NICE National Institute for Health and Care Excellence

NICE guidance

Idelalisib for treating refractory follicular lymphoma

Technology appraisal guidance

Published: 2 October 2019
www.nice.org.uk/guidance/ta604

Your responsibility

The recommendations in this guidance represent the view of NICE, arrived at after careful consideration of the evidence available. When exercising their judgement, health professionals are expected to take this guidance fully into account, alongside the individual needs, preferences and values of their patients. The application of the recommendations in this guidance are at the discretion of health professionals and their individual patients and do not override the responsibility of healthcare professionals to make decisions appropriate to the circumstances of the individual patient, in consultation with the patient and/or their carer or guardian.

Commissioners and/or providers have a responsibility to provide the funding required to enable the guidance to be applied when individual health professionals and their patients wish to use it, in accordance with the NHS Constitution. They should do so in light of their duties to have due regard to the need to eliminate unlawful discrimination, to advance equality of opportunity and to reduce health inequalities.

Commissioners and providers have a responsibility to promote an environmentally sustainable health and care system and should assess and reduce the environmental impact of implementing NICE recommendations wherever possible.

Contents

This guidance replaces TA328.

1 Recommendations

1.1 Idelalisib is not recommended, within its marketing authorisation, for treating follicular lymphoma that has not responded to 2 prior lines of treatment in adults.

1.2 This recommendation is not intended to affect treatment with idelalisib that was started in the NHS before this guidance was published. People having treatment outside this recommendation may continue without change to the funding arrangements in place for them before this guidance was published, until they and their NHS clinician consider it appropriate to stop.

Why the committee made these recommendations

The options after 2 lines of treatment for follicular lymphoma in the NHS include a range of chemotherapy treatments. The choice of the specific treatment depends on what the person has had already.

Idelalisib has not been compared directly with current individual chemotherapeutic treatments. There are several complex indirect comparisons of idelalisib, but these are based on sparse data and have other problems. So, it is unclear whether idelalisib is better than individual chemotherapeutic regimens currently offered by the NHS and, if so, by how much.

There is a wide range of cost-effectiveness estimates but, because the evidence is weak, idelalisib is not considered to represent a cost-effective use of NHS resources. Therefore, idelalisib cannot be recommended for routine use in the NHS.

Idelalisib cannot be recommended for inclusion in the Cancer Drugs Fund. This is because data collected in the Cancer Drugs Fund cannot resolve the key problems determining whether idelalisib is more effective that chemotherapy, nor document the adverse effects associated with individual chemotherapeutic regimens.

2 Information about idelalisib

Marketing authorisation indication	Idelalisib (Zydelig, Gilead) has a marketing authorisation 'as monotherapy for the treatment of adult patients with follicular lymphoma (FL) that is refractory to two prior lines of treatment'.
Dosage in the marketing authorisation	Idelalisib is administered orally at a dose of 150 mg, twice daily. Treatment is continued until disease progression or unacceptable toxicity.
Price	The list price for idelalisib is £3,114.75 per pack of 60 x 150 mg film-coated tablets (excluding VAT, company submission). The company has a commercial arrangement for idelalisib, which would apply if the technology had been recommended.

3 Committee discussion

The appraisal committee (section 4) considered evidence submitted by Gilead and a review of this submission by the evidence review group (ERG). See the committee papers for full details of the evidence.

Treatment pathway

This appraisal focuses on idelalisib as a treatment used after an anti-CD20 monoclonal antibody and an alkylating agent

3.1 Idelalisib is an oral treatment for follicular lymphoma. The marketing authorisation specifies a population with follicular lymphoma that is 'refractory to 2 prior lines of treatment'. However, the committee noted that the clinical evidence for idelalisib included a population in which there was a wide range in the numbers of prior chemotherapeutic regimens previously used by patients. Furthermore, the committee noted that the population considered in the company's submission was narrower than the population specified by the marketing authorisation, which is 'double-refractory'. This means that the follicular lymphoma has not responded, or showed only a limited response, to at least 2 previous treatments. The company specified that the previous treatments had to include an anti-CD20 monoclonal antibody (rituximab or obinutuzumab) and chemotherapy containing an alkylating agent (for example, cyclophosphamide). The clinical experts considered this double-refractory population to have an unmet need for treatment options. The committee agreed with the company that idelalisib is likely to be used after an anti-CD20 monoclonal antibody and an alkylating agent.

There is no single standard-of-care chemotherapy for the population in this appraisal

3.2 In the population that would be offered idelalisib (see section 3.1), most people are currently offered chemotherapy including (as listed in the NICE scope) cyclophosphamide-containing regimens, fludarabine-containing regimens, bendamustine and chlorambucil. The clinical experts explained that these vary in effectiveness and toxicity. The choice of chemotherapy regimen depends on individual circumstances, and takes into account previous chemotherapy, plus

clinician and patient preference. Because of this, there is no single standard-of-care chemotherapy regimen in this population. The committee concluded that all the chemotherapeutic regimens listed in the scope were appropriate comparators.

