (Appendix H).[1, 4]

4.1.1 Searches performed for cost effectiveness section

The following paragraphs contain summaries and critiques of the searches related to CE presented in the CS. The CADTH evidence-based checklist for the Peer Review of Electronic Search Strategies (PRESS), was used to inform this critique.[10, 11] The CS was checked against the single technology appraisal (STA) specification for company/sponsor submission of evidence.[12]

Appendices G and I of the CS reported the literature searches used to identify CE studies, and cost and healthcare resource use studies.[4] Appendix H reported the literature searches used to identify HRQoL studies.[4] All searches were conducted on 16 May 2021.

A summary of the resources searched for CE studies, HRQoL studies, and cost and healthcare resource use studies is provided in Table 4.1.

Table 4.1: Resources searched for cost effectiveness studies, HRQoL studies, and cost and healthcare resource use studies (as reported in CS)

Resource	Host/Source	Date Ranges	Date Searched
Electronic databases			
MEDLINE and Epub Ahead of Print, In-Process, In-Data-Review & Other Non-Indexed Citations	Ovid	1946 to 14 May 2021	16/05/21
Embase	Ovid	1974 to 14 May 2021	16/05/21
CENTRAL	EMB Reviews, Ovid	April 2021	16/05/21
CDSR	EMB Reviews, Ovid	2005 to 12 May 2021	16/05/21
EconLit	Ovid	1886 to 6 May 2021	16/05/21
Additional resources			
HERC Database of Mapping Studies	NR	NR	NR
ScHARRHUD	NR	NR	NR
Conference proceedings			
ASCO	Northern Light, Ovid	2016-2020	16/05/21
ESMO	Northern Light, Ovid	2016-2020	16/05/21
ISPOR Annual European Conference	Northern Light, Ovid	2016-2020	16/05/21
ISPOR Annual Asian Conference	Northern Light, Ovid	2016-2020	16/05/21

Resource	Host/Source	Date Ranges	Date Searched
ISPOR Annual International Meeting North America	Northern Light, Ovid	2016-2020	16/05/21
NCCN	Northern Light, Ovid	2016-2020	16/05/21
SABCS	Northern Light, Ovid	2016-2020	16/05/21
HTA organisations			
AHRQ	NR	NR	NR
NIHR HTA	NR	NR	NR
INAHTA	NR	NR	NR
SMC	NR	NR	NR
AWMSG	NR	NR	NR
CADTH	NR	NR	NR
HAS	NR	NR	NR
IQWIG	NR	NR	NR
ICER	NR	NR	NR
NICE	NR	NR	NR

AHRQ = Agency for Healthcare Research and Quality; ASCO = American Society of Clinical Oncology; AWMSG = All Wales Medicines Strategy Group; CADTH = Canadian Agency for Drugs and Technologies in Health; CDSR = Cochrane Database of Systematic Reviews; CENTRAL = Cochrane Central Register of Controlled Trials; CS = company submission; ESMO = European Society for Medical Oncology congress; HAS = French National Authority for Health (Haute Autorité de Santé); HERC = Health Economics Research Centre; HTA = health technology assessment; ICER = Institute for Clinical and Economic Review; INAHTA = Health Technology Assessment database of the International Network of Agencies for Health Technology Assessment; IQWIG = Institute for Quality and Efficiency in Healthcare; ISPOR = International Society for Pharmacoeconomics and Outcomes Research; NCCN = National Comprehensive Cancer Network Annual Conference; NICE = National Institute for Health and Care Excellence; NIHR = National Institute of Health Research Health; NR = not reported; SABCS = San Antonio Breast Cancer Symposium; ScHARRHUD = School of Health and Related Research health utilities database; SMC = Scottish Medicines Consortium

ERG comment:

- The CS provided full details of the literature searches for the ERG to appraise.[1, 4]
- A comprehensive range of databases, supplementary resources, conference proceedings, and health technology assessment (HTA) organisation websites were searched.
- Full details of the database searches, including the database name, host platform, date range and date searched, were provided.
- Economic specific resources were searched, but details of the search strategies or search terms used, dates of searches, and results were not reported in the CS.[1, 4] The search terms used, and results, were provided in response to the ERG clarification letter; the full search strategies, and dates of searches, were not provided.[6]
- Conference proceedings were searched via the Northern Light Life Sciences Conference Abstracts database. The search strategies, date of searches, and results were not reported in the CS.[1, 4] No further details were provided in response to the ERG request for clarification.[6]
- A comprehensive list of HTA organisation websites was searched. The search strategies or search terms used, date of searches, and results, were not reported in the CS.[1, 4] In response to

the ERG request for clarification, the search terms used were provided. Full details of the search strategies were not provided, because *"across these resources, inconsistent formatting and search functionality often precluded the determination of the magnitude of the available materials. Thus, in accordance with historical precedent, detailed records of grey literature searches were not recorded in a manner analogous to that of the traditional database searches of Embase, MEDLINE, and CENTRAL"*.[6]

- The database search strategies were well structured, transparent and reproducible. They included truncation, proximity operators, synonyms, and subject headings (MeSH and EMTREE). There were no date limits.
- MeSH terms were used instead of EMTREE in the Embase search strategy, though Ovid does map to the correct subject heading when the search is conducted. Several MeSH and EMTREE terms were exploded when there were no terms beneath them in the tree hierarchy.
- The population facet of search terms could have been improved with more synonyms, fewer exact phrases, better use of proximity operators, and the removal of redundant terms/phrases. The combination of search terms for 'triple negative breast cancer' with search terms for 'breast cancer' using the Boolean AND was incorrect but had barely any impact on the search results.
- Study design search filters for economic studies were included in the CE, and costs and healthcare resources search strategies. Study design search filters for utilities studies were included in the health-related quality of life searches. Neither of the search filters used were cited, as current practice recommends.[16]
- The economic studies search filter used was designed to identify CE studies, and not to capture cost and healthcare resource use studies. More relevant search terms such as 'cost', 'resource use', 'employment', 'carers', etc., should have been included in the search strategy.
- The searches were limited to English language only studies and this may have introduced language bias. Best practice states that *"to reduce the risk of introducing bias, searches should not be restricted by language"*.[13] Any limits (including language) should be reported and justified according to Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) 2020 and PRISMA-S guidelines.[14-16]
- It would have been preferable for the database search strategies to be presented exactly as run, rather than copied into a tabular format, as Item 8 of the PRISMA-S checklist recommends.[16] The Cochrane Handbook also recommends that *"...bibliographic database search strategies should be copied and pasted into an appendix exactly as run and in full, together with the search set numbers and the total number of records retrieved by each search strategy. The search strategies should not be re-typed, because this can introduce errors"*.[17]
- The Cochrane Central Register of Controlled Trials (CENTRAL) and the Cochrane Database of Systematic Reviews (CDSR) were searched for economic studies, when one database consists of trials and the other consists of systematic reviews. It is possible that this was a reporting error, and that both databases were searched for the clinical effectiveness SLR.
- The searches were conducted in May 2021. An update of the searches immediately prior to submission to NICE would have been appropriate and could have identified potentially relevant records published since May 2021.
- In order to identify OS data for the economic model the company referred to a SLR conducted for another ongoing NICE submission.[38] Brief details of this SLR were reported in Appendix M.1.3.[4]

4.1.2 Inclusion/exclusion criteria

In- and exclusion criteria for the review on CE studies, HRQoL studies and costs and resource use studies are presented in Table 4.2.

Table 4.2: Eligibility criteria for the systematic literature reviews

	Inclusion criteria	Exclusion criteria
Patient population	• Early-stage locally advanced non-metastatic TNBC • Metastatic TNBC	
Intervention	Not restricted	
Comparator	Not restricted	
Outcomes(s) 1 (Published economic evaluations)	Costs combined with clinical endpoints (e.g. clinical outcomes, utilities, QALY, resource use, burden of illness)	
Outcomes(s) 2 (HRQoL studies)	**Treatment effects in terms of generic and disease-specific patient-reported outcomes and utilities:** • Generic PRO measures (EQ-5D, HUI-2, HUI-3, SF-6D, SF-36, EORTC QLQ-C30, PROMIS-Fatigue SF1, Q-TWIST, CTSQ, etc.) • Disease-specific HRQoL (EORTC QLQ-BR23, FACT-B—FBSI) • Utility measures **Utility values for different health states, disutility associated with AEs, and mapping algorithms:** • Preference measures (both generic and disease-specific non-preference-based measures not converted to utilities will be considered) • Utility values for health states stable disease, pre-progression, post-progression, responders, and by time prior to death • Disutility values associated with AEs **Patient-specific disease burden:** • Recommendations regarding use of PRO measures • PRO measures used in the target populations across different regions	
Outcomes(s) 3 (Cost/resource use studies)	• Direct costs • Indirect costs • Healthcare resource utilisation	
Study design 1 (Cost effectiveness analysis studies)	Primary research studies: • Full economic evaluations (e.g. CEA, cost-utility analyses, cost-benefit analyses, cost-consequence analyses) • Partial economic evaluations (e.g. cost-of-illness analyses, budget impact analyses, cost-minimization analyses) • Observational studies (e.g. prospective and retrospective cohort studies, case-control studies,	• Results are not available • Publication type not of interest (e.g. comment, editorial, letter, case report, animal study)

	Inclusion criteria	Exclusion criteria
	cross-sectional studies, controlled and uncontrolled longitudinal studies) • RCTs and non-RCTs HTAs Pooled analysis presenting the cost or resource use estimates Literature reviews summarizing results of primary research studies and/or economic evaluations	
Study design 2 (HRQoL studies)	**Treatment effects in terms of generic and disease-specific patient-reported outcomes and utilities:** • RCTS and non-RCTs • Economic evaluations reporting patient utility values (studies must provide extractable results) **Utility values for different health states, disutility associated with AEs, and mapping algorithms:** • Mapping algorithms that would allow a non-preference-based measure to be mapped onto a generic preference-based measure • Mapping algorithms between different generic preference-based health state utility values **Patient-specific disease burden:** • Observational studies reporting HRQoL/utility (e.g. controlled before-and-after studies, interrupted time series studies, historically controlled studies, prospective and retrospective cohort studies, case-control studies, cross-sectional studies, controlled and uncontrolled longitudinal studies) **All topics:** • Literature reviews summarizing results of primary research studies • Pooled analyses presenting QoL/utility data	• Results are not available • Publication type not of interest (e.g. comment, editorial, letter, case report, animal study)
Study design 3 (Cost/resource use studies)	• Full economic evaluations (e.g. CEA, cost-utility analyses, cost-benefit analyses, cost-consequence analyses) • Partial economic evaluations (e.g. cost-of-illness analyses, budget impact analyses, cost-minimization analyses) • Observational studies (e.g. prospective and retrospective cohort studies, case-control studies, cross-sectional studies, controlled and uncontrolled longitudinal studies) • RCTs and non-RCTs • Literature reviews summarizing results of primary research studies and/or economic evaluations	
Region	Global	
Publication date	No restriction	

	Inclusion criteria	Exclusion criteria
Language	Studies published in English will be included	

Based on Appendices G, H, and I of the CS[4]
CEA = cost effectiveness analysis, CS = company submission; EORTC = European Organisation for Research and Treatment of Cancer; EQ-5D = European Quality of Life-5 Dimensions; FACT-B-FBSI: Functional Assessment of Cancer Therapy Breast Symptom Index; HTA = health technology assessment; HUI = Health Utility Index; PRO = patient-reported outcomes; PROMIS-Fatigue SF1 = Patient-Reported Outcomes Measurement Information System Fatigue Short Form-1; QALY = quality adjusted life year; QLQ-BR23 = Cancer-Specific Quality of Life Questionnaire; QLQ-C30 = Quality of Life Questionnaire-Core 30; QoL = quality of life; CTSQ = Cancer Treatment Satisfaction Questionnaire; Q-TWIST = Quality-adjusted time without symptoms or toxicity; RCTs = randomised controlled trials; SF-6D = Short-Form Six-Dimension; SF-36 = 36-Item Short Form Survey; TNBC = triple-negative breast cancer

ERG comment: The ERG agrees that the eligibility criteria are suitable to fulfil the company's objective to identify CE studies. The rationales for excluding CE studies after full paper reviewing are considered appropriate given the defined in- and exclusion criteria.

4.1.3 Conclusions of the cost effectiveness review

The CS provides an overview of the included CE, HRQoL and resource use and costs studies, but no specific conclusion was formulated.

ERG comment: The CS and response to request for clarification provided sufficient details for the ERG to appraise the literature searches conducted to identify CE, HRQoL, cost and healthcare resource use studies for the treatment of patients in neoadjuvant and adjuvant TNBC.[1, 6] The searches were conducted in May 2021. Searches were transparent and reproducible, and comprehensive strategies were used. A good range of databases and grey literature resources were searched. The search strategies included validated search filters for study design. The ERG was concerned about the language bias of restricting searches to English language only.

4.2 Summary and critique of company's submitted economic evaluation by the ERG

4.2.1 NICE reference case checklist

Table 4.3: NICE reference case checklist

Element of HTA	Reference case	ERG comment on CS
Perspective on outcomes	All direct health effects, whether for patients or, when relevant, carers	In line with NICE reference case
Perspective on costs	NHS and PSS	In line with NICE reference case
Type of economic evaluation	Cost utility analysis with fully incremental analysis	In line with NICE reference case
Time horizon	Long enough to reflect all important differences in costs or outcomes between the technologies being compared	In line with NICE reference case
Synthesis of evidence on health effects	Based on systematic review	In line with NICE reference case

Element of HTA	Reference case	ERG comment on CS
Measuring and valuing health effects	Health effects should be expressed in QALYs. The EQ-5D is the preferred measure of HRQoL in adults.	In line with NICE reference case
Source of data for measurement of health-related quality of life	Reported directly by patients and/or carers	In line with NICE reference case
Source of preference data for valuation of changes in health-related quality of life	Representative sample of the UK population	In line with NICE reference case
Equity considerations	An additional QALY has the same weight regardless of the other characteristics of the individuals receiving the health benefit	In line with NICE reference case
Evidence on resource use and costs	Costs should relate to NHS and PSS resources and should be valued using the prices relevant to the NHS and PSS	In line with NICE reference case
Discounting	The same annual rate for both costs and health effects (currently 3.5%)	In line with NICE reference case

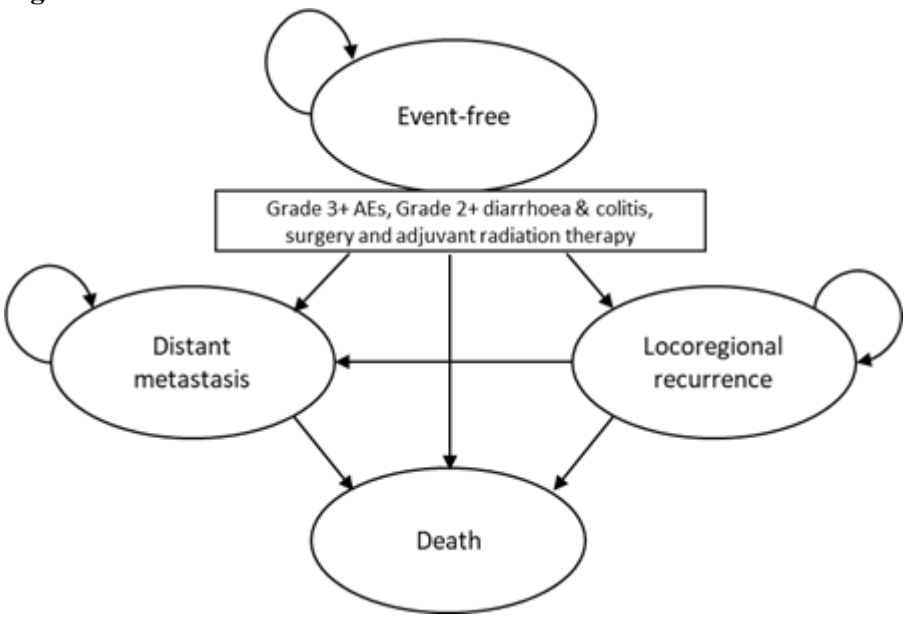
CS = company submission; EQ-5D = European Quality of Life-5 Dimensions; ERG = Evidence Review Group; HTA = health technology assessment; HRQoL = health-related quality of life; NHS = National Health Service; NICE = National Institute for Health and Care Excellence; PSS = Personal Social Services; QALY = quality-adjusted life-year; UK = United Kingdom

4.2.2 Model structure

A 4-state Markov cohort model was used in the economic analysis. The model was developed in Microsoft Excel™. The model structure consists of four mutually exclusive health states; "event-free (EF)", "locoregional recurrence (LR)", "distant metastasis (DM)", and "death". All patients begin in the "EF" health state. Movement through the model is determined by transition probabilities estimated using patient-level data from KEYNOTE-522 and KEYNOTE-355 (RCT of pembrolizumab + chemotherapy (paclitaxel or nab-paclitaxel or carboplatin/gemcitabine combination) versus chemotherapy for advanced/metastatic PD-L1+ve TNBC) trials. Grade 3+ AEs and Grade 2 AEs diarrhoea and colitis were modelled on the background.