Best supportive care is also a comparator for idelalisib

3.3 The clinical experts stated that, after second-line treatment, people who could not have chemotherapy would be offered best supportive care instead. The clinical experts stated that some of those people could take idelalisib because it has a different toxicity profile to chemotherapy. So, while people may not tolerate the specific adverse effects of chemotherapy, they may tolerate those of idelalisib. The committee concluded that best supportive care is an option and a relevant comparator for people who cannot have chemotherapy, but who can take idelalisib.

The effect of autologous stem cell transplantation after idelalisib or chemotherapy should be considered

3.4 The clinical experts stated that response to second-line treatment is consolidated with autologous stem cell transplantation in people who are considered fit enough, in line with NICE's guideline on non-Hodgkin's lymphoma: diagnosis and management. They also noted that, if a patient was healthy, autologous stem cell transplantation would be considered at later lines of therapy, as soon as the cancer responded to treatment. The clinical experts considered that autologous stem cell transplantation can improve prognosis, and that it could be used after idelalisib. The committee concluded that autologous stem cell transplantation was not a comparator to idelalisib, but that it was possible to use idelalisib as a bridge to autologous stem cell transplantation if there was a sufficient response to idelalisib. It also understood that patients could have a transplant before disease progression. The committee further concluded that it would need to consider the effects of autologous stem cell transplantation used after idelalisib or chemotherapy in the clinical- and cost-effectiveness results provided by the company.

Clinical evidence

There is no trial evidence on how much better idelalisib is

compared with current clinical practice

3.5 The key clinical evidence for idelalisib came from the single-arm phase II DELTA study. This study included 125 patients with indolent (slowly growing) non-Hodgkin's lymphoma, 72 of whom had follicular lymphoma refractory to 2 or more lines of treatment. The primary outcome measured overall response rate. Overall survival and progression-free survival were among the secondary outcomes. DELTA had no control group; all patients in the study had idelalisib. The committee understood that this study was designed to show that idelalisib was a safe treatment. It also understood that the company had not carried out further effectiveness trials planned to support its application for a marketing authorisation (see section 3.6), which the European Medicines Agency granted on the basis of the DELTA study. The data from DELTA first provided by the company (cut-off date June 2015) showed an overall response rate of 55.6% and a median progression-free survival of 11.0 months; median overall survival was not reached, but the company estimated it to be 38.1 months. During consultation, the company submitted updated data based on a later cut-off date (August 2018); they showed that the median overall survival result was better than the company's original estimate (results are academic in confidence). The company stated that the latest data suggested that the overall survival benefit was largely as a result of survival after the cancer had progressed. The company supplemented the DELTA study with another source of evidence for idelalisib: the Compassionate Use Programme (CUP). This provided retrospective observational data from patients with follicular lymphoma having compassionate treatment in the UK and Ireland. The company took a subset of 79 patients with relapsed or refractory follicular lymphoma that had been treated with idelalisib. In these patients, median progression-free survival was 7.1 months, and median overall survival was not reached. The committee concluded that the evidence in DELTA and the CUP was not adequate enough for using to determine how well patients on idelalisib fared compared with people who had not taken idelalisib.

DELTA is a non-comparative safety study and is not designed to determine the clinical effectiveness of idelalisib compared with current NHS treatment

3.6 The committee discussed the lack of direct comparative evidence for idelalisib. The company explained that it had originally planned to develop a confirmatory

randomised controlled trial after DELTA. However, the European Medicines Agency granted a marketing authorisation based on the non-comparative DELTA study. The company therefore chose not to pursue the planned studies of effectiveness of idelalisib in follicular lymphoma. The committee considered that this had resulted in an important gap in the evidence base for idelalisib, making it difficult to carry out an informed assessment of the effectiveness of idelalisib. It reflected that a randomised controlled trial in the intended population would help to resolve this. The committee concluded that the paucity of data, and the lack of comparative data, were key shortcomings to determining the clinical effectiveness of idelalisib compared with currently offered treatments in the NHS.

It is unclear whether the DELTA population or the Compassionate Use Programme cohort more closely reflects clinical practice

3.7 Despite the absence of a controlled trial, the committee discussed the evidence presented by the company. The committee discussed whether the populations in DELTA and the CUP are generalisable to people who take idelalisib in clinical practice:

- The committee questioned why only 79% (and not 100%) of patients in DELTA had cancer refractory to 2 or more lines of therapy. The company explained that some patients had had an anti-CD20 monoclonal antibody and an alkylating agent together in the same line of therapy (rather than sequentially), which the company defined as a single treatment regimen. The committee accepted this.