Figure 4.1 shows the model structure of the 4-state Markov cohort model.

Figure 4.1: Model structure



Based on Figure 8 of the CS
AE; adverse effect; CS = company submission

ERG comment: In the CS, it is stated that the model structure of a previous appraisal of pertuzumab for the neoadjuvant treatment in HER2-positive BC (TA424) was used to inform the model structure of the current model.[39] The stage of disease in TA424 was identical to that in the decision problem for this appraisal i.e. locally advanced, inflammatory, or early-stage with a high risk of recurrence. However, the current model is a simplified version of the model structure used in TA424, excluding remission from LR and no differentiation between not progressed and progressed metastatic patients in the DM state. In clarification question B3b, the ERG requested a scenario analysis based on the same model structure as used in TA424.[5] However, the company did not provide the scenario in the response, which was justified by the fact that the clinical data from KEYNOTE-522 do not support the modelling structure used in TA424.[6]

The ERG acknowledges that the model structure of TA424 has its limitations, and it would be complex to use the exact same model structure for the current submission with the available KEYNOTE data. However, the ERG is concerned about the fact that the model: a) does not include the option for remission of LR; and b) does not differentiate between not progressed and progressed DM.

a) In TA424, patients moved from the 'LR' state through tunnel states to the 'remission' health state. The tunnel states (12 months) were used to 'hold' patients in the LR state for a certain duration before progressing to the remission state. In the 'LR' state patients received further treatment with pertuzumab. After completing the treatment, patients were assumed to be in remission and transitioned to the 'remission' health state. Similar to the current model, patient could progress from the remission health state (i.e., after a first LR) to the metastatic not progressed or death states, i.e., a second LR event was not possible. The company justifies the exclusion of a 'remission' health state based on the fact that the 'remission' state from TA424 in fact resembles the LR state in the current submission. The company argues that there are three reasons for this deviation from TA424. First, the NeoSphere trial - which informed TA424 - explored complete response (pCR) as the primary outcome, while the KEYNOTE-522 explored pCR and EFS as primary outcomes.[39, 40] Second, in contrast to TA424, subsequent retreatment with therapy at LR was not allowed in the KEYNOTE-522 trial design. Finally, the

LR health state in TA424 did not allow for patients to move to the death state, which may have led to an overestimation of the QALYs. The current model avoids this unrealistic assumption. The ERG acknowledges the differences between TA424 and the current submission and agrees with the company that the introduction of a remission state is not ideal, as it would increase the model's complexity by introducing multiple tunnel states to the model. However, assuming patients with LR cannot experience remission does simply not reflect clinical practice. Though the company assumes no further treatment effect in the LR state (i.e., transition probabilities to DM and death are treatment independent), the current model assumes that patients remain in the LR state until progression to metastatic disease or death, and therefore patients accrue health utilities and costs related to LR for the remaining time in this state. As patients in the placebo arm have a relatively higher probability to move from the event-free state to the LR state (because of relatively lower EFS and a relatively higher proportion of events being LR (year 1: pembrolizumab [redacted], placebo [redacted] and year 2+: pembrolizumab [redacted], placebo [redacted]) compared with the pembrolizumab arm, the ERG concludes that exclusion of remission from LR may lead to overestimation of pembrolizumab's effect, underestimating the incremental cost effectiveness ratio (ICER). The ERG was not able to include a remission health state within the timeframe of this appraisal, and therefore the exact effect on the ICER is unclear.

b) Differentiating between not progressed and progressed metastatic patients is essential to correctly reflect clinical disease progression and CE, since mortality, costs, and QoL differ considerably between pre-progression and post-progression metastatic patients. In TA424, the model differentiates between a not-progressed metastasis state (first line (1L) treatment) and a progressed metastasis state (>second line (2L+) treatment) using the line of treatment as a proxy. In TA424, non-progressed patients were assumed to have the general population mortality and rate of progression to and death in the progressed metastasis state; >second line (2L+) were estimated based on a weighted average of treatments informed by CLEOPATRA (RCT of trastuzumab and docetaxel versus trastuzumab, docetaxel and Pertuzumab). The current model used one DM state (including both not-progressed and progressed patients) with OS based on patients who received 1L in the KEYNOTE-355. Within the DM health state, patients who receive 1L treatments (% based on the KEYNOTE-355 and expert opinion) were also assumed to receive 2L+ treatment for which a lump sum cost was included in the model based on a weighted average of patients receiving 2L, 3L, or 4L treatments in the cost-effectiveness model for pembrolizumab as 1L treatment in patients with metastatic TNBC (based on KEYNOTE-355 and being used in the NICE appraisal ID1546).[41] It should be noted that OS as estimated from KEYNOTE-355 is a function of death in the 1L state plus rate of progression to and death in the progressed metastasis state (>second line (2L+) and the costs for 1L and 2L+ have been included in the model. However, the company did not account separately for OS, costs, and QoL related to progressed patients (receiving 2L+) as opposed to not progressed (1L) patients. This potentially leads to under or overestimation of the ICER. The ERG was not able to include separate health states for not-progressed and progressed DM since the company was not able to provide the ERG with data on the progression status for patients with DM as this was not recorded in the KEYNOTE-522. Therefore, the exact effect of this remains unclear. The ERG believes that this creates considerable uncertainty in the model.

4.2.3 Population

The patient population included in the economic evaluation consisted of [redacted]

[redacted]. This definition is narrower than the population defined in the final scope issued by NICE, i.e., *"adults with previously untreated locally advanced, nonmetastatic triple-negative breast cancer"*.[3] The proposed marketing authorisation is pembrolizumab with chemotherapy for neoadjuvant and adjuvant treatment of untreated locally advanced non-metastatic TNBC. The main body of clinical evidence for pembrolizumab was derived from KEYNOTE-522 which included patients with untreated newly diagnosed, locally advanced, centrally confirmed TNBC and have an ECOG PS of 0 or 1, see Section 2.1 for more details.[2] The key baseline patient characteristics in the economic model are listed in Table 4.4 below.

Table 4.4: Key baseline patient characteristics used in the economic model

	Mean (SD)	Source
Starting age (year)	[redacted]	KEYNOTE-522[2]
Female weight (kg), mean	[redacted]	KEYNOTE-522[2]
Female weight (kg), standard deviation	[redacted]	KEYNOTE-522[2]
Body surface area (BSA; m^2), mean	[redacted]	KEYNOTE-522[2]
Body surface area (BSA; m^2), standard deviation	[redacted]	KEYNOTE-522[2]
Based on Table 31 of the CS and the company model BSA = body surface area; CS = company submission; SD = standard deviation		

ERG comment: The population in the economic evaluation is narrower than the population defined in the NICE final scope.[3] The company stated in its response to the request for clarification that 'high-risk TNBC' within KEYNOTE-522 is synonymous with 'locally advanced TNBC'. The ERG agree that the wording is comparable with the final NICE scope.

4.2.4 Interventions and comparators

The intervention considered in the CS was pembrolizumab in combination with standard neoadjuvant chemotherapy followed by adjuvant pembrolizumab as a single regimen, administered IV at a fixed dose of 200 mg over 30 minutes Q3W in the neoadjuvant and adjuvant phases. The neoadjuvant chemotherapy component was: carboplatin (AUC 5 Q3W or AUC 1.5 weekly on days 1, 8 and 15) and paclitaxel (80 mg/m^2 weekly on days 1, 8 and 15) followed by doxorubicin (60 mg/m^2 Q3W) or epirubicin (90 mg/m^2 Q3W) and cyclophosphamide (600mg/m^2 Q3W).

The comparators considered were standard neoadjuvant chemotherapy (as described above) and placebo as adjuvant therapy. The NICE scope listed the following comparators: Standard neoadjuvant/adjuvant chemotherapy without pembrolizumab. The standard neoadjuvant therapy recommended by NICE is: platinum added to an anthracycline-containing neoadjuvant chemotherapy regimen. For adjuvant treatment after surgery, NICE recommends offering a regimen that contains both a taxane and an anthracycline. Standard chemotherapy options used for neoadjuvant and adjuvant treatment of TNBC include doxorubicin, epirubicin, docetaxel, paclitaxel and carboplatin. The company stated that the exclusion of chemotherapy as adjuvant therapy reflects the current UK practice, where no active treatments are given after surgery.

The neoadjuvant and adjuvant therapy was continued until completion of study treatment (17 cycles of pembrolizumab/placebo), disease progression in the neoadjuvant phase or until recurrence (local or distance) after surgery, unacceptable adverse event(s) or physician's decision to withdraw treatment

ERG comment: For adjuvant treatment after surgery, NG101 recommends offering a regimen that contains both a taxane and an anthracycline.[42] Moreover, the CS stated that recent evidence has shown

that capecitabine in the adjuvant phase may improve disease survival and recurrence-free survival, see Section 2.3.[1]

The company stated in its response to the request for clarification that a taxane and anthracycline regimen for the treatment of early-stage BC is generally given either before or after surgery. For chemotherapy, neoadjuvant versus adjuvant administration of a taxane and anthracycline regimen is considered equivalent in terms of distant recurrence, BC mortality or death from any cause for BC patients. Moreover, the company stated that in common clinical practice, a patient would not be treated with the same neoadjuvant and adjuvant chemotherapy regimen. Regarding the treatment with capecitabine as an adjuvant therapy, the company stated in its clarification response that the National Comprehensive Cancer Network (NCCN) guidelines were updated in 2017 to include adjuvant capecitabine as an option for patients with TNBC who do not achieve pCR after neoadjuvant chemotherapy.[43] Optional use of adjuvant capecitabine in patients who do not achieve pCR after neoadjuvant therapy may confound the EFS endpoint, due to the potential for imbalanced capecitabine use between the two treatment arms.

The ERG partly agrees with the company approach of excluding the taxane and anthracycline regimen from the adjuvant phase. Although the statement of the company that a taxane and anthracycline regimen for the treatment of early-stage breast cancer is generally given either before or after surgery, is not supported by any reference, this may still be common clinical practice. However, the use of a taxane and anthracycline regimen as adjuvant treatment could majorly change the EFS and therefore the ICER. The ERG does not agree with excluding capecitabine as adjuvant treatment because of the imbalance between the two arms. Excluding capecitabine as adjuvant therapy for patients who do not achieve pCR after adjuvant chemotherapy does not reflect the general practice and the used guidelines in the UK. The company stated that including capecitabine may increase EFS rate for patients with poor prognosis to 74%. This would majorly change the ICER.

The ERG considers additional scenarios where taxane and anthracycline are used as adjuvant therapy instead of neoadjuvant therapy would have been informative to see the impact on CE, as well as a scenario where capecitabine is used as an adjuvant therapy.

4.2.5 Perspective, time horizon and discounting

The analysis is performed from the NHS and Personal Social Services (PSS) perspective. Discount rates of 3.5% are applied to both costs and benefits. The model cycle length is 1 week with a time horizon of 51 years and a half-cycle correction is applied.

ERG comment: Perspective, time horizon and discounting are appropriate.

4.2.6 Treatment effectiveness and extrapolation

The primary source of treatment effectiveness for the intervention and comparator is KEYNOTE-522, a phase III RCT to evaluate pembrolizumab in combination with chemotherapy versus chemotherapy alone in the neoadjuvant phase followed by pembrolizumab monotherapy versus placebo in the adjuvant phase.[2] Patient level data of the KEYNOTE-522 trial was used to determine transition probabilities from the event-free and locoregional states. Due to immaturity of the KEYNOTE-522 OS data, transition probabilities from DM to death were based on the KEYNOTE-355 trial for those receiving 1L treatment for metastatic TNBC.[28] For patients who did not receive 1L metastatic TNBC treatment, OS data from the recent Surveillance, Epidemiology, and End Results (SEER) Medicaid database publication ('no treatment' subgroup) was used.[44] Time-on-treatment and relative dose intensity for the intervention and comparator were based on patient-level data from KEYNOTE-522.

4.2.6.1 Transition probabilities from event-free health state

The transition probabilities from the event-free health state were estimated based on the extrapolated EFS data, along with the probabilities of experiencing LR, DM, or death as the first EFS event in each treatment arm derived from the KEYNOTE-522 clinical trial (data cut-off date: 23 March 2021). Extrapolation of the EFS data beyond the trial duration to lifetime horizon was done using survival curve fitting, carried out in line with the NICE Decision Support Unit (DSU) guidelines.[45]

Statistical testing showed that the proportional hazard assumption for EFS did not hold. Therefore, standard parametric models were fitted to the patient level EFS data from the pembrolizumab arm and placebo arm in the KEYNOTE-522 trial separately. All standard parametric models (i.e., exponential, Weibull, Gompertz, log-logistic, log-normal and generalised gamma) were fitted to the patient level EFS data from the pembrolizumab arm and placebo arm in the KEYNOTE-522 trial to extrapolate the endpoints from the trial over a lifetime time horizon. Since the standard parametric distributions did not provide a good fit to the observed EFS data, two-phase parametric functions fit to the data were conducted. Hazard plots were used to identify potential cut-off points for two-phase models. Visual examination of the cumulative hazard plot suggested week 50 as a potential turning point of the EFS curves in both treatment arms. Hazard plots also suggested week 43 and 68 as turning points for the hazard function. Chow statistical tests showed two additional turning points, 93 and 109 weeks. From these five cut-off points, week 50 was used in the base case as it provides plausible visual fit and has a good balance of robust Kaplan-Meier (KM) data used to directly calculate transition probabilities in the first phase whilst enough data remaining can be used to fit a parametric curve in the second phase (week 50 to life-time horizon). Other cut-off points were included in the economic model.

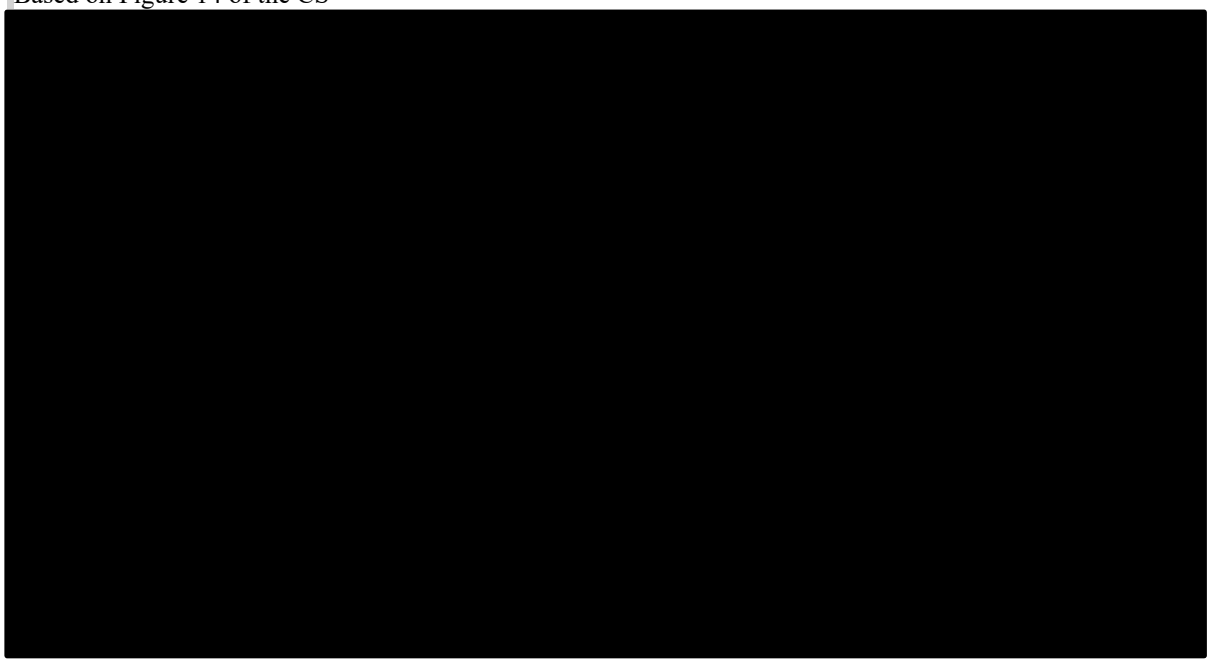
All parametric models were assessed against the Akaike information criterion (AIC) and Bayesian information criterion (BIC) criterion. AIC/BIC statistics, in combination with visual inspection and clinical plausibility based on expert opinion, were used to identify the best-fitted parametric distribution from week 50 onward. As the proportional hazards assumption did not hold, individual distributions were fitted for pembrolizumab and placebo EFS. In addition, the company chose these distributions to be of a different type between the two arms. The company argued that, given the unique mode of action of immunotherapy, pembrolizumab and placebo could have different parametric extrapolations because the underlying hazard for the parametric curves does not need to be the same.