- The committee noted the difference in Eastern Cooperative Oncology Group (ECOG) performance status and Follicular Lymphoma International Prognostic Index (FLIPI) I and II scores between DELTA and the CUP. Notably, 8% of patients in DELTA had an ECOG score of 2 to 4 compared with 25% of patients in the CUP, reflecting poorer performance among patients in the CUP. The clinical experts stated that the ECOG performance status in the CUP more closely reflected clinical practice than that in DELTA.

- The clinical experts noted that the time since completing the last therapy was shorter in DELTA than in the CUP, suggesting that patients in DELTA had a poorer prognosis.

  The committee agreed that the populations in DELTA and the CUP were different. The studies differed in design (for example, how they defined disease progression; see section 3.10). Also, patient and disease characteristics at baseline differed, with some suggesting a more favourable prognosis in DELTA than in the CUP, and others suggesting the opposite. The committee queried why the company had not chosen to combine the 2 studies. The company explained that it considered the DELTA data, being a trial, to be too different from clinical practice-based data. The company argued that the population enrolled in DELTA better reflected clinical practice. The committee did not accept the company's rationale for not combining the data and wondered why the company did not, for the same reason, find it inappropriate to compare DELTA with the Haematological Malignancy Research Network (HMRN, see section 3.8), a clinical practice-based cohort. The clinical experts suggested that the CUP cohort was more likely to reflect the intended UK treatment population because it was a 'real-world' study with patients from Britain and Ireland. However, the clinical experts acknowledged that such studies lack the methodological rigour typical of a clinical trial. The committee recognised that the company submitted data from the CUP in its original submission, but not in its response to the consultation, despite the committee having concluded that it would consider both studies in the appraisal consultation document. The committee concluded that it was unclear whether the DELTA population or the CUP cohort more closely reflected clinical practice and took both into account for decision making.

Evidence of effectiveness for chemotherapy is limited because it does not provide enough information on individual regimens or best supportive care

3.8 Because the company had not done a controlled trial, it sourced evidence for comparators from a registry: the UK HMRN. This registry comprised a retrospective observational cohort. It included 26 patients with follicular lymphoma refractory to 2 previous lines of therapy, including rituximab and chemotherapy at first or second line, who had chemotherapy at third line. The committee noted that it was unclear whether the registry included a 'double-refractory' population and how many lines of chemotherapy each person had previously had. The committee was concerned that the HMRN population did not reflect the population defined for the current appraisal. The company did

not have access to the data directly; instead, it requested analyses from the HMRN, based at the University of York. The company combined several comparators into a single chemotherapy comparator arm (see section 3.15). The original data cut (August 2013) submitted by the company for the committee's first meeting showed an estimated median progression-free survival of about 17 months, and a median overall survival of about 20 months. During consultation, the company submitted updated data based on a later data cut (August 2016), which showed that median overall survival was longer than the original data cut (detailed results are academic in confidence). The latest data cut had 34 patients, which the committee recognised as being a small number of patients; this number fell further in the matching analyses (see sections 3.13 and 3.15). The company explained that the analyses did not censor patients at transplantation. The committee understood that, because transplantation can improve prognosis (see section 3.4, not censoring patients at transplantation had the potential to invalidate clinical-effectiveness results. It acknowledged that it was likely that the HMRN was the only source of comparative data for the UK. Nevertheless, it concluded that the HMRN had limited value because it:

- did not define the number of prior therapies each person had
- did not provide information on individual chemotherapy regimens or best supportive care
- was likely to differ from DELTA in ways that influenced patient outcomes.

It is more likely that patients in DELTA would have a better prognosis than those in the HMRN

3.9 The company explained that it considered the DELTA population (that is, those having idelalisib) at baseline to have a worse prognosis than the HMRN population (that is, those having chemotherapy) at the time of reaching registry eligibility criteria. However, the clinical experts believed that the clinical practice-based HMRN population was likely to have a worse prognosis. The committee agreed that it was difficult to compare the 2 patient populations because of differences in baseline patient characteristics. It also agreed that the interplay of these differences made it difficult to judge which way overall population health would be biased. Either way, the committee appreciated that these differences could have confounded the association between treatment and survival. It considered the differences in how some variables had been

defined between the 2 studies (see section 3.13), and that people in the HMRN had similar or worse disease in a shorter time from diagnosis. Based on these considerations, the committee concluded it more likely that patients in DELTA would have a better prognosis than those in the HMRN.