Both AIC/BIC and visual inspection suggested generalised gamma was the best fit for the pembrolizumab arm. A log-normal distribution was suggested as the second-best option and was explored in scenario analysis. For the placebo arm, AIC/BIC statistics were lowest for the Gompertz distribution with log-normal distribution ranked second. However, the Gompertz distribution is associated with a flat tail potentially leading to overestimation of long-term EFS, which suggests an implausible extrapolation. Clinical experts and visual inspection of the curves confirmed the use of log-normal distribution in the base case analysis. Generalised gamma distribution was also suggested as plausible option and was explored in a scenario analysis.

Figure 4.2 shows the modelled and observed EFS extrapolation for the pembrolizumab (generalised gamma distribution) and placebo (log-normal distribution) arm from KEYNOTE-522.

Figure 4.2: Modelled versus observed EFS for pembrolizumab and placebo arm from KEYNOTE-522

Based on Figure 14 of the CS[1]



CS = company submission; EFS = event-free survival; KM = Kaplan-Meier.

The estimation of the transition probabilities from the event-free health state to LR, DM, or death were estimated using Gray's method considering competing risks. Competing risk analysis of the time to first EFS event was used to determine the distribution between the EFS event being LR, DM, or death. Within each cycle, the cause-specific probability of each transition (i.e., EF to LR, EF to DM, and EF to death) was calculated based on the estimated probability of an EFS event, and the probability that the EFS event being LR, DM or death (Table 4.5). The probability of EF to death was constrained by the general population mortality, adjusted for the transition probabilities from EF to LR and EF to DM.

Table 4.5: Probability of the first EFS event being LR, DM, or Death.

Treatment arm	Year 1			Year 2+		
	% LR	% DM	% Death	% LR	% DM	% Death
Pembrolizumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Placebo	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Based on Table 36 of CS[1]
CS = company submission; DM = distant metastasis; LR = locoregional recurrence; % = percentage

The predicted cumulative incidence of EF to LR, EF to DM, and EF to death were validated with the observed cumulative incidence from the KEYNOTE-522 trial. Based on Figures 15 and 16 as well as Tables 37 and 38 of the CS, the company concludes that the modelled cumulative incidence rates are comparable to the observed data.[1]

4.2.6.2 Transition probabilities from locoregional recurrence health state

The transition probabilities of LR to DM and LR to death were estimated based on the pooled data from the two treatment arms from the KEYNOTE-522 trial. Parametric models were fitted to the time from LR to DM or death, and the exponential distribution was found to be the best fit. When asked for statistics of the fit in the clarification phase, the company responded that the selection of the exponential

parametric distribution selected to model LR to DM or death was not based in isolation to the AIC/BIC statistics, but also on visual fit to the observed KM curve alongside balanced assessment of clinical plausibility of long-term predictions generated by each of the alternative parametric models.[6] The company stated that the few number of events which have taken place from which extrapolations are based could make the AIC/BIC statistics unreliable and therefore rankings based on AIC/BIC may change as more data become available. Whilst the exponential model yields the highest AIC/BIC statistics the difference versus the lowest average AIC/BIC produced by the log-normal model was small (6 to 7 points). The company decided to choose the exponential distribution because they considered it to better fit the tail of the KM-curve, despite the fact that it would overestimate OS for the observed period. See also Figure 4.3.

Figure 4.3: Long term parametric extrapolations using the combined KEYNOTE-522 arms time from LR to DM or Death

Based on Figure 4 of the response to the request for clarification[6]
DM = distant metastasis; LR = locoregional recurrence; OS = overall survival

The company assumed constant transition probabilities from the LR state. The transition probabilities to DM and death were calculated based on the transition probabilities of LR to DM *or* death, and the proportions of DM and death respectively, which were all obtained from the KEYNOTE-522 trial. The probability of LR to death was constrained by the general population mortality, adjusted for the transition probability from LR to DM.

4.2.6.3 Transition probabilities from the distant metastasis health state

In the DM state it was assumed that a proportion of patients would receive 1L treatment for metastatic disease. This proportion was obtained from the KEYNOTE-522 trial and was [redacted] for the pembrolizumab arm and [redacted] for the placebo arm.[2] Because of the current immaturity of the KEYNOTE-522 OS data, data from KEYNOTE-355 were used to estimate transition probabilities from DM to death.[28] KEYNOTE-522 data were used in a scenario.

The base case analysis assumed that patients could not be rechallenged with pembrolizumab but could receive other IOs in the DM setting 2 years post initiation of neoadjuvant treatment. This assumption is explored in two scenarios; one where pembrolizumab rechallenge, or treatment with another IO, is possible for the PD-L1 positive population 2 years post initiation of neoadjuvant treatment, and another where patients cannot receive any IOs and would receive a mix of non-IO chemotherapies. Patients in the first two scenarios who relapse within 2 years of neoadjuvant treatment initiation will be managed as in this last 'IO ineligible' scenario. Based on KEYNOTE-355, the proportion of PD-L1 positive patients was estimated at [redacted].

The base case treatment mix of each of the above scenarios was obtained from UK market research and clinical expert input, see Table 4.6 for details on treatment mix per scenario.

Table 4.6: Treatment mix of 1L metastatic TNBC used in the model

Type of 1L treatment	Pembrolizumab + chemotherapy			Chemotherapy
	Pembrolizumab rechallenge for PD-L1 positive	IO-eligible (pembrolizumab ineligible)	IO-ineligible	IO-eligible (pembro ineligible)[#]
Pembrolizumab + taxanes (paclitaxel or nab-paclictaxel)	[redacted]	[redacted]	[redacted]	[redacted]
Paclitaxel	[redacted]	[redacted]	[redacted]	[redacted]
Carboplatin (or containing regimens)	[redacted]	[redacted]	[redacted]	[redacted]
Carboplatin + paclitaxel	[redacted]	[redacted]	[redacted]	[redacted]
Gemcitabine + carboplatin	[redacted]	[redacted]	[redacted]	[redacted]
Atezolizumab + Nab-paclitaxel[*]	[redacted]	[redacted]	[redacted]	[redacted]
Capcitabine	[redacted]	[redacted]	[redacted]	[redacted]

Based on Table 42 of the CS and company model[1]

[*] assumes PD-L1 SP132 positive as per Impassion130 study; [#] See point f) in the ERG comment below

CS = company submission; 1L = first line; IO = immune oncology; N/A = not applicable; TNBC = triple-negtaive breast cancer

Mean OS in the DM state was estimated as a weighted average of OS for patients who received 1L treatments and OS for patients who did not receive 1L treatments. The transition probability from DM to death was then calculated by fitting an exponential curve to this mean OS and taking the coefficient of this fitted curve as the constant death rate, over the entire time horizon of the model.

The mean OS for patients who did receive 1L treatments was based on the pembrolizumab metastatic TNBC model, with HRs from an NMA (Appendix M of the CS) applied for carboplatin and atezolizumab + nab-paclitaxel.[4, 41] As can be seen in Table 4.7, the company stated that the NMA HRs were versus taxanes. However, the full NMA report provided with the clarification letter response

shows that the comparison was with nab-paclitaxel only. This was because networks were constructed according to subgroups that would be suitable for each of the investigator choice compactors in KEYNOTE 355, i.e., paclitaxel, nab-paclitaxel, and gemcitabine/carboplatin. Therefore, from these sources, it is unclear how this HR versus any taxane was estimated and how valid the estimate is when applied to survival with any taxane. The factual accuracy check (FAC) stated that: *"the studies informing the carboplatin comparison and link carboplatin into the rest of the network, only contain nab-paclitaxel. However, data from the pooled KEYNOTE-355 taxane data were used considering that the AC have previously concluded on taxane efficacy equivalence during prior HTAs"*. The company provided no clarification regarding the comparison with atezolizumab + nab-paclitaxel, but presumably this was also via the pooled KEYNOTE-355 taxane data as opposed to those for nab-paclitaxel only. The ERG considers that, despite the claim that there is equivalence between taxanes, the fact that the KEYNOTE 355 trial was stratified by investigator choice including taxane type (paclitaxel or nab-paclitaxel), which enabled the NMA to also be structured by subgroup according to investigator choice, means that the most appropriate KEYNOTE 355 data source for comparison with carboplatin or atezolizumab-paclitaxel is that for nab-paclitaxel only.

Time on treatment for each of the 1L treatments was derived in a similar way as OS according to the CS. That is, it was based on the pembrolizumab metastatic TNBC model. No further details on this are available.

The study by Aly et al. 2019 used to obtain OS for patients not treated with 1L treatments contained a sample of elderly mBC patients who were on average 79 years of age when they entered the study.[44] The company stated in its response to clarification that this high age should not bias estimates of DM survival since the metastases would likely be the leading cause of death even in high age.

Table 4.7 represents the OS estimates for the different 1L treatment options and Table 4.8 shows the resulting transition probability as used in the model for the base case and in the scenario using KEYNOTE-522 OS data.

Table 4.7: Mean OS by 1L metastatic TNBC treatment

Type of 1L treatment	Mean OS (weeks)	Source/method
Pembrolizumab + taxanes (paclitaxel or nab-paclictaxel)	[redacted]	Taken directly from KEYNOTE-355 1L mTNBC model
Paclitaxel	[redacted]	Taken directly from KEYNOTE-355 1L mTNBC model for taxanes (paclitaxel plus nab-paclitaxel) pooled arm in line with previous NICE assumptions
Carboplatin (or containing regimens)	[redacted]	HR estimated from NMA. Applied OS HR of carboplatin versus taxanes (paclitaxel/nab-paclitaxel) in mTNBC model
Carboplatin + paclitaxel	[redacted]	Assumed equal to gemcitabine + carboplatin arm of KEYNOTE-355 1L mTNBC model
Gemcitabine + carboplatin	[redacted]	Taken directly from KEYNOTE-355 1L mTNBC model
Atezolizumab + Nab-paclitaxel *	[redacted]	HR estimated from NMA. Applied OS HR of atezolizumab + nab-paclitaxel versus taxanes (paclitaxel/nab-paclitaxel) from KEYNOTE-355 1L mTNBC model

Type of 1L treatment	Mean OS (weeks)	Source/method
Capcitabine	[redacted]	Assumed equal to taxanes arm of KEYNOTE-355 1L mTNBC model
No 1L treatment	[redacted]	SEER Medicare, 'no treatment' group[44]

Based on Table 43 of the CS and company model[1]
1L = first line; CS = company submission; HR = hazard ratio; mTNBC = metastatic triple-negative breast cancer; NICE = National Institute for Health and Care Excellence; NMA= network meta-analysis; OS = overall survival

Table 4.8: Transition probabilities used in base case and scenarios

Treatment arm	Eligibility for IOs	Weighted mean OS (weeks)	Transition probability (weekly) from DM to death
Based on KEYNOTE-355 data and NMA			
Pembrolizumab	IO-eligible*	[redacted]	[redacted]
Pembrolizumab	Pembrolizumab rechallenge-eligible	[redacted]	[redacted]
Pembrolizumab	IO ineligible	[redacted]	[redacted]
Placebo	IO-eligible	[redacted]	[redacted]
Based on KEYNOTE-522 data			
Pembrolizumab	N/A^	[redacted]	[redacted]
Placebo	N/A^	[redacted]	[redacted]

Based on Table 44 and 45 of the CS[1]
* IO-eligible assumed in base case for the pembrolizumab arm
^ in the scenario using KEYNOTE-522 data, OS was not based on treatment mix but taken as observed and therefore the scenarios for IO eligibility do not apply
CS = company submission; DM = distant metastasis; N/A = not applicable; NMA = network meta-analysis; OS= overall survival

ERG comment: The main concerns of the ERG relate to: a) the use of differential distributions to extrapolate EFS; b) the difference between EFS gain obtained in the observed period versus EFS gain the extrapolated period; c) the use of constant transitions from LR and DM states; d) the use of the 50-week cut-off point for the EFS curve fitting; e) the use of KEYNOTE-355 as base case for the DM survival; f) no option to receive pembrolizumab as 1L treatment for patients in the placebo arm; and g) adjustment of general mortality in the formula for EF and LR to death by subtracting transitions to other states.

a) The ERG is satisfied that the company has followed the general approach to survival analysis and extrapolation of individual participant data recommended by NICE DSU TSD 14. However, the company used different distributions for the curve fitting for extrapolation of the EFS data. The TSD recommends that where parametric models are fitted separately to individual treatment arms it is sensible to use the same 'type' of model.[45] The use of different types of distributions should be justified using clinical expert judgement, biological plausibility, and robust statistical analysis.

   Therefore, in clarification question B8, the ERG asked for a clear explanation why different distributions in this case would be justifiable.[5] The company justified the use of different distributions based on the argument that the unique mode of action of immunotherapy (with or without chemotherapy) is not comparable to chemotherapy alone; therefore, the underlying hazard

assumption for the parametric curve does not need to be the same.[6] The company argues this has been observed across a number of metastatic and adjuvant submissions with IO agents to date—although they do not mention which—and that clinicians have noted that IO therapies used in neoadjuvant/adjuvant setting may have an effect of improving 'Immune surveillance'. Furthermore, the company explains clinical plausibility of different parametric models was discussed during an advisory board meeting. In response to clarification question B8, the company mentions that clinical experts *"based on the unique mode of action of IO therapies as well as the characteristics of patients with early TNBC disease, clinical experts noted that they would expect EFS to start to plateau across both treatment arms since most recurrences occur within the first 3 to 5 years and that pembrolizumab + chemotherapy EFS would sit above that of placebo"*.[6] The ERG does not consider this the same as the clinical experts confirming that different distributions are clinically plausible. Based on AIC and BIC values, generalised gamma was the best option in the pembrolizumab arm and Gompertz in the placebo arm. However, in both arms lognormal was the second-best option, which does not suggest strong evidence for different distributions. The company mentions in its response to clarification question B8 that the statistical fit was validated using real-world data, however there is only real-world data available for the placebo arm (and not for the pembrolizumab arm) and therefore this validation says nothing about the justification for the use of different distributions.

b) Related to the above issue, the company model and the extrapolations implemented result in a substantial gain in EFS which is mostly obtained in the unobserved part of the time horizon. It is important to take the realism of the extrapolated marginal gain into consideration when selecting the best model as an unrealistic marginal gain would create bias in the economic analysis. To evaluate the realism of the post-extrapolation survival gain, the 'rule-of-thumb' from Tremblay et al., 2015 can be used, stating that the ratio of the marginal relative difference in the extrapolated period (post cut-off) divided by the number of months post-cut-off should not be higher than the ratio of marginal difference on the number of months in the pre-extrapolation period.[46] In other words, the average "rate of survival gain" per month between treatments should be equal or inferior in the post-extrapolation period compared to the pre-extrapolation period. In the current model using different distributions for the extrapolation of EFS, the pre-extrapolation (up to week 205, based on KM data of KEYNOTE-522) rate of survival gain is 0.2367, while the post-extrapolation (from week 206) rate is 0.3340, suggesting lack of realism of the extrapolated marginal gain according to the rule-of-thumb. This is also seen in the model: chancing the time horizon from 51 years to a short-term horizon (e.g. 5 years, which reflects the period for which KM data of the KEYNOTE-522 is available) causes a considerable increase in the ICER. The ERG believes this is a major uncertainty in the model.