It is appropriate to assume equal progression-free survival for chemotherapy and idelalisib, but progression-free survival should be taken from the HMRN

3.10 Based on the latest data for DELTA, idelalisib was associated with longer overall survival than was chemotherapy in the HMRN. However, these data showed that patients on idelalisib from 12 months onwards had a shorter progression-free survival than patients on chemotherapy. The committee agreed that shorter progression, then longer survival, was improbable. The company stated that it believed this counterintuitive result because of differences in the frequency with which each study measured disease progression. In both studies, progression-free survival was defined as 'time from initiation of treatment to date of first disease progression' or 'death from any cause'. However, in DELTA, patients had more regular scans compared with patients in the HMRN cohort; the company proposed that this is likely to have identified progression at an earlier stage for patients having idelalisib than having chemotherapy. The committee was not presented with data on the frequency of scanning in the HMRN cohort. The company stated that 'the HMRN dataset will systematically overpredict PFS [progression-free survival] in comparison to DELTA. Along with other issues of comparability across the two datasets, PFS comparison is rendered almost meaningless'. The ERG stated that this would have affected all methods of matching analyses submitted by the company (see section 3.12) equally because they all used progression-free survival from both samples. The company explained that the analyses did not account for stem cell transplantation before disease progression; specifically, it heard from the company that the analyses for progression-free survival did not censor patients at transplantation (see section 3.8). The committee considered that this may have biased the results. To address this issue, in some (but not all) of its analyses for cost effectiveness, the company assumed that the results for progression-free survival for chemotherapy and idelalisib were equivalent by using progression-free survival from DELTA in both treatment arms. The committee recognised that identifying progression at an earlier stage (for patients having idelalisib) was likely to have underestimated the treatment costs of idelalisib

because people have idelalisib until disease progression (or unacceptable toxicities). The committee agreed that, in the absence of a trial, time to progression should reflect that seen in the NHS, that is, using data from the HMRN. The committee concluded that, in the absence of data, it was satisfied with the assumption that progression-free survival was equivalent between chemotherapy and idelalisib. However, it considered that the company should have taken estimates of progression-free survival from the HMRN, and not DELTA.

Indirect treatment comparisons

Comparisons with previous lines of therapy without clear methods do not give reliable estimates of effectiveness

3.11 In the absence of a head-to-head randomised comparison of idelalisib with chemotherapy, in its original submission, the company compared progression and survival on idelalisib with the last line of chemotherapy before idelalisib, for DELTA and the CUP. The company used the time to progression with the previous line of chemotherapy (second line) as a proxy for time to progression with chemotherapy at third line (that is, when idelalisib would be a treatment option). The committee heard at the first meeting that the company did not perform a paired matching analysis. The committee was aware that the company used data from the CUP to compare idelalisib with previous lines of therapy but, at the second meeting, the company stated that it did not have access to the CUP data. The committee discussed several issues around the analysis:

- The ERG commented that this comparison should be considered with caution because of the potential bias from including only patients who survive to have idelalisib; these patients would have been healthier than the entire chemotherapy population that existed at the previous line of therapy. The committee also recognised that patients on second-line chemotherapy may be healthier than patients on third-line chemotherapy. The committee agreed that these were potential sources of bias.
- The committee recognised that the fitness of a patient and the effectiveness of treatment could decline between 1 treatment line and the next. Therefore, comparing the idelalisib group with the group who previously had chemotherapy was not comparing like with like.

- For DELTA, the data on the previous line of therapy were based on 'clinician recall' (because people entered the study at the point at which they would have idelalisib), which may have been subject to bias. The committee was aware that the definition of trial-based and historical progression may have differed, and that progression may be more quickly identified during a clinical trial when patients are actively monitored and scanned (see section 3.10). The committee recognised this would bias results against idelalisib.

- The clinical experts recognised that it was difficult to draw conclusions on the comparative effect of idelalisib by looking retrospectively at previous lines of therapy, and more so in the CUP. This was because time to progression on idelalisib was already determined retrospectively, and treatment previous to that was even more distant in time.

- For both studies, previous lines of therapy reflected a range of chemotherapeutic regimens, at a range of different points in the treatment pathway (up to fourteenth line). These regimens may differ from 1 another in clinical effectiveness and adverse effects, and may not be generalisable to the population in the appraisal.

- The committee noted that DELTA began recruiting patients in 2011 from a range of countries. It recognised that chemotherapy options available for previous treatment in DELTA may have changed over time and may have differed in other countries. Therefore, the chemotherapy treatments used may not represent current UK clinical practice.

- The committee recognised that this comparison did not compare the same patients with each other because the company had not done a paired matching analysis. The committee considered that a paired matching analysis would have minimised confounding.

- The committee was aware that the company had not analysed idelalisib against the individual chemotherapy agents identified in the NICE scope.

  The committee was aware that many of these issues also applied to the matching comparisons with the HMRN data submitted by the company (see section 3.12). The committee concluded that, because the company's comparison with previous lines of therapy lacked clear methods and had multiple sources of bias, its results were not reliable.