   Taken this and the issues discussed under point a) into account, the ERG is not convinced that there is a strong enough justification for using different distributions for the extrapolation of EFS based on the information provided in the CS. Therefore, the ERG uses lognormal distributions (second-best option) in both arms in its base case analysis, and additionally conducted several scenarios to explore the effect of different distributions.

c) The company assumed transition probabilities to move to the DM state (from LR) and to the death state (from LR and DM) to be constant over the entire time horizon of the model. According to page 84 of the CS, this was necessary because of the memoryless feature of the Markov model. The company stated it would be reasonable to assume a constant transition probability since an exponential distribution provided the best fit to the LR survival.[1] For DM, the transition to death was based on the constant hazard assumption without further explanation. No justification based on

clinical plausibility was provided though, also not in response to question B10b of the response to request for clarification.[6] Moreover, from the response to request for clarification to this question, it also became apparent that the exponential distribution did not actually provide the best fit for the LR survival – as almost all other parametric distributions resulted in lower AIC and BIC (although differences were small).[6] The ERG is concerned that oversimplifying assumptions for these transitions, which are mostly relevant to the placebo arm as relatively more patients in the placebo end up in LR and DM, will distort incremental CE while uncertainty around this issue is not captured in the sensitivity analyses.

d) The ERG agrees that the KEYNOTE-522 is the best available source for the extrapolation of EFS data. In accordance with the NICE DSU TSD 21, the company has explored the hazard plots for turning points, which suggested a turning point in week 43 for the pembrolizumab + chemotherapy arm and week 68 for the chemotherapy arm.[47] Visual inspection of the cumulative hazard plots suggested a divergence of curves with a potential turning point at approximately 50 weeks (Figure 10 of the CS).[1] Additionally, the Chow test suggested week 93 and 109 as potential turning points. For the base case analysis, the company used the week 50 cut-off point, justified by the fact that it provides plausible visual fit and a good balance of robust KM data to be used directly in the first phase and enough remaining data to be used to fit a parametric curve in the second phase. However, this does not explain why the week-50 cut-off is preferred over the other cut-off points with sufficient data left to inform survival extrapolations (i.e., week 43 and week 68). Although in response to clarification question B6 the company mentions that the other cut-off points are included in the model, they were not included as scenarios in the CS.[6]

e) The ERG considers the use of KEYNOTE-355 data as base case for the DM survival to be a potential source of bias. Although the company argues that KEYNOTE-355 is to be preferred over KEYNOTE-522 data because KEYNOTE-522 data are not sufficiently mature in the DM state, there are quite substantial differences in observed survival between these two studies (see Table 4.8), which raises doubts about comparability of the populations and therefore on appropriateness of using KEYNOTE-355 OS data for this appraisal. The ERG therefore prefers to use the company's scenario using KEYNOTE-522 data as a base case.

f) In the base case company model, patients in the chemotherapy arm are assumed to not receive pembrolizumab when they metastasize, and so all patients that are IO-eligible receive atezolizumab (see also Table 4.7 above where the treatment mix in 1L metastatic mTNBC is specified – for placebo, the pembrolizumab is set to N/A). The ERG believes this to be an error in the model and corrected for this in its base case.

g) The probabilities of EF and LR to death were constrained by the general population mortality. However, the general mortality in the formula for the transition probability from EF to death was adjusted for the transitions from EF to LR and DM. Similar, the general mortality in the formula for the transition probability from LR to death was adjusted for the transition from LR to DM. The ERG believes the adjustment of the general mortality by subtracting transition probabilities from the EF and LR state to states other than death to be an error in the model and corrected for this in its base case.

4.2.7 Adverse events

The main source of evidence on incidence of treatment-related AEs used for intervention and comparator is the KEYNOTE-522 trial. The model considers all-cause Grade 3+ AEs with an incidence

of ≥5%. Additionally, two Grade 2+ AEs, diarrhoea and colitis, were included in the economic model as these were deemed as clinically relevant.

ERG comment: The main concern of the ERG relates to the inclusion of Grade 2+ AEs colitis and diarrhoea because these were deemed clinically relevant. In response to clarification question B16a, in which the ERG asked for clarification why these Grade 2 AEs were deemed clinically relevant, the company explained that these specific AEs were included in addition to Grade 3+ AEs as they expect these AEs to be associated with a high management cost (i.e. hospitalisation) and HRQoL burden, and to ensure consistency with previous NICE appraisals for IO therapies.[6] The ERG agrees that the inclusion of Grade 2 diarrhoea was indeed in line with other appraisals (e.g. ID1546, pembrolizumab + chemotherapy for untreated, locally recurrent unresectable or metastatic TNBC).[38] However, it remains unclear how the clinical relevance of these AEs was defined.

4.2.8 Health-related quality of life

Health state utility values were estimated for the following health states: EF, LR, and DM and were treatment independent.

4.2.8.1 Health-related quality of life data identified in the review

According to the CS, the SLR identified only one study (Huang et al. 2020) reported EQ-5D-5L utility values for metastatic TNBC.[1, 48] However, no studies reported EQ-5D derived utilities for eBC. Therefore, utility values from the KEYNOTE-522 trial were used to inform the economic model.

4.2.8.2 Health state utility values

In the absence of studies from the SLR (see Section 4.2.8.1), the primary source for the HSUVs was the KEYNOTE-522 trial. HRQoL was collected in KEYNOTE-522 using the EQ-5D-5L questionnaire. In the neoadjuvant phase, the questionnaire was administered on day 1 of cycle 1 of treatment 1 (carboplatin + paclitaxel with or without pembrolizumab) and on day 1 of cycles 1 and 4 of treatment 2 (doxorubicin/epirubicin + cyclophosphamide with or without pembrolizumab). In the adjuvant phase, the questionnaire was administered on day 1 of cycles 1, 5 and 9. Assessments were also conducted at the early discontinuation visits and for long-term follow-up visits every 12 months for 2 years or until progressed disease (PD). The analysis of the EQ-5D-5L data was based on FAS population (pembrolizumab + chemotherapy n=762 and chemotherapy alone n=384) and compliance remained high throughout the trial ([redacted]).

EQ-5D-5L scores were retrieved for: 1) by health state and by treatment status; 2) by treatment status within EF state and by treatment arm, and 3) by AE status within EF on treatment period and by treatment arm. No statistically significant difference in utilities between the two arms was found. Therefore, health state utilities used in the economic model were estimated based on the pooled treatment arm set by health state and for the EF state by treatment status (on or off treatment). For the EFS on treatment health state the utility was only of patients without Grade3+ AE, to avoid double counting as Grade 3+ AE-related disutilities were included in the model separately.

As per the NICE reference case, the EQ-5D-5L data were mapped back to the 3L tool using crosswalk method by van Hout et al., 2012.[49] The 3L value set was used in the base case analysis. The 5L value set was explored in scenario analysis.

A summary of all utility values used in the CEA is provided in Table 4.9.

Table 4.9: Health state utility EQ-5D-3L values used in the model

Health state	Utility value (95% CI)	Reference
Event-free, on treatment	[redacted]	KEYNOTE-522[2] and UK crosswalk tariffs[49]
Event-free, off treatment	[redacted]	
Local recurrence	[redacted]	
Distant metastasis	[redacted]	
Based on base case analysis from Table 52 of the CS[1] CI = confidence interval; CS = company submission		

4.2.8.3 Disutility values

All Grade 3+ events with an incidence of ≥5% were included in the economic model. Additionally, Grade 2 events diarrhoea and colitis were included. Disutilities associated with the AEs were implemented in the model by calculating a QALY loss which was the product of the disutility and the duration of the AE and applied in the first cycle of the model. Grade 3+ AE-related disutilities were obtained from KEYNOTE-522 patient-level data. The disutility associated with AEs from the pooled utility analysis was estimated at [redacted] (standard error (SE): [redacted]). The Grade 3+ AE disutility was also applied to the Grade 2+ AEs included in the model.

An age-related disutility was applied using calculations from Ara and Brazier et al., 2010.[50]

ERG comment: The main concerns of the ERG relate to: a) the use of pooled health state utilities; and b) the relatively low utility value for DM health state.

a) The company used the pooled health state utilities in the base case analysis, as there was no statistically significant or clinically meaningful difference between the treatment arms. However, the HSUVs were slightly but consistently lower in the pembrolizumab + chemotherapy arm compared to the chemotherapy arm (Table 7 in Appendix N of the CS).[4] In response to clarification question B18a, the company explains this may be in part due to the more complex treatment regimen since pembrolizumab is an add on therapy to the neoadjuvant current standard of care (SoC).[6] As such patients randomised in this arm experience more AEs which subsequently may reduce utility scores. However, the company argues that treatment related HRQoL decrement associated with pembrolizumab is applied through AE disutility (modelled as a one-off QALY decrement). The effect of using treatment-related health state utilities was explored in a scenario analysis but showed to have a minimal effect on the ICER. Therefore, the ERG does not consider this is a major issue.
b) The utility for the health state DM is relatively low ([redacted]) compared to other studies. As mentioned in Appendix H of the CS, one other study (Huang et al. 2020) assessing EQ-5D in metastatic TNBC patients is available, which examined the EQ-5D-3L data collected from patients enrolled in KEYNOTE-119 (previously treated metastatic TNBC patients).[48] The mean utility for progression-free and progressed patients was 0.715 and 0.606, respectively. The difference (0.104) between the two was considered clinically meaningful. The utility value used for the 'DM' health state in the model is very comparable to that of the progressed metastatic TNBC patients in the KEYNOTE-119 trial, however, the 'DM' state in the current model includes both progressed and not-progressed patients. The company justifies the use of the relatively low utility value from the KEYNOTE-522 by the fact that as the NICE reference case stipulates a preference for HRQoL data collected alongside the pivotal trial to be used for the decision problem when these are available.[51] However, the company also reported that the number of EQ-5D questionnaires from patients what have experienced DM is very limited ([redacted] across both treatments). Moreover, it is unclear whether the relatively low utility value may be related to a relatively high proportion of patients with PD within

the DM state, because the KEYNOTE-522 does not record the progression status for patients with DM and the company argues that utility data stratified by subsequent treatment line as a proxy for progression status was not possible as these data remain immature and considering the already limited EQ-5D data available, these analyses would not be meaningful if conducted. However, the company did acknowledge that the small number of questionnaires available at DM setting may explain why utility values are relatively lower than reported elsewhere in literature. Therefore, in response to clarification question B19c, the company conducted two scenarios to explore the effect of using alternative data sources to test the impact of the DM utility estimate on the ICER using utility values from: 1) KEYNOTE-355 ([redacted] weighted average based on total predicted LYs gained during pre-progression and post-progression of chemotherapy arm); and 2) KEYNOTE-119 (0.715, pre-progression utility value). In both scenarios the ICER increased, however, the difference was marginal.

4.2.9 Resources and costs

The cost categories included in the model were treatment acquisition costs for intervention, comparator and subsequent treatments, medical costs (treatment administration, disease management, costs of LR and DM states, costs of surgery, and costs of terminal care and end of life), and costs of managing AEs.

Unit prices were based on the NHS reference prices, British National Formulary (BNF), Personal Social Services Research Unit (PSSRU) and the Monthly Index of Medical Specialities (MIMS). Costs were inflated to the current price year using the Hospital and Community Health Services (HCHS) index published by PSSRU where necessary.[52]

4.2.8.1 Resource use and costs data identified in the review

According to the CS, the SLR did not identify any studies reporting UK relevant resource use and cost information for the population of interest.

4.2.8.2 Treatment costs

As per the anticipated license, the model uses a 200 mg fixed dose of pembrolizumab administered as a 30-minute IV infusion Q3W, in combination with chemotherapy (carboplatin + paclitaxel, followed by doxorubicin/epirubicin + cyclophosphamide) in the neoadjuvant phase for 8 cycles, and pembrolizumab monotherapy in the adjuvant phase for 9 cycles.

The intervention and comparator drug acquisition costs used in the model were reported in Table 53 of the CS.[1] The list price for pembrolizumab used is £2,630 per 100 mg/4 ml vial. A confidential PAS is in place.

No vial sharing was included in the base case model, but this assumption was varied in a scenario where vial sharing was allowed.

Relative dose intensity as reflected in the pembrolizumab arm of KEYNOTE-522 was applied to the drug acquisition costs. The detailed dosing schedule, relative dose intensity, and treatment allocation, can be found in Table 54 of the CS.[1]

KEYNOTE-522 patient level data were used to estimate time to end of neoadjuvant treatment, time to end of surgery, and time to end of treatment course. The proportion of patients on neoadjuvant treatment was derived directly from the time to end of neoadjuvant treatment KM curve. The proportion of patients on adjuvant treatment was derived by subtracting the survival function for time to end of surgery from the KM curve for end of treatment course. All three curves are displayed in Figure 4.4

below which was provided by the company in its response to request for clarification.[6] The company also explained in its response to clarification that there is a 2 to 6 week wait time between end of neoadjuvant treatment and surgery, and that resource use associated with the EF state was applied to patients waiting for surgery. In its response to the request for clarification, the company stated that there was a waiting time of about 2 to 6 weeks after the last cycle of the neoadjuvant phase.[6]

Figure 4.4: Time to end of neoadjuvant treatment/surgery/treatment course

Based on Figure 5 of the response to request for clarification[6]

Costs of drug acquisition and administration for subsequent treatments were applied as one-time costs upon entry into the DM state. A proportion of patients entering the DM state were assumed to receive an active 1L treatment. KEYNOTE-355 was used to estimate these 1L treatment costs in the base case, while KEYNOTE-522 was used in a scenario. Patients who received 1L treatments were also assumed to receive subsequent lines (2L, 3L, and 4L). The weighted average costs for each treatment arm was calculated by multiplying the proportion of patients who received 1L treatments (Table 40 of the CS) by the weighted average costs of patients who receive 1L treatments derived from KEYNOTE-355.[1]

Administration costs for intervention/comparators and subsequent treatments were included in the model, depending on complexity and treatment type. Detailed administration costs were presented in Tables 59 and 60 of the CS.[1]

4.2.8.3 Health state costs

Health state costs consisted of disease management costs and included oncologist visits, visits to the general practitioner (GP), clinical nurse specialist and community nurse contacts, imaging (mammogram, computed tomography (CT) and magnetic resonance imaging (MRI) scans), and laboratory monitoring. The frequency for these types of resource use was based on TA424 for the EF health state,[39] TA569 for LR,[53] and ID1546 and TA639 for the DM health state.[38, 54] In addition to these costs which were applied weekly in the model, there were also additional disease management costs for the EF state whilst on treatment (based on assumption from clinical expert opinion), and a one-off cost for the LR and DM states. A one-off cost was also applied at the point of death. Table 4.10 reflects these various health state costs applied.

Table 4.10: Health state costs

Health state	Costs	Reference	Justification
Weekly health state costs			
Event free whilst on treatment (year 0-1) pembrolizumab arm	£81.99	Assumption from clinical expert opinion	Annually 17 oncologist visits, 17 nurse specialist visits, and 25 FBC
Event free whilst on treatment (year 0-1) placebo arm	£38.06	As above adjusted for chemo arm	Annually 8 oncologist visits, 8 nurse specialist visits, and 16 FBC
Event free (year 1-3)	£7.55	TA424 Table 90[39]	Annually 2 oncologist visits, 2 GP visits, 1 mammogram
Event free (year 4-5)	£3.89	TA424 Table 90[39]	Annually 1 oncologist visit, 1 GP visit, 1 mammogram
Event free (year 6-10)	£0.75	TA424 Table 90[39]	Annually 1 GP visit
Locoregional recurrence	£14.50	TA569 Table 42[53]	Annually 2 oncologist visits, 1 mammogram, 2 CT scans, 1 MRI
Distant metastasis	£69.00	ID1546 Table 65[38] and TA 639 Table 64[54]	Annually 12 oncologist visits, 1 GP visit, 4 CT scans, 12 nurse specialist visits, 3 community nurse visits, 17 FBC
One-off costs			
Locoregional recurrence	£474.76	Assumed equal to DM state	Equal to DM

Health state	Costs	Reference	Justification
Distant metastasis	£474.76	ID1546 Table 64[38] and TA639 Table 63[54]	1 oncologist visit, 1 CT scan, 1 FBC, 1 MRI
End of life	£8,347.03	Georghiou & Bardsley et al. 2014 inflated to 2020 value[55]	Including district nurse, nursing and residential care, hospice care, and nursing service
Based on Tables 62, 63, 64 and 65 of the CS[1] CS = company submission; CT = computed tomography; DM = distant metastasis; FBC = full blood count; GP = general practitioner; MRI = magnetic resonance imaging; TA = technology appraisal			

4.2.8.4 Surgery costs and adverse event costs

Surgery costs were applied within the model as a one-time cost calculated based on the unit costs of surgery and the proportion of patients receiving surgery in each arm. A weighted average of £5,823.04 was derived from the unit costs of breast surgery from the NHS reference costs.[56] The proportion of patients receiving surgery was obtained from the KEYNOTE-522 trial and was [redacted] and [redacted] for the pembrolizumab and placebo arm, respectively.