The company performed several matching indirect comparisons

with the HMRN data

3.12 In response to the consultation, the company submitted the following matching indirect comparisons using updated data (see sections 3.5 and 3.8) from DELTA and the HMRN:

- Unanchored matching adjusted indirect comparisons with the HMRN cohort (see section 3.8) including:
  - An 'updated' matching adjusted indirect comparison: the company matched individual patient-level data from the HMRN cohort to the population-level descriptive characteristics of patients in the DELTA study. This estimated the effects of chemotherapy in a DELTA-like population.
  - A 'reverse' matching adjusted indirect comparison: the company matched individual patient-level data from the DELTA study to the population-level descriptive characteristics of patients in the HMRN cohort. This estimated the effects of idelalisib in a HMRN-like population.
- Propensity score matching analyses (see section 3.14):
  - With DELTA as the 'treated' group and the HMRN as the 'control' group: it matched individual patient-level data from the DELTA study to individual patient-level data from the HMRN cohort. This estimated the effects of chemotherapy in a DELTA-like population.
  - With the HMRN as the 'treated' group and DELTA as the 'control' group: it matched individual patient-level data from the HMRN cohort to individual patient-level data from the DELTA study. This estimated the effects of idelalisib in a HMRN-like population.

    The company submitted the propensity score analyses in response to the appraisal consultation document and because the committee had recognised that patient-level data were available from both the DELTA study and the HMRN cohort. The company submitted the 'reverse' matching adjusted indirect comparison in response to the statements in the appraisal consultation document from the ERG and committee, noting that estimating the effect of idelalisib in the HMRN (UK) population would provide more individual patient data, and better represent NHS clinical practice.

The 'updated' matching adjusted indirect comparison is preferred over the 'reverse' comparison, although both are unreliable

3.13 The committee discussed the factors affecting prognosis that the 'updated' and 'reverse' matching adjusted indirect comparison analyses should have included (see section 3.12), and the sample sizes available. The committee was aware that it had not been presented with a systematic review of risk factors for progression and death. In addition to history of autologous stem cell transplantation (see section 3.4), the committee understood that there are other factors associated with progression and death. These include, but are not limited to, the components of FLIPI I and FLIPI II, notably: age, serum beta 2 microglobulin levels, bone marrow involvement, size of the largest involved lymph node, haemoglobin levels and the presence of bulky disease. Other factors include time in previous remission, time since completing the last therapy, comorbid conditions and previous chemotherapeutic agents. The clinical experts suggested that the FLIPI index is the best validated prognostic tool to use when diagnosing follicular lymphoma, but has limited value at third line. They proposed that a key prognostic indicator would be response to previous therapy, but this input was not captured in the variables chosen in the 'updated' matching adjusted indirect comparison. The company matched only 5 of 7 variables and assumed that these 5 accounted for all prognostic factors and treatment-effect modifiers. The committee was aware that a technical support document published by the Decision Support Unit recommends that, when only single-arm trial data are available, all the characteristics that could influence the outcomes of interest should be adjusted. However, increasing the number of matched characteristics reduced the effective sample size and the precision of the estimates, and the committee understood that the results were sensitive to this. For example, when the company removed the variable 'median time since diagnosis' from the analyses, estimated 2-year overall survival fell by more than 20%. The committee also appreciated that there were unobserved differences between study populations that the analyses could not take into account. It noted that this would have biased the estimates of relative effectiveness if these unknown factors were associated with progression or death. The committee was satisfied that, in the 'reverse' matching adjusted indirect comparison, the company matched all 7 variables. However, in both 'updated' and 'reverse' comparisons, the committee noted that DELTA and the HMRN cohort defined 2 of the matched variables ('bulky disease' and 'time to diagnosis') differently, so the estimates of relative effectiveness from the

'updated' and 'reverse' comparison were likely to be biased. The committee was aware that the effective sample size for the 'reverse' matching adjusted indirect comparison was 26.7 people (26.7 represents a statistic rather than an actual number of people). The committee concluded that, although it preferred the 'updated' over the 'reverse' comparison, sparse data and potential confounding meant that the clinical-effectiveness results reflecting both were unreliable.

Propensity score matching analysis is more reliable than the 'updated' and 'reverse' matching adjusted indirect comparisons, but all are problematic

3.14 In response to the consultation, the company submitted 2 analyses with propensity scores using individual patient-level data from both DELTA and the HMRN, and using both populations applied as the 'treated group' (see section 3.12). The company matched all 7 variables (see section 3.13). The committee considered that matching on bulky disease and 'time to diagnosis' caused problems because each study defined those variables differently (see section 3.13). The committee noted that the propensity score matching analyses excluded patients without a suitable match (50% of patients in DELTA), which reduced the idelalisib sample size to 39 (DELTA as 'treated group') and 35 (the HMRN as 'treated group'). It also noted that the company submitted only 1 of many methods that can be used for propensity score matching, the 'three-nearest-neighbour' matching method (that is, matching baseline characteristics to the 3 closest patients in the comparison group). The company did not provide a rationale for its choice, nor had it conducted sensitivity analyses using alternative methods. The ERG commented that this was an important limitation because using other methods could change the results. In addition, the ERG noted that the summary measures from DELTA and HMRN groups using propensity matching were not as similar at baseline as they were in the 'updated' and 'reverse' matching adjusted indirect comparisons. The committee recognised that propensity scoring matches individual patient-level data together, rather than matching individual patient-level data to the mean, as was done in the matching adjusted indirect comparisons. Therefore, it did not expect comparable baseline characteristics between groups after matching. The ERG explained that, although all the analyses had problems, it preferred the 'reverse' to the 'updated' matching adjusted indirect comparison and propensity score matching analyses. However, the committee noted that sample sizes in the analyses using propensity score matching were more balanced across treatment