Modelled AEs and its corresponding incidence were presented in Section 4.2.7. The resource use and costing approach was based on previous technology appraisals in IO. See Table 4.11 below for details.

Table 4.11: Unit costs associated with AE management

Type of AE	AE cost	Justification
Grade 3+ AEs		
Neutropenia	£635.68	Costing as per TA519[57]
Neutrophil count decreased	£635.68	Equal to Neutropenia as in TA519[57]
Anaemia	£762.54	Costing as per TA519[57]
Febrile neutropenia	£3,580.80	Costing as per TA737 approach[57]
White blood cell count decreased	£635.68	Equal to Neutropenia as in TA519[57]
AAT increased	£0.00	Costing as per TA684 (previously TA558); Assumption of zero cost for laboratory abnormalities; (already considered under health-state management costs)[58]
Other AEs		
Diarrhoea (Grade 2+)	£2,166.42	Costing as per TA581 approach[59]
Colitis (Grade 2+)	£2,166.42	Equal to Diarrhoea (Grade 2+) as in TA581[59]
Based on Table 66 of the CS[1] AAT = alanine aminotransferase increased; AE = adverse effect; CS = company submission; TA = technology appraisal		

ERG comment: The main concerns of the ERG relate to: a) treatment costs for the DM state may be overestimated; and b) waiting time for surgery seems longer than anticipated

a) As already discussed in Section 4.2.6, modelled survival in the DM state was based on KEYNOTE-355. Given the differences in observed survival between KEYNOTE-355 and KEYNOTE-522, the ERG believes that KEYNOTE-522 would be a more accurate source to inform

DM. However, in the company scenario where KEYNOTE-522 was used for survival in DM, costs (and time on treatment) were still based on KEYNOTE-355 as time on treatment in DM is not a model parameter but assumed to be fixed and costs for DM treatment were implemented as a one-off cost. Therefore, in the scenario where KEYNOTE-522 data are used to inform survival in DM, costs for treatment in DM would be overestimated since patients will have shorter survival while costs are not adjusted. The ERG considers that even when KEYNOTE-355 data would be appropriate, the approach to estimating 1L treatment costs as a one-off in the DM state is not sufficiently precise given the rather substantial impact these costs have on the ICER. An additional comment to this is that the proportion of patients assumed to receive 1L treatment in the DM state was derived from KEYNOTE-522 data in the company base case and was higher for the placebo arm, driving up costs. No clinical or other rationale was provided for the difference in proportion of patients receiving 1L treatment.

b) The ERG asked the company in the clarification phase whether there was a waiting time for surgery after neoadjuvant treatment, and if so, how would patients be managed in the meantime.[5] The company, in its response to the request for clarification, said that indeed according to the KEYNOTE-522 protocol, patients underwent definitive surgery 2 to 6 weeks after the last cycle of the neoadjuvant phase, and thus there was a waiting time before surgery. According to the time on treatment curves however (see Figure 4.4 above), the waiting time appears to be much longer, at least 10 weeks. Although the ERG is puzzled by this apparent difference between protocol and reality, there may not be a large impact on CE as the difference is seen in both arms and in the model the patients waiting for surgery were assumed to have resource use as associated with the EF state.

5. COST EFFECTIVENESS RESULTS

5.1 Company's cost effectiveness results

The CS base case CE results indicated that pembrolizumab is both more effective (incremental QALYs of [redacted]) and more costly (additional costs of [redacted]) than current care amounting to an ICER of £5,940 per QALY gained (Table 5.1)

Table 5.1: Company's deterministic base case results using pembrolizumab PAS price

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER versus baseline (£/QALY)
Placebo arm	[redacted]	13.82	[redacted]	-	-	-
Pembrolizumab arm	[redacted]	16.89	[redacted]	[redacted]	[redacted]	£5,940

ICER = incremental cost-effectiveness ratio; LYG = life years gained; PAS = Patient Access Scheme; QALY = quality-adjusted life year

Overall, the technology is modelled to affect QALYs by:

- An increase in EF survival at a relatively high utility
- A relatively lower utility in the LR and DM states where proportionally more chemotherapy patients reside

Overall, the technology is modelled to affect costs by:

- Its higher treatment acquisition price compared to chemotherapy alone in both the neoadjuvant and the adjuvant phase
- The higher metastatic (one-off) treatment costs for the chemotherapy arm

ERG comment: The main concerns of the ERG relate to: a) the extrapolated gains being substantially higher than observed gains; and b) the metastatic treatment costs being more than three times higher in the chemotherapy arm compared to pembrolizumab.

a) The issue of the extrapolated versus observed gains was already discussed earlier in Section 4.2.6 (see ERG comment b)
b) The base case model results show that for the chemotherapy arm, almost half of the total costs consisted of metastatic treatment costs. The ERG considers this to be unlikely and may be a result of the potentially biased and imprecise way of estimating the one-off metastatic treatment costs as discussed in Section 4.2.9 (see ERG comment a)

5.2 Company's sensitivity analyses

The company performed and presented the results of probabilistic sensitivity analyses (PSAs), deterministic sensitivity analyses (DSAs) as well as scenario analyses. The PSA with 1,000 iterations resulted in a higher ICER. The results of PSA analysis are presented in Table 5.2.

Table 5.2: Company's probabilistic base case results using pembrolizumab PAS price

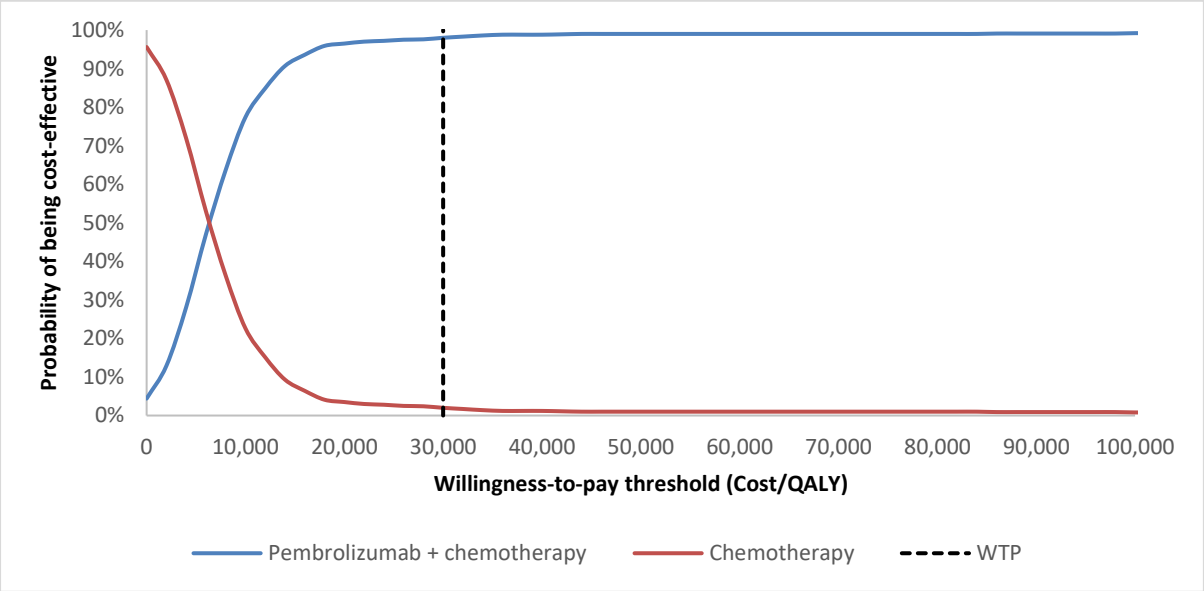
Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER versus baseline (£/QALY)
Placebo arm	[redacted]	13.79	[redacted]	-	-	

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER versus baseline (£/QALY)
Pembrolizumab arm	[redacted]	16.72	[redacted]	[redacted]	[redacted]	£6,128

ICER = incremental cost-effectiveness ratio; LYG = life years gained; PAS = patient access scheme; QALY = quality-adjusted life year

The cost effectiveness acceptability curve (CEAC) in the untreated analysis showed that pembrolizumab had a 98% probability of being cost effective at willingness to pay thresholds of £30,000. The CEAC is presented in Figure 5.1.

Figure 5.1: Company's CEAC with pembrolizumab PAS price

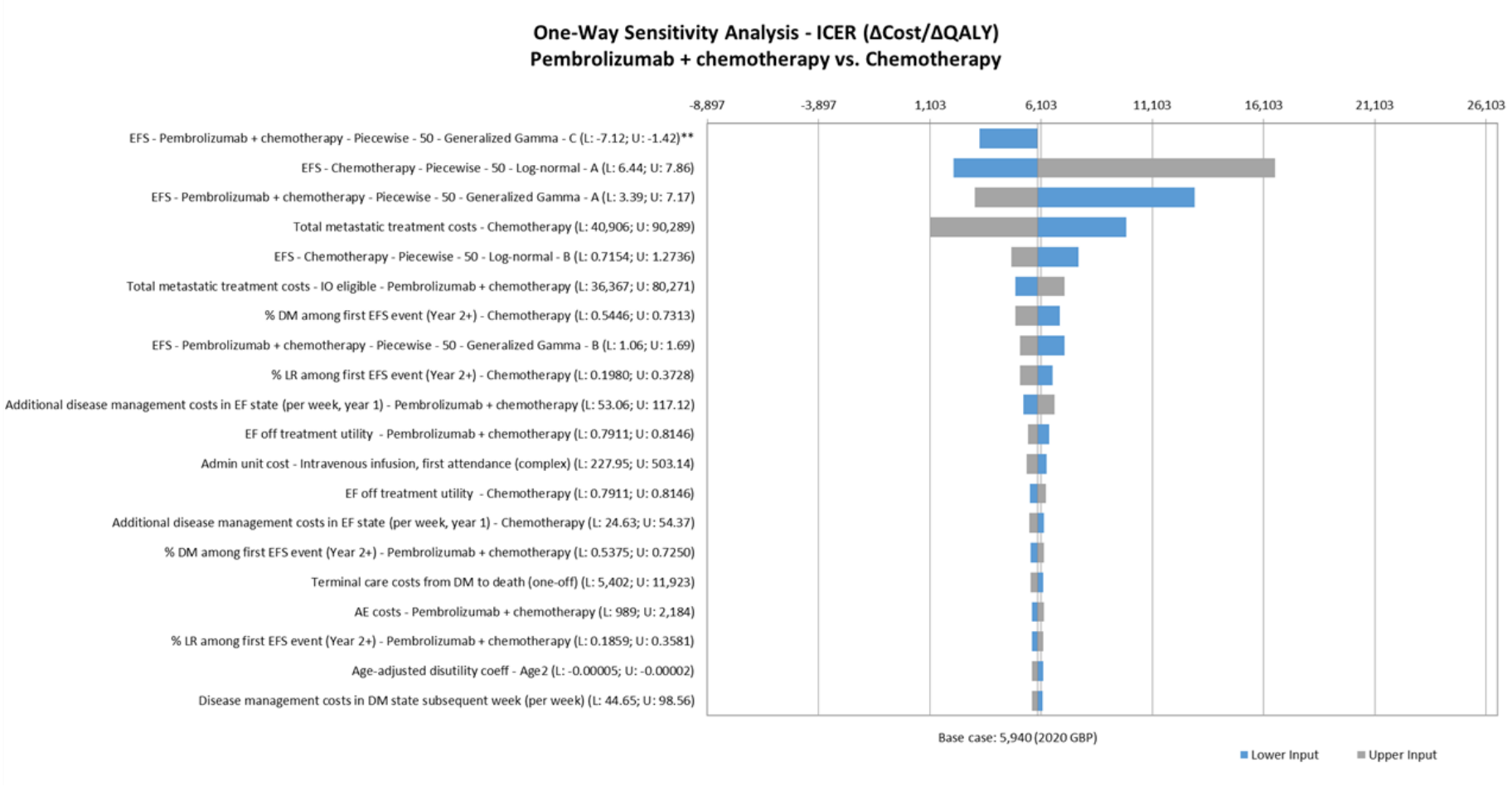


Based on Figure 22 of the CS[1]

CEAC = cost effectiveness acceptability curve; CS = company submission; PAS = patient access scheme; QALY = quality-adjusted life years; WTP = willingness-to-pay

The DSA was performed to investigate key drivers of the base case results. Each input parameter was varied to its respective upper or lower bound and the deterministic results for the model recorded. The base case parameter values were varied across their 95% CI where possible. The results of the DSA are presented in Figure 5.2 below. The inputs that have most impact on the ICERs are those related to parameters linked to EFS extrapolations followed by metastatic treatment costs. CS scenarios that have a substantial impact on the ICER are the scenarios varying the distributions for the extrapolation of EFS, and the scenario with a limited time horizon (20 years).

Figure 5.2: Company's tornado diagram for the 20 most sensitive parameters with pembrolizumab PAS price



**Upper limit parameter pembrolizumab arm is dominated i.e. more costly and less effective; therefore an ICER statistic cannot be presented for the tornado diagram

Based on Figure 23 of the CS[1]

CS = company submission; DM = distant metastasis; EFS = event-free survival; ICER = incremental cost effectiveness analyis; IO = immune oncology; LR = locoregional recurrence; PAS = patient access scheme; QALY = quality-adjusted life year

5.3 Model validation and face validity check

5.3.1 Face validity assessment

Efficacy (EFS) outcomes from KEYNOTE-522 were compared with the modelled EFS curves produced from the economic model. This was possible until a 3-year time horizon, as there were no observed data beyond this point. The company concluded that the modelled EFS curves matched well with the observed EFS curves (Tables 74 and 75 and Figure 24 of the CS).[1]

In addition, OS as modelled in KEYNOTE-355 was compared to OS as observed in KEYNOTE-522. Again, comparison was only possible up until the 3-year time point. The company concluded that the modelled and observed curves matched well. There was slightly more deviation between modelled and observed outcomes than for EFS though.

5.3.2 Technical verification

No details on technical verification were provided, other than a statement that clinical expert opinion was sought to validate certain aspects of the model, including internal review and quality control for model inconsistencies and errors performed.

5.3.3 Comparisons with other technology appraisals

No comparison with other technology appraisals was reported.

5.3.4 Comparison with external data used to develop the economic model

No comparison with external data used to develop the model was reported.

5.3.5 Comparison with external data not used to develop the economic model

For EFS, two external sources were identified as sources of validation for the placebo modelled EFS, a randomized phase II trial reported by Sikov et al. 2019 and a retrospective cohort by Walsh et al. 2019[60, 61] The company concluded that the placebo arm EFS curve matched well with the DFS curve from Walsh et al. 2019 and was reasonably close to the EFS curve from Sikov at al. 2019.

For OS, the same two studies by Sikov et al. 2019 and Walsh et al. 2019 were identified for validation purposes.[60, 61] The company concluded that given reasonable visual alignment, the model produces robust estimates of OS for the chemotherapy arm. The company also noted that using KEYNOTE-522 data provided slightly better visual alignment than using KEYNOTE-355 OS (which was used in the company base case to inform OS).

In the absence of clinical or real-world long-term data for early-stage TNBC patients receiving pembrolizumab, plausibility of modelled long-term EFS and OS was validated with clinical experts.[29]

ERG comment: The main concerns of the ERG relate to the absence of explicit clinical validation (using landmark estimates of survival, for instance) and the questionable appropriateness of validating KEYNOTE-355 model OS with KEYNOTE-522 OS given that these are different populations.

6. EVIDENCE REVIEW GROUP'S ADDITIONAL ANALYSES

6.1 Exploratory and sensitivity analyses undertaken by the ERG

Table 6.1 summarises the key issues related to the CE categorised according to the sources of uncertainty as defined by Grimm et al. 2020:[62]

- Transparency (e.g. lack of clarity in presentation, description, or justification)
- Methods (e.g. violation of best research practices, existing guidelines, or the reference case)
- Imprecision (e.g. particularly wide CIs, small sample sizes, or immaturity of data)
- Bias & indirectness (e.g. there is a mismatch between the decision problem and evidence used to inform it in terms of population, intervention/comparator and/or outcomes considered)
- Unavailability (e.g. lack of data or insight)

Identifying the source of uncertainty can help determine what course of action can be taken (i.e., whether additional clarifications, evidence and/or analyses might help to resolve the key issue). Moreover, Table 6.1 lists suggested alternative approaches, expected effects on the CE, whether it is reflected in the ERG base case as well as additional evidence or analyses that might help to resolve the key issues.

Based on all considerations in the preceding Sections of this ERG report, the ERG defined a new base case. This base case included multiple adjustments to the original base case presented in the previous sections. These adjustments made by the ERG form the ERG base case and were subdivided into three categories (derived from Kaltenthaler et al. 2016):[63]

- Fixing errors (FE; correcting the model where the company's submitted model was unequivocally wrong)
- Fixing violations (FV; correcting the model where the ERG considered that the NICE reference case, scope or best practice had not been adhered to)
- Matters of judgement (MJ; amending the model where the ERG considers that reasonable alternative assumptions are preferred)

6.1.1 ERG base case

Adjustments made by the ERG, to derive the ERG base case (using the CS base case as starting point) are listed below. Table 6.2 shows how individual adjustments impact the results plus the combined effect of all abovementioned adjustments simultaneously, resulting in the ERG base case. The 'FE' adjustments were combined, and the other ERG analyses were performed also incorporating these 'FE' adjustments given the ERG considered that the 'FE' adjustments corrected unequivocally wrong issues.

6.1.1.1 Fixing errors

1. Enable pembrolizumab 1L treatment in DM state for IO-eligible patients in the placebo arm
2. Adjustment to formulas correcting for general population mortality

6.1.1.2 Fixing violations

Not applicable.

6.1.1.3 Matters of judgement

1. Issue 3 (Section 3.2.1)
   Correction for efficacy of pembrolizumab adjusting for Europe versus rest of the world HR. The ERG implemented a simple fix to the efficacy in the model, assuming the HR to remain constant over time.
2. Issue 10 (Section 4.2.6)
   Use KEYNOTE-522 data to inform survival in DM state and alongside this adjust treatment costs according to the shorter survival.
3. Issue 8 (Section 4.2.6)
   Use lognormal distributions in EFS for both arms.

6.1.2 ERG exploratory scenario analyses

The ERG performed the following exploratory scenario analyses to explore the impact of alternative assumptions conditional on the ERG base case.

Exploratory scenario analyses

1. Limit time horizon to 5 years (similar to the observed period)
2. Set the cut-off of the piecewise model at 68 weeks instead of 50 weeks
3. Use generalised gamma distributions for EFS in both arms (Issue 8, Section 4.2.6)
4. Use lognormal distribution for pembrolizumab and generalized gamma distribution for placebo EFS (Issue 8, Section 4.2.6)
5. Adjust utility in DM health state to [redacted] based on KEYNOTE-355 (Issue 11, Section 4.2.8)
6. Adjust utility in DM health state to 0.715 based on KEYNOTE-119 (Issue 11, Section 4.2.8)

6.1.3 ERG subgroup analyses

No subgroup analyses were performed by the ERG.

Table 6.1: Overview of key issues related to the cost effectiveness (conditional on fixing errors highlighted in Section 5.1)

Key issue	Section	Source of uncertainty	Alternative approaches	Expected impact on ICER[a]	Resolved in ERG base case[b]	Required additional evidence or analyses
Model structure not including locoregional remission and no differentiation between pre-progression and post-progression distant metastatic patients.	4.2.2.	Unavailability	Structural model adjustment	+/-	No	May not be possible with available data
Modelled treatment effectiveness and extrapolation for EFS state likely overestimates effectiveness of pembrolizumab	4.2.6	Bias and indirectness; extrapolated part of the model generates most of the EFS gain compared to observed part	Change distributions	+	Partly in ERG analysis 4, and ERG scenarios 3 and 4	Mature data for better validation of long-term extrapolations
Constant transition probabilities from LR and DM states assumed without clinical justification	4.2.6	Unavailability & imprecision; lack of mature comparative data on OS	Use well-informed OS distributions	+/-	No	Mature LR and DM survival data, clinical justification
The use of KEYNOTE-355 data for DM survival may not be appropriate	4.2.6	Bias and indirectness; lack of mature comparative data observed in KEYNOTE-522	Use KEYNOTE-522 data for OS in DM	+/-	Partly in ERG analysis 3; however, KEYNOTE-522 not mature	Mature LR and DM survival data
Relatively low utility in the DM health state	4.2.8	Bias & indirectness as utility may not be appropriate to the population and health state in question	Explore impact of higher utility	+	Yes, in ERG scenarios 5 and 6, but utility is still an estimate	Not applicable

[a] Likely conservative assumptions (of the intervention versus all comparators) are indicated by '-'; while '+/-' indicates that the bias introduced by the issue is unclear to the ERG and '+' indicates that the ERG believes this issue likely induces bias in favour of the intervention versus at least one comparator; [b] Explored
ERG = Evidence Review Group; FE = Fixing errors; FV = fixing violations; MJ = matters of judgement; ICER = incremental cost effectiveness ratio

6.2 Impact on the ICER of additional clinical and economic analyses undertaken by the ERG

In Section 6.1 the ERG base case was presented, which was based on various changes compared to the company base case. Table 6.2 shows how individual changes impact the results plus the combined effect of all changes simultaneously. The exploratory scenario analyses are presented in Table 6.3 deterministically and in Table 6.4 probabilistically. These are all conditional on the ERG base case (except the scenarios where EFS distributions are varied – these override the base case distributions). The analyses numbers in Tables 6.2, 6.3 and 6.4 correspond to the numbers reported in Section 6.1. The CEAC of the ERG base case and the exploratory scenario analyses are presented in Figure 6.1 to 6.7 The submitted model file contains technical details on the analyses performed by the ERG (e.g. the "ERG" sheet provides an overview of the cells that were altered for each adjustment).

Table 6.2: Deterministic ERG base case

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
CS base case					
Pembrolizumab + chemotherapy	[redacted]	[redacted]			
Chemotherapy	[redacted]	[redacted]	[redacted]	[redacted]	£5,940
Fixing errors 1: Enable pembrolizumab 1L treatment in DM state for IO-eligible patients in the placebo arm					
Pembrolizumab + chemotherapy	[redacted]	[redacted]			
Chemotherapy	[redacted]	[redacted]	[redacted]	[redacted]	£9,346
Fixing errors 2: Adjustment to formulas correcting for general population mortality					
Pembrolizumab + chemotherapy	[redacted]	[redacted]			
Chemotherapy	[redacted]	[redacted]	[redacted]	[redacted]	£5,976
Matters of judgement 1: Correction for efficacy of pembrolizumab adjusting for Europe versus rest of the world hazard ratio					
Pembrolizumab + chemotherapy	[redacted]	[redacted]			
Chemotherapy	[redacted]	[redacted]	[redacted]	[redacted]	£7,801
Matters of judgement 2: Use KEYNOTE-522 data to inform survival in DM state and alongside this adjust treatment costs according to the shorter survival					
Pembrolizumab + chemotherapy	[redacted]	[redacted]			
Chemotherapy	[redacted]	[redacted]	[redacted]	[redacted]	£8,976
Matters of judgement 3: Use lognormal distributions in EFS for both arms					
Pembrolizumab + chemotherapy	[redacted]	[redacted]			
Chemotherapy	[redacted]	[redacted]	[redacted]	[redacted]	£16,444

1L = first line; CS = company submission; DM = distant metastasis; EFS = event-free survival; ERG = Evidence Review Group; ICER = incremental cost effectiveness ratio; IO = immune oncology; QALY = quality-adjusted life year

Table 6.3: Deterministic scenario analyses (conditional on ERG base case)

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
ERG base case					
Pembrolizumab + chemotherapy	[redacted]	[redacted]			
Chemotherapy	[redacted]	[redacted]	[redacted]	[redacted]	£43,621
Scenario 1: Limit time horizon to 5 years (similar to the observed period)					
Pembrolizumab + chemotherapy	[redacted]	[redacted]			
Chemotherapy	[redacted]	[redacted]	[redacted]	[redacted]	£397,435
Scenario 2: Set the cut-off of the piecewise model at 68 weeks instead of 50 weeks					
Pembrolizumab + chemotherapy	[redacted]	[redacted]			
Chemotherapy	[redacted]	[redacted]	[redacted]	[redacted]	£27,172
Scenario 3: Use generalized gamma distributions for EFS in both arms					
Pembrolizumab + chemotherapy	[redacted]	[redacted]			
Chemotherapy	[redacted]	[redacted]	[redacted]	[redacted]	£15,447
Scenario 4: Use lognormal distribution for pembrolizumab and generalized gamma distribution for placebo EFS					
Pembrolizumab + chemotherapy	[redacted]	[redacted]			
Chemotherapy	[redacted]	[redacted]	[redacted]	[redacted]	£53,592
Scenario 5: Adjust utility in DM health state based on KEYNOTE-355					
Pembrolizumab + chemotherapy	[redacted]	[redacted]			
Chemotherapy	[redacted]	[redacted]	[redacted]	[redacted]	£44,259
Scenario 6: Adjust utility in DM health state based on KEYNOTE-119					
Pembrolizumab + chemotherapy	[redacted]	[redacted]			
Chemotherapy	[redacted]	[redacted]	[redacted]	[redacted]	£44,362
ERG base case					
Pembrolizumab + chemotherapy	[redacted]	[redacted]			
Chemotherapy	[redacted]	[redacted]	[redacted]	[redacted]	£43,621

CS = company submission; EFS = event-free survival; ERG = Evidence Review Group; ICER = incremental cost effectiveness ratio; QALY = quality-adjusted life year

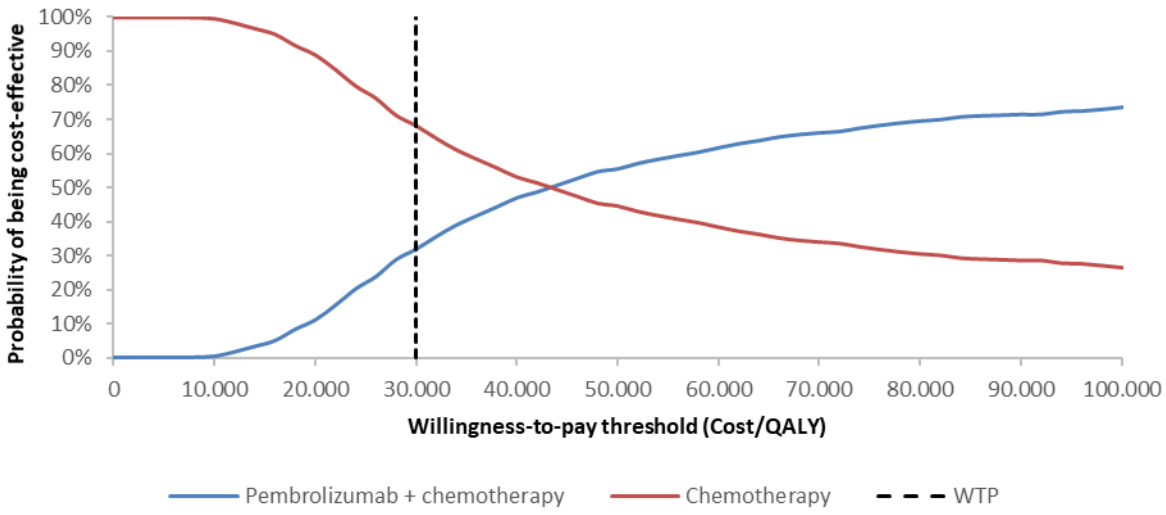
Table 6.4: Probabilistic scenario analyses (conditional on ERG base case)

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)	Prob-ability
ERG base case						
Pembrolizumab + chemotherapy	[redacted]	[redacted]				
Chemotherapy	[redacted]	[redacted]	[redacted]	[redacted]	£43,621	31.9%
Scenario 1: Limit time horizon to 5 years (similar to the observed period)						
Pembrolizumab + chemotherapy	[redacted]	[redacted]				
Chemotherapy	[redacted]	[redacted]	[redacted]	[redacted]	£381,768	0.0%
Scenario 2: Set the cut-off of the piecewise model at 68 weeks instead of 50 weeks*						
Pembrolizumab + chemotherapy	[redacted]	[redacted]				
Chemotherapy	[redacted]	[redacted]	[redacted]	[redacted]	£37,272	50.8%
Scenario 3: Use generalized gamma distributions for EFS in both arms						
Pembrolizumab + chemotherapy	[redacted]	[redacted]				
Chemotherapy	[redacted]	[redacted]	[redacted]	[redacted]	£16,697	79.0%
Scenario 4: Use lognormal distribution for pembrolizumab and generalized gamma distribution for placebo EFS						
Pembrolizumab + chemotherapy	[redacted]	[redacted]				
Chemotherapy	[redacted]	[redacted]	[redacted]	[redacted]	£58,421	28.1%
Scenario 5: Adjust utility in DM health state based on KEYNOTE-355						
Pembrolizumab + chemotherapy	[redacted]	[redacted]				
Chemotherapy	[redacted]	[redacted]	[redacted]	[redacted]	£44,568	31.4%
Scenario 6: Adjust utility in DM health state based on KEYNOTE-119						
Pembrolizumab + chemotherapy	[redacted]	[redacted]				
Chemotherapy	[redacted]	[redacted]	[redacted]	[redacted]	£44,685	31.4%

CS = company submission; DM = distant metastasis; EFS = event-free survival; ERG = Evidence Review Group; ICER = incremental cost effectiveness ratio; QALY = quality-adjusted life year

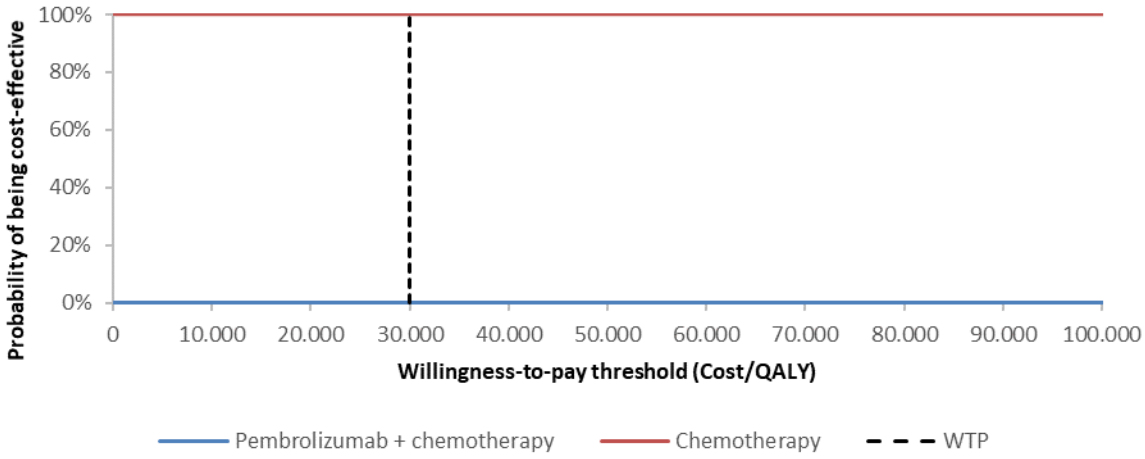
*Errors in approximately ten PSA runs. Errors were excluded from the analysis to obtain the results

Figure 6.1: Cost-effectiveness acceptability curve of pembrolizumab + chemotherapy versus chemotherapy based on ERG base case