arms than in the matching adjusted indirect comparisons, although they were small. The committee agreed that, because propensity score analyses match individual patient data for both populations, they provide more precise estimates than those from matching adjusted indirect comparisons. Despite this, the committee concluded that analyses using propensity score matching were associated with high levels of uncertainty. It also concluded that the matching comparisons submitted by the company did not provide the committee with robust information to evaluate the relative effectiveness of idelalisib compared with chemotherapy or best supportive care.

The 'blended' comparator assumes that different chemotherapeutic treatments are similarly effective and tolerated, but this assumption is not justified

3.15 The committee was aware that the NICE scope listed the chemotherapeutic agents separately, and heard from the clinical experts that the therapies are likely to differ in effectiveness and tolerability. It recognised that the company had combined treatments together, which reflected a 'blended' comparator. The committee considered that evidence for the effectiveness of separate chemotherapeutic agents might exist from the trials that provided evidence for using chemotherapy, or from registries other than the HMRN. The Cancer Drugs Fund clinical lead considered that currently used individual chemotherapy regimens would be based on controlled clinical trial evidence. However, the committee also considered that this evidence might be difficult to source. At the second meeting, the company explained that it had not identified other registries or trials providing information on individual chemotherapy regimens. It further explained that it chose not to collect this information from the HMRN cohort because the sample size was already small (n=34, based on the latest data cut). The committee recalled that clinicians did not consider the chemotherapeutic regimens to be equally effective or have the same profile of adverse effects (see section 3.2). It understood that there were practical issues associated with estimating the effects of individual chemotherapy regimens. However, it noted that it had not been presented with evidence that justified the company's 'blended' comparator assumption that chemotherapeutic agents in UK practice could be considered similarly effective and tolerated, and at the same costs.

There are no data for best supportive care

3.16 The committee recognised that the company did not provide data to describe the natural history of disease in patients having best supportive care. The committee asked the company whether the HMRN cohort could provide this, or whether clinical trial data were available from control arms of clinical trials of chemotherapy. The committee was aware that the company did not provide clinical data with which to compare idelalisib with best supportive care. Even so, the company compared idelalisib with best supportive care in its cost-effectiveness analyses by relying on an assumption that patients would progress instantly in the absence of an active treatment (see section 3.18).

The current analyses based on the data provided are insufficient

3.17 The committee appreciated the company's attempts to reduce uncertainty in determining the clinical effectiveness of idelalisib compared with chemotherapy. The company did this by submitting longer-term data from DELTA and the HMRN, and by conducting additional indirect analyses, in response to concerns identified by the committee in the appraisal consultation document. However, the committee noted several concerns with the additional data and indirect analyses:

- The indirect analyses were sensitive to changes in assumptions, and did not confirm a difference between treatments (see section 3.13).
- There were differences in the definition of key variables (see section 3.13) and outcomes (see section 3.10).
- The company explained that it had not taken into account autologous stem cell transplantation before disease progression or death. The committee agreed this had the potential to invalidate results (see section 3.8).
- In relation to the updated data it had submitted, the company stated that 'the HMRN dataset will systematically overpredict progression-free survival in comparison to DELTA. Along with other issues of comparability across the two datasets, progression-free survival comparison [between DELTA and the HMRN] is rendered almost meaningless'. The ERG stated that the overprediction of progression-free survival would apply equally to all matching analyses because they all used progression-free survival from both samples (see section 3.10).

- The committee noted its previous consideration that the company's decision not to pursue a randomised controlled trial of idelalisib in this population had left an important gap in the evidence base (see section 3.6).

  The committee appreciated statements from professional organisations and clinical experts that accepted the potential for idelalisib to extend life compared with current treatments for 'double-refractory' follicular lymphoma. A clinical expert suggested that idelalisib might prolong life by 12 months or more in some patients. However, a professional organisation suggested that estimating the extension to life with certainty is difficult. The committee acknowledged the efforts of the company, but noted that concerns about the comparative effectiveness evidence made it difficult to establish the benefit of idelalisib compared with chemotherapy or best supportive care (see section 3.14). It took this into account in its decision making.