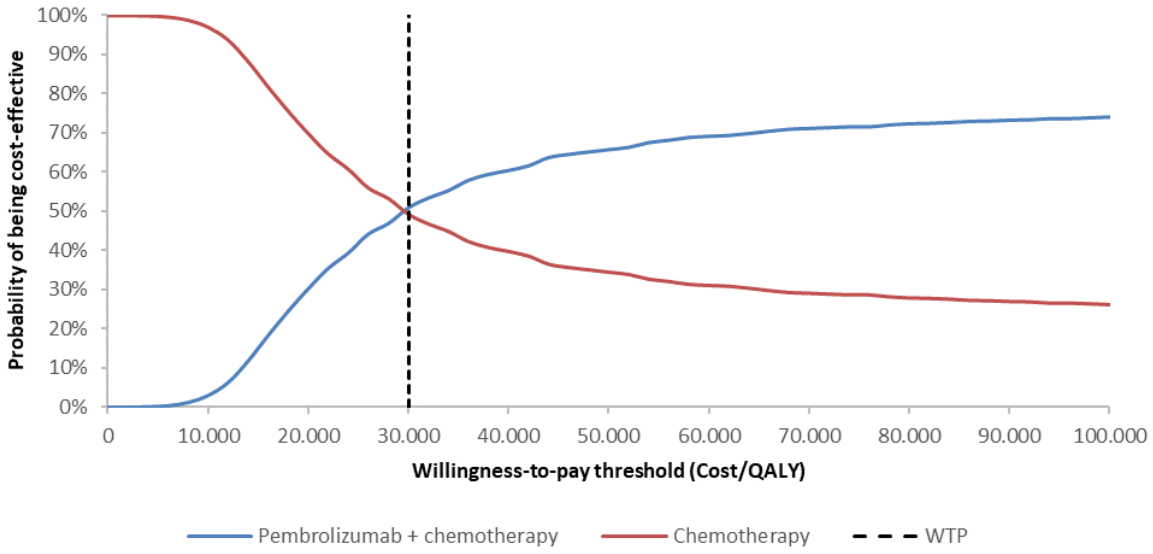
Based on the company model with ERG adjustments
ERG = Evidence Review Group; QALY = quality-adjusted life year; WTP = willingness-to-pay

Figure 6.2: Cost-effectiveness acceptability curve of pembrolizumab + chemotherapy versus chemotherapy based on scenario 1



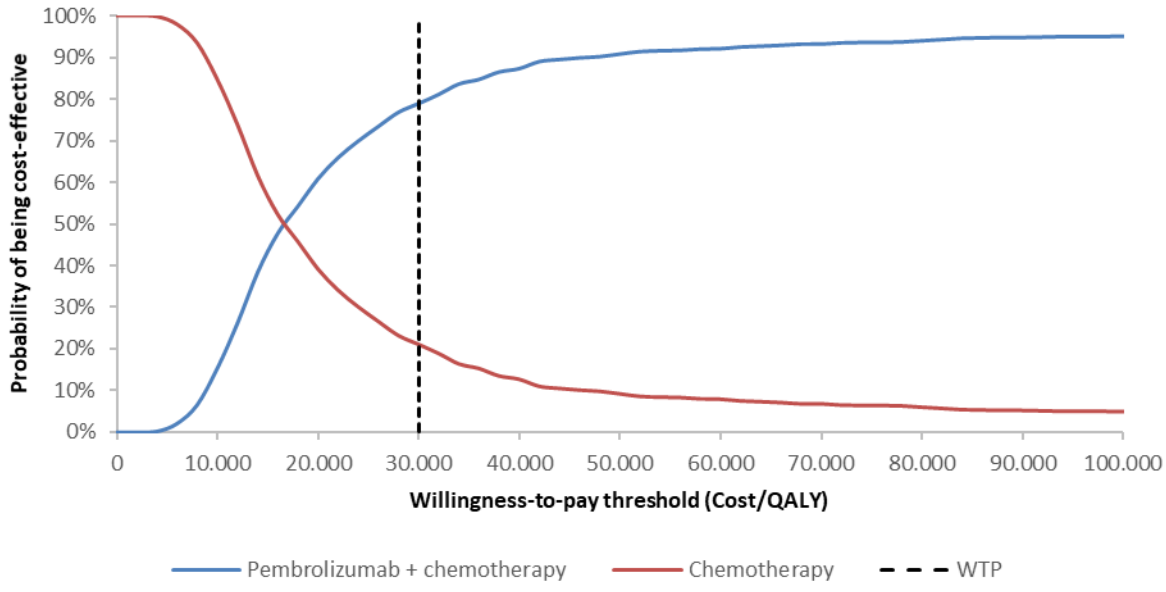
Based on the company model with ERG adjustments
ERG = Evidence Review Group; QALY = quality-adjusted life year; WTP = willingness-to-pay

Figure 6.3: Cost-effectiveness acceptability curve of pembrolizumab + chemotherapy versus chemotherapy based on scenario 2



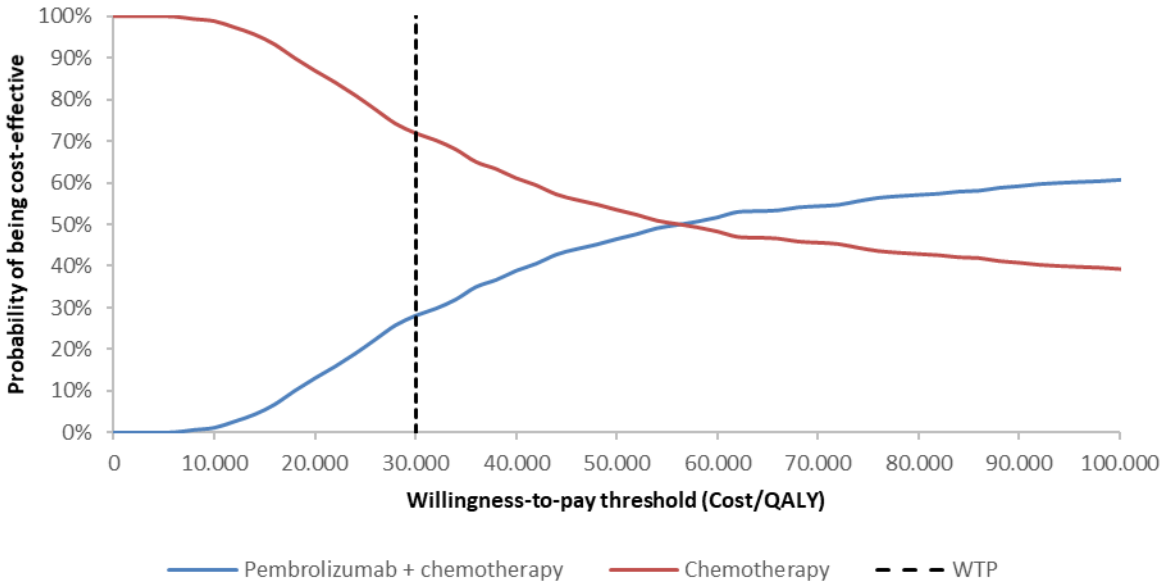
Based on the company model with ERG adjustments
ERG = Evidence Review Group; QALY = quality-adjusted life year; WTP = willingness-to-pay

Figure 6.4: Cost-effectiveness acceptability curve of pembrolizumab + chemotherapy versus chemotherapy based on scenario 3



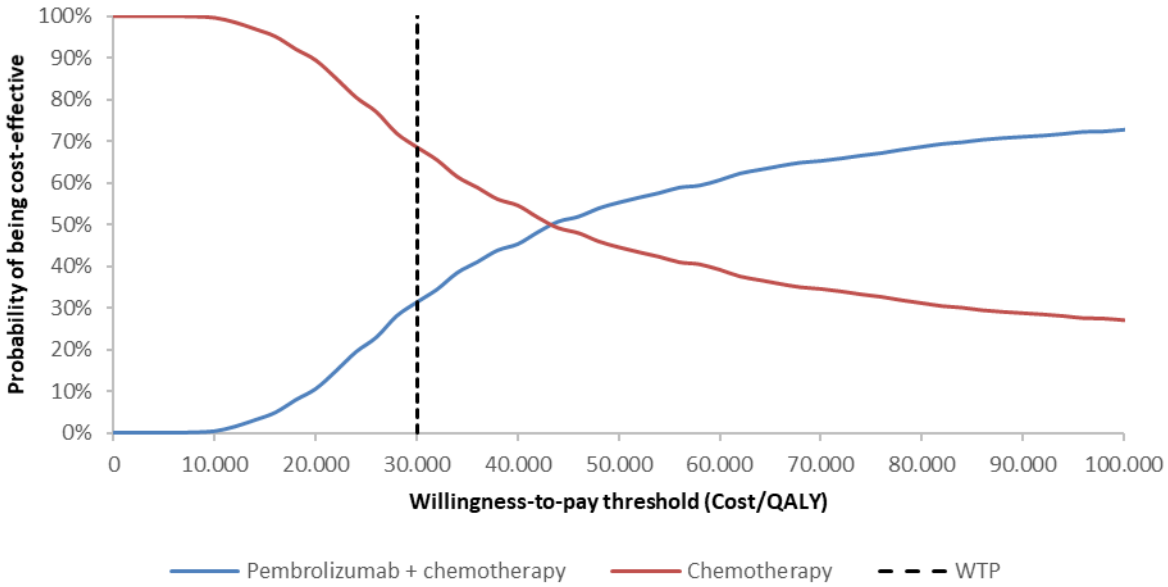
Based on the company model with ERG adjustments
ERG = Evidence Review Group; QALY = quality-adjusted life year; WTP = willingness-to-pay

Figure 6.5: Cost-effectiveness acceptability curve of pembrolizumab + chemotherapy versus chemotherapy based on scenario 4



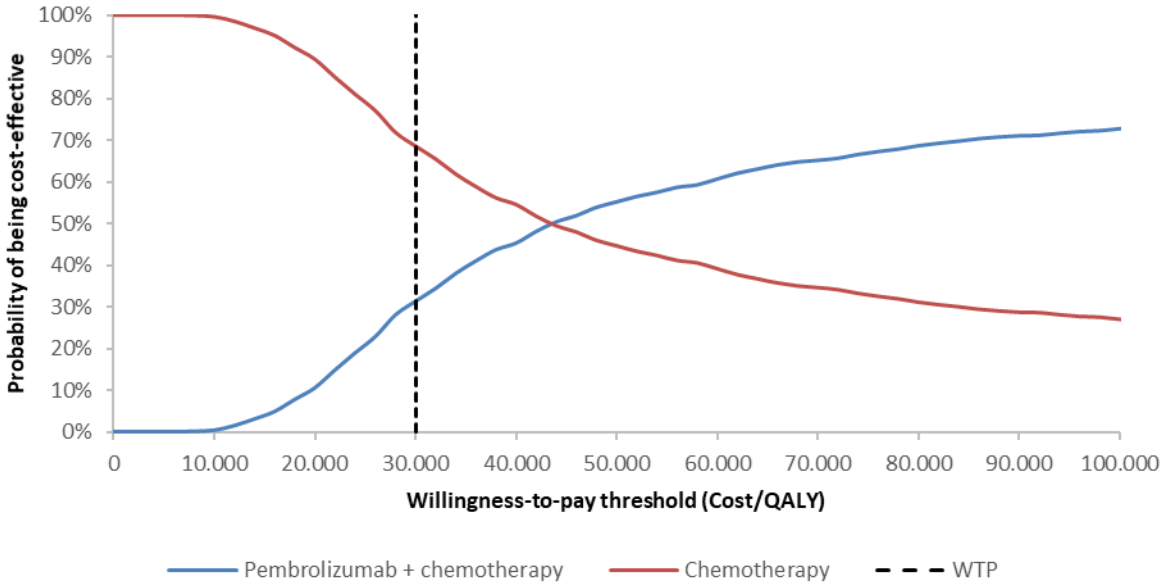
Based on the company model with ERG adjustments
ERG = Evidence Review Group; QALY = quality-adjusted life year; WTP = willingness-to-pay

Figure 6.6: Cost-effectiveness acceptability curve of pembrolizumab + chemotherapy versus chemotherapy based on scenario 5



Based on the company model with ERG adjustments
ERG = Evidence Review Group; QALY = quality-adjusted life year; WTP = willingness-to-pay

Figure 6.7: Cost-effectiveness acceptability curve of pembrolizumab + chemotherapy versus chemotherapy based on scenario 6



Based on the company model with ERG adjustments
ERG = Evidence Review Group; QALY = quality-adjusted life year; WTP = willingness-to-pay

6.3 ERG's preferred assumptions

The estimated ERG base case ICER (probabilistic), based on the ERG preferred assumptions highlighted in Section 6.1, was £43,621 per QALY gained. The probabilistic ERG base case analysis indicated a CE probability of 31.9% at WTP thresholds of £30,000 per QALY gained. The most influential adjustments were using lognormal distributions in EFS for both arms. The ICER increased most in the scenario analysis with alternative assumptions regarding the time horizon.

6.4 Conclusions of the cost effectiveness section

The company's CE estimates rest heavily on QALY gains in the extrapolated part of the model, while QALY gains in the observed part of the model were only very modest (Issue 8). The ERG adjusted the distributions used for EFS extrapolation in its base case but not all uncertainty caused by this issue may be resolved with this adjustment. As the model structure does not include separate health states for remission from LR and pre- and post-progression in the metastatic phase, it may not sufficiently capture relevant changes in HRQoL and costs in these states (Issue 7). The ERG could not resolve this issue in its analyses. Issue 3, the fact that pembrolizumab may not be as effective in the European population, has been addressed in the ERG model, but to properly explore the impact of regional difference in effectiveness, the model structure would need to be adapted more elaborately. Resolving Issue 2, the exclusion of a potentially relevant comparator, would also require structural changes and additional evidence which was not available.

Given that relatively more patients in the placebo arm reside in the locoregional and metastatic health states (because of the substantial EFS advantage modelled for pembrolizumab), costs and utilities in these states have an important impact on the ICER. However, the way the locoregional and metastatic health states were modelled was quite crude, with transition probabilities assumed constant over the full-time horizon of the model (Issue 9), and the metastatic health state being mostly informed by the KEYNOTE-355 data and model, with treatment costs calculated as one-off based on fixed treatment durations (Issue 10) and a relatively low utility value (Issue 11). Most of the uncertainty around these issues remains in the ERG analyses, although the ERG explores the impact of some assumptions in its scenarios.

In conclusion, with the current model, CE estimates of pembrolizumab + chemotherapy compared with chemotherapy alone are uncertain and likely are subject to bias. Although part of the issues were addressed, substantial uncertainty remains, especially on the long-term EFS benefit and on the costs and QALYs in the metastatic health state. Both are not supported by mature comparative data.

7. END-OF-LIFE

According to the CS, *"pembrolizumab plus chemotherapy followed by adjuvant pembrolizumab monotherapy does not meet the end-of-life criteria"*.[1]

8. REFERENCES

[1] Merck Sharp & Dohme. *Pembrolizumab with chemotherapy for neoadjuvant and adjuvant treatment of untreated locally advanced non-metastatic triple negative breast cancer [ID1500]. Document B. Company evidence submission. Single technology appraisal (STA)*: Merck Sharp & Dohme, 2021

[2] Merck Sharp & Dohme. *Clinical study report. Study of pembrolizumab (MK-3475) plus chemotherapy vs placebo plus chemotherapy as neoadjuvant therapy and pembrolizumab vs placebo as adjuvant therapy in participants with triple negative breast cancer (TNBC) (MK-3475-522/KEYNOTE-522) - Data on File*: Merck Sharp & Dohme, 2021

[3] National Institute for Health and Care Excellence. *Pembrolizumab with chemotherapy for neoadjuvant and adjuvant treatment of locally advanced non-metastatic triple-negative breast cancer. Final scope [Internet]*. London: National Institute for Health and Care Excellence, 2021 [accessed 21.3.22] Available from: https://www.nice.org.uk/guidance/gid-ta10399/documents/final-scope

[4] Merck Sharp & Dohme. *Pembrolizumab with chemotherapy for neoadjuvant and adjuvant treatment of untreated locally advanced non-metastatic triple negative breast cancer [ID1500]. Appendices. Company evidence submission. Single technology appraisal (STA)*: Merck Sharp & Dohme, 2021

[5] National Institute for Health and Care Excellence. *Pembrolizumab with chemotherapy for neoadjuvant and adjuvant treatment of locally advanced non-metastatic triple-negative breast cancer [ID1500]. Clarification letter*. London: NICE, 2022

[6] Merck Sharp & Dohme. *Pembrolizumab with chemotherapy for neoadjuvant and adjuvant treatment of untreated locally advanced non-metastatic triple negative breast cancer [ID1500]. Response to request for clarification from the ERG*: Merck Sharp & Dohme, 2022

[7] European Medicines Agency. *Keytruda: pembrolizumab. Opinion [Internet]*. Amsterdam, The Netherlands: European Medicines Agency, 2022 [accessed 4.5.22] Available from: https://www.ema.europa.eu/en/medicines/human/summaries-opinion/keytruda-8

[8] Early Breast Cancer Trialists' Collaborative Group (EBCTCG). Long-term outcomes for neoadjuvant versus adjuvant chemotherapy in early breast cancer: meta-analysis of individual patient data from ten randomised trials. *Lancet Oncol* 2018;19(1):27-39.

[9] U.S. Food & Drug Administration. FDA approves pembrolizumab for high-risk early-stage triple-negative breast cancer [Press release] [Internet]. U.S. Food & Drug Administration, 2021 [accessed 27.5.22]. Available from: https://www.fda.gov/drugs/resources-information-approved-drugs/fda-approves-pembrolizumab-high-risk-early-stage-triple-negative-breast-cancer

[10] McGowan J, Sampson M, Salzwedel DM, Cogo E, Foerster V, Lefebvre C. PRESS Peer Review of Electronic Search Strategies: 2015 Guideline statement. *J Clin Epidemiol* 2016;75:40-6.

[11] Canadian Agency for Drugs and Technologies in Health. *PRESS - Peer Review of Electronic Search Strategies: 2015 Guideline Explanation and Elaboration (PRESS E&E) [Internet]*. Ottawa: CADTH, 2016 [accessed 21.3.22] Available from: https://www.cadth.ca/resources/finding-evidence/press

[12] National Institute for Health and Care Excellence. *Single Technology Appraisal: company evidence submission template [Internet]*. London: NICE, 2015 [accessed 21.3.22] Available from:

http://www.nice.org.uk/Media/Default/About/what-we-do/NICE-guidance/NICE-technology-appraisals/specification-for-company-submission-of-evidence-2015-version.docx

[13] Lefebvre C, Glanville J, Briscoe S, Littlewood A, Marshall C, Metzendorf M-I, et al. Chapter 4: Searching for and selecting studies. In: Higgins JPT, Thomas J, Chandler J, Cumpston MS, Li T, Page MJ, et al., editors. *Cochrane handbook for systematic reviews of interventions version 6.3 (updated February 2022) [Internet]*: Cochrane, 2022 [accessed 21.3.22]. Available from: www.training.cochrane.org/handbook

[14] Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* 2021;372:n71.

[15] Page MJ, Moher D, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. PRISMA 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews. *BMJ* 2021;372:n160.

[16] Rethlefsen ML, Kirtley S, Waffenschmidt S, Ayala AP, Moher D, Page MJ, et al. PRISMA-S: an extension to the PRISMA statement for reporting literature searches in systematic reviews. *Syst Rev* 2021;10(1):39.

[17] Higgins J, Thomas J, Chandler J, Cumpston M, Li T, Page M, et al. *Cochrane handbook for systematic reviews of interventions version 6.3 (updated February 2022) [Internet]*: Cochrane, 2022 [accessed 21.3.22] Available from: https://training.cochrane.org/handbook

[18] Golder S, Peryer G, Loke YK. Overview: comprehensive and carefully constructed strategies are required when conducting searches for adverse effects data. *J Clin Epidemiol* 2019;113:36-43.

[19] Brenton JD, Carey LA, Ahmed AA, Caldas C. Molecular classification and molecular forecasting of breast cancer: ready for clinical application? *J Clin Oncol* 2005;29(29):7350-60.

[20] Higgins JPT, Savović J, Page MJ, Sterne JAC, on behalf of the RoB2 Development Group, eds. *Revised Cochrane risk-of-bias tool for randomized trials (RoB 2). 22 August 2019 [Internet]*, 2019 [accessed 21.3.22]. Available from: https://www.riskofbias.info/welcome/rob-2-0-tool/current-version-of-rob-2

[21] Gianni L, Mansutti M, Anton A, Calvo L, Bisagni G, Bermejo B, et al. Comparing neoadjuvant nab-paclitaxel vs paclitaxel both followed by anthracycline regimens in women with ERBB2/HER2-negative breast cancer-the Evaluating Treatment with Neoadjuvant Abraxane (ETNA) trial: a randomized phase 3 clinical trial. *JAMA Oncol* 2018;4(3):302-8.

[22] Untch M, Jackisch C, Schneeweiss A, Conrad B, Aktas B, Denkert C, et al. Nab-paclitaxel versus solvent-based paclitaxel in neoadjuvant chemotherapy for early breast cancer (GeparSepto-GBG 69): a randomised, phase 3 trial. *Lancet Oncol* 2016;17(3):345-56.

[23] Mittendorf EA, Zhang H, Barrios CH, Saji S, Jung KH, Hegg R, et al. Neoadjuvant atezolizumab in combination with sequential nab-paclitaxel and anthracycline-based chemotherapy versus placebo and chemotherapy in patients with early-stage triple-negative breast cancer (IMpassion031): a randomised, double-blind, phase 3 trial. *Lancet* 2020;396(10257):1090-100.

[24] Schmid P, Cortes J, Pusztai L, McArthur H, Kümmel S, Bergh J, et al. Pembrolizumab for early triple-negative breast cancer. *N Engl J Med* 2020;382(9):810-21.

[25] Chen X, Ye G, Zhang C, Li X, Chen Y, Xie X, et al. Superior outcome after neoadjuvant chemotherapy with docetaxel, anthracycline, and cyclophosphamide versus docetaxel plus

cyclophosphamide: results from the NATT trial in triple negative or HER2 positive breast cancer. *Breast Cancer Res Treat* 2013;142(3):549-58.

[26] Ademuyiwa FO, Gao F, Chen I, Northfelt DW, Wesolowski R, Arora M, et al. NCI 10013 - A randomized phase 2 study of neoadjuvant carboplatin and paclitaxel, with or without atezolizumab in triple negative breast cancer (TNBC). *Cancer Res* 2021;81(4 Suppl):PD14-09.

[27] Vriens BE, Aarts MJ, de Vries B, van Gastel SM, Wals J, Smilde TJ, et al. Doxorubicin/cyclophosphamide with concurrent versus sequential docetaxel as neoadjuvant treatment in patients with breast cancer. *Eur J Cancer* 2013;49(15):3102-10.

[28] Merck Sharp & Dohme. *Clinical study report. Study of pembrolizumab (MK-3475) plus chemotherapy vs. placebo plus chemotherapy for previously untreated locally recurrent inoperable or metastatic triple negative breast cancer (MK-3475-355/KEYNOTE-355) - Data on file*, 2021

[29] Merck Sharp & Dohme. *Meeting report. MSD UK early-stage TNBC virtual advisory board meeting*: Merck Sharp & Dohme, 2021

[30] Jack RH, Davies EA, Renshaw C, Tutt A, Grocock MJ, Coupland VH, et al. Differences in breast cancer hormone receptor status in ethnic groups: a London population. *Eur J Cancer* 2013;49(3):696-702.

[31] National Disease Registration Service, NHS Digital. *Case-mix adjusted percentage of cancers diagnosed at stages 1 and 2 by CCG in England, 2019 [Internet]*: National Disease Registration Service, 2021 [accessed 4.5.22] Available from: https://digital.nhs.uk/data-and-information/publications/statistical/case-mix-adjusted-percentage-cancers-diagnosed-at-stages-1-and-2-by-ccg-in-england/2019

[32] Advisory Committee on Breast Cancer Screening. Screening for breast cancer in England: past and future. *J Med Screen* 2006;13(2):59-61.

[33] Cardoso F, Paluch-Shimon S, Senkus E, Curigliano G, Aapro MS, Andre F, et al. 5th ESO-ESMO international consensus guidelines for advanced breast cancer (ABC 5). *Ann Oncol* 2020;31(12):1623-49.

[34] Sterne JAC, Savović J, Page MJ, Elbers RG, Blencowe NS, Boutron I, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ* 2019;366:l4898.

[35] Citrome L. Quantifying clinical relevance. *Innov Clin Neurosci* 2014;11(5-6):26-30.

[36] Sashegyi A, Ferry D. On the interpretation of the hazard ratio and communication of survival benefit. *Oncologist* 2017;22(4):484-6.

[37] U.S. Department of Health and Human Services, National Institutes of Health, National Cancer Institute. *Common Terminology Criteria for Adverse Events (CTCAE). Version 4.0 [Internet]*: U.S. Department of Health and Human Services, National Institutes of Health, National Cancer Institute, 2009 [accessed 27.5.22] Available from: https://evs.nci.nih.gov/ftp1/CTCAE/CTCAE_4.03/Archive/CTCAE_4.0_2009-05-29_QuickReference_8.5x11.pdf

[38] National Institute for Health and Care Excellence. *Pembrolizumab plus chemotherapy for untreated, triple-negative, locally recurrent unresectable or metastatic breast cancer [ID1546]. In development [GID-TA10417] [Internet]*. London: National Institute for Health and Care Excellence, [accessed 22.3.22] Available from: https://www.nice.org.uk/guidance/indevelopment/gid-ta10417

[39] National Institute for Health and Care Excellence. *Pertuzumab for the neoadjuvant treatment of HER2-positive breast cancer. Technology appraisal guidance 424 [Internet]*. London: National Institute for Health and Care Excellence, 2016 [accessed 22.3.22] Available from: https://www.nice.org.uk/guidance/ta424

[40] Gianni L, Pienkowski T, Im YH, Roman L, Tseng LM, Liu MC, et al. Efficacy and safety of neoadjuvant pertuzumab and trastuzumab in women with locally advanced, inflammatory, or early HER2-positive breast cancer (NeoSphere): a randomised multicentre, open-label, phase 2 trial. *Lancet Oncol* 2012;13(1):25-32.

[41] Merck Sharp & Dohme. KEYNOTE-355. Metastatic TNBC cost-effectiveness model [NICE ID1546]. 2020.

[42] National Institute for Health and Care Excellence. *Early and locally advanced breast cancer: diagnosis and management. NICE guideline 101 [Internet]*. London: National Institute for Health and Care Excellence, 2018 [accessed 21.3.22] Available from: https://www.nice.org.uk/guidance/ng101

[43] Masuda N, Lee SJ, Ohtani S, Im YH, Lee ES, Yokota I, et al. Adjuvant capecitabine for breast cancer after preoperative chemotherapy. *N Engl J Med* 2017;376(22):2147-59.

[44] Aly A, Shah R, Hill K, Botteman MF. Overall survival, costs and healthcare resource use by number of regimens received in elderly patients with newly diagnosed metastatic triple-negative breast cancer. *Future Oncol* 2019;15(9):1007-20.

[45] Latimer N. *NICE DSU Technical Support Document 14: Undertaking survival analysis for economic evaluations alongside clinical trials - extrapolation with patient-level data*. Sheffield: Decision Support Unit, ScHARR, 2011 Available from: http://www.nicedsu.org.uk

[46] Tremblay G, Haines P, Briggs A. A criterion-based approach for the systematic and trensparent extrapolation of clinical trial survival data. *JHEOR* 2015;2(2):147-60.

[47] Rutherford MJ, Lambert PC, Sweeting MJ, Pennington R, Crowther MJ, Abrams KR, et al. *NICE DSU Technical Support Document 21: Flexible methods for survival analysis*. Sheffield: Decision Support Unit, ScHARR, 2020 Available from: http://www.nicedsu.org.uk

[48] Huang M, Haiderali A, Hu P, Mejia J. PCN308 Health utility in patients with previously treated metastatic TNBC. *Value Health* 2020;23:S78.

[49] van Hout B, Janssen MF, Feng YS, Kohlmann T, Busschbach J, Golicki D, et al. Interim scoring for the EQ-5D-5L: mapping the EQ-5D-5L to EQ-5D-3L value sets. *Value Health* 2012;15(5):708-15.

[50] Ara R, Brazier JE. Populating an economic model with health state utility values: moving toward better practice. *Value Health* 2010;13(5):509-18.

[51] National Institute for Health and Care Excellence. *Guide to the methods of technology appraisal 2013 (2018 revision) [Internet]*. London: NICE, 2013 (revised 2018) [accessed 21.3.22] Available from: https://www.nice.org.uk/process/pmg9/

[52] Curtis LA, Burns A. *Unit costs of health and social care 2020 [Internet]*. Canterbury: PSSRU, University of Kent, 2020 [accessed 4.5.22] Available from: https://www.pssru.ac.uk/project-pages/unit-costs/unit-costs-2020/

[53] National Institute for Health and Care Excellence. *Pertuzumab for adjuvant treatment of HER2-positive early stage breast cancer. Technology appraisal guidance 569 [Internet]*. London: National

Institute for Health and Care Excellence, 2019 [accessed 22.3.22] Available from: https://www.nice.org.uk/guidance/ta569

[54] National Institute for Health and Care Excellence. *Atezolizumab with nab-paclitaxel for untreated PD-L1-positive, locally advanced or metastatic, triple-negative breast cancer. Technology appraisal guidance 639 [Internet]*. London: National Institute for Health and Care Excellence, 2020 [accessed 22.3.22] Available from: https://www.nice.org.uk/guidance/ta639

[55] Georghiou T, Bardsley M, Trust N. *Exploring the cost of care at the end of life [Internet]*: Nuffield Trust, 2014 [accessed 30.4.20] Available from: https://www.nuffieldtrust.org.uk/research/exploring-the-cost-of-care-at-the-end-of-life

[56] NHS. *NHS reference costs 2019-2020 [Internet]*: NHS, 2021 [accessed 22.3.22] Available from: https://improvement.nhs.uk/resources/national-cost-collection/

[57] National Institute for Health and Care Excellence. *Pembrolizumab for treating locally advanced or metastatic urothelial carcinoma after platinum-containing chemotherapy. Technology appraisal guidance 519 [Internet]*. London: National Institute for Health and Care Excellence, 2018 [accessed 22.3.22] Available from: https://www.nice.org.uk/guidance/ta519

[58] National Institute for Health and Care Excellence. *Nivolumab for adjuvant treatment of completely resected melanoma with lymph node involvement or metastatic disease. Technology appraisal guidance 684 [Internet]*. London: National Institute for Health and Care Excellence, 2021 [accessed 22.3.22] Available from: https://www.nice.org.uk/guidance/TA684

[59] National Institute for Health and Care Excellence. *Nivolumab with ipilimumab for untreated advanced renal cell carcinoma. Technology appraisal guidance 581 [Internet]*. London: National Institute for Health and Care Excellence, 2019 [accessed 22.3.22] Available from: https://www.nice.org.uk/guidance/TA581

[60] Sikov WM, Polley M-Y, Twohy E, Perou CM, Singh B, Berry DA, et al. CALGB (Alliance) 40603: long-term outcomes (LTOs) after neoadjuvant chemotherapy (NACT) +/- carboplatin (Cb) and bevacizumab (Bev) in triplenegative breast cancer (TNBC). *J Clin Oncol* 2019;37(15 Suppl):591.

[61] Walsh EM, Shalaby A, O'Loughlin M, Keane N, Webber MJ, Kerin MJ, et al. Outcome for triple negative breast cancer in a retrospective cohort with an emphasis on response to platinum-based neoadjuvant therapy. *Breast Cancer Res Treat* 2019;174(1):1-13.

[62] Grimm SE, Pouwels X, Ramaekers BLT, Wijnen B, Knies S, Grutters J, et al. Development and Validation of the TRansparent Uncertainty ASsessmenT (TRUST) Tool for assessing uncertainties in health economic decision models. *Pharmacoeconomics* 2020;38(2):205-16.

[63] Kaltenthaler E, Carroll C, Hill-McManus D, Scope A, Holmes M, Rice S, et al. The use of exploratory analyses within the National Institute for Health and Care Excellence single technology appraisal process: an evaluation and qualitative analysis. *Health Technol Assess* 2016;20(26):1-48.