The company's economic models

The model structures are appropriate for modelling

3.18 The company submitted 4 economic analyses using different model structures and sources of clinical data, which are summarised in table 1. In its models, the company used secondary endpoints (overall survival and progression-free survival) rather than the primary outcome from DELTA (overall response rate) to determine cost effectiveness.

Table 1 Summary of comparisons used in the economic modelling

Comparison	Idelalisib data source	Comparator data source	Model type
A (company base case)	DELTA	All chemotherapy combined: DELTA data from 'self-control' previous line of treatment as a proxy for current chemotherapy	Markov cohort – state transition

B	DELTA	All chemotherapy combined: matching adjusted survival data from chemotherapy regimens of the Haematological Malignancy Research Network cohort (either matching adjusted indirect comparison or propensity score matching analysis; see section 3.12)	Partitioned survival model
C	Data from the Compassionate Use Programme cohort and DELTA	All chemotherapy combined: time-to-progression data from 'self-control' previous line of treatment as a proxy for current chemotherapy	Markov cohort – state transition
D	DELTA	Best supportive care: no data (see section 3.16, company assumes instant disease progression	Markov cohort – state transition

Comparisons A, C and D reflected state transition models. In these, 'transition probabilities' determined the movements between states (clinical events) and time spent by patients in each state determined expected costs and quality-adjusted life years (QALYs). Comparison B used a 'partitioned survival analysis' approach in which survival curves (rather than transition probabilities) determined the proportion of patients in each state at each time point. The committee agreed in its first meeting that it would focus on comparisons A (the company's base case) and B; the company did not provide best supportive care data for comparison D. The committee noted that, in comparisons A and C, the company applied a 'hazard ratio' of 0.75 to adjust for the expected decline in the effectiveness of chemotherapy compared with idelalisib. This meant that, at each successive treatment line, patients would have expected their prognosis to worsen by 75% compared with the previous line of therapy. It also noted that the company had not justified its choice of the value of 0.75, which affected comparisons A and C to different extents. In its updated cost-effectiveness analysis submitted in response to the consultation, the company submitted results exclusively based on comparison B. It explained that, in its view, only this comparison could reflect the survival benefit of idelalisib after disease progression (see section 3.5). The committee preferred comparison B, but noted that all the comparisons were limited by:

- sparse clinical data

- no comparative evidence
- no censoring for autologous stem cell transplantation
- no comparison with or data on best supportive care
- likely confounding (see section 3.9)
- variables defined differently by study (see section 3.13)
- using a blended comparator and the arbitrary choice of 0.75 to reflect the expected worsening in prognosis compared with previous line of therapy (see sections 3.8, 3.13, 3.14, 3.15 and 3.16).

  In response to the consultation, the company did not submit analyses for idelalisib compared with best supportive care, so the committee did not consider this further.

The 'blended comparator' masks cost-ineffective treatments

3.19 The committee considered that using a blend of chemotherapeutic regimens (see section 3.15) as a comparator meant averaging the cost effectiveness of the treatments included, and potentially masked cost-ineffective individual treatments. The committee concluded that the cost effectiveness of idelalisib needed to reflect the comparison with each chemotherapy treatment individually.

Utility values in the economic models

There are no mapped utilities for the committee to consider

3.20 The company submitted utility values to reflect health-related quality of life from a published study (Wild et al. 2006), even though it had collected quality-of-life data in DELTA. The ERG submitted scenario analyses with alternative utility values from other published studies (Bec et al. 2014 and the GADOLIN trial for non-Hodgkin's indolent lymphoma). The clinical experts agreed that all values seemed reasonable because people with follicular lymphoma can expect a good quality of life once the bulk of the disease has decreased. The committee was aware that DELTA collected health-related quality-of-life data using the Functional Assessment of Cancer Therapy - Lymphoma (FACT-Lym) instrument (an extension of FACT - General [G]). This has 15 questions specific to patients with lymphoma. It was also aware that there is a mapping algorithm to map from

FACT-G to EQ-5D, which the ERG requested of the company at the clarification stage. The company did not provide the mapped utilities, arguing that it was not useful because there was no mapping algorithm that specifically matched FACT-Lym to EQ-5D. The company added that the mapped utilities would have been limited to the clinical symptoms captured in EQ-5D. The committee expressed its interest in seeing the mapped utilities in the appraisal consultation document, but the company chose not to submit the values at the second meeting. The committee concluded that it would have liked to see mapped utility values to validate literature-derived values, but that this would not have overcome the other problems associated with the clinical evidence.

Adverse effects with idelalisib are different than those with chemotherapy, but the adverse effects of individual chemotherapy regimens also differ

3.21 The committee noted the adverse effects in people having idelalisib outlined in the European Medicine Agency's risk management plan, which include severe drug-related colitis, pneumonitis and organising pneumonia, and serious infections. The company did not provide relative safety data for chemotherapy. The clinical experts commented that adverse effects with chemotherapy are qualitatively different to those with idelalisib, so it is difficult to compare the safety profiles. The clinical experts commented that idelalisib is generally well tolerated. The committee considered that the adverse effects of chemotherapy are unlikely to differ between haematological malignancies, and thought that data from other malignancies would be valuable. The company included adverse events as disutilities within all the economic models. These disutility values were based on various sources and estimating the incidence using DELTA. The company models assumed the same incidence of adverse effects for both idelalisib and chemotherapy. The committee noted that adverse event disutility is not a key driver in the model. It concluded that idelalisib is unlikely to have the same effect on quality of life because of an adverse effect as chemotherapy, which themselves differed by individual regimen, but that this is unlikely to markedly affect cost-effectiveness results.

Cost-effectiveness estimates

The company submitted several cost-effectiveness estimates

3.22 In its response to the consultation, the company submitted 3 deterministic incremental cost-effectiveness ratios (ICERs) for idelalisib compared with chemotherapy using comparison B (see section 3.18) incorporating a patient access scheme discount; probabilistic ICERs were broadly similar:

- Using the updated matching adjusted indirect comparison data, and including the ERG's corrections to the model, resulted in an estimated ICER of £16,481 per QALY gained.
- Using a propensity score matching analysis with the DELTA group as the treated group (based on updated data), and including the ERG's corrections to the model, with the assumption that progression-free survival with chemotherapy and idelalisib are equal (see section 3.10), resulted in an estimated ICER of £25,605 per QALY gained.
- Using a propensity score matching analysis with the HMRN group as the treated group (based on updated data), and including the ERG's corrections to the model, with the assumption that progression-free survival with chemotherapy and idelalisib are equal (see section 3.10), resulted in an estimated ICER of £26,627 per QALY gained.

The ERG carried out an exploratory cost-effectiveness analysis

3.23 The committee considered that all matching analyses (see sections 3.13 and 3.14) were associated with high levels of uncertainty. It therefore would have liked to see the cost-effectiveness results associated with all matching analyses, which it considered would help to explore the effect of clinical-effectiveness uncertainty on the cost-effectiveness results. The committee noted that the company had not submitted cost-effectiveness analyses using the 'reverse' matching adjusted indirect comparison (see sections 3.12 and 3.22), even after NICE requested this. The ERG was not able to estimate pseudo-patient-level data for the 'reverse' matching adjusted indirect comparison (to estimate parametric survival curves). This was because the company did not submit numbers at risk or survival curves, and declined to provide a model with the option to choose between matching characteristics. At the second committee meeting, the company stated it had not submitted cost-effectiveness results for the 'reverse' matching adjusted indirect comparison because they would have been based on clinically implausible effectiveness

data, so the results would have added to the uncertainty. The company further stated that it submitted the results only for the 'updated' matching adjusted indirect comparison because it had submitted these at the first meeting. The ERG carried out an exploratory analysis to give a crude estimate of an ICER for the 'reverse' matching adjusted indirect comparison. This estimated an ICER of £86,161 per QALY gained, which is well above what NICE normally considers to be an effective use of NHS resources. The committee concluded that it would take into account the ERG's exploratory analysis in its decision making.

The company's and ERG's estimates of cost effectiveness lack robustness

3.24 The committee stated that it considered propensity score matching analyses to be more reliable than matching adjusted indirect comparisons or 'reverse' matching adjusted indirect comparisons, but all had problems (see section 3.13). However, it would have preferred to see estimates from the company using all the submitted clinical-effectiveness estimates. The committee noted that the ERG's exploratory analysis contained all the uncertainties associated with the company's model in addition to potential confounding (see section 3.18). It also recalled all the limitations associated with propensity score matching analyses (see section 3.14). It considered that the cost-effectiveness estimates lacked robustness and concluded that it needed to account for this in its decision making.

Idelalisib has not been shown to represent a cost-effective use of NHS resources

3.25 The committee appreciated that the company had submitted new clinical data and attempted additional indirect analyses. However, the committee noted that these additional data did not include any comparative evidence that appropriately adjusted for potential confounders. The analyses presented other problems including sparse data, matching on variables defined differently and analyses that did not censor patients having a transplant. The committee also appreciated that the ERG considered that, rather than reducing uncertainty, the company's additional data and analyses had instead increased uncertainty. The committee considered whether information from the clinical experts could inform estimates of overall survival with the new treatment (see sections 3.17 and 3.26). However, the company provided no data to support any estimates.

Therefore, the committee concluded that idelalisib was not considered to represent a cost-effective use of NHS resource because of the range of ICERs presented (between £16,481 and £86,161 per QALY gained) and its concerns with the quality of the evidence.

End of life

End-of-life criteria are not met

3.26 The committee considered the advice about life-extending treatments for people with a short life expectancy in NICE's guide to the methods of technology appraisal:

- The committee discussed whether the life expectancy using standard care was likely to be less than 24 months. It was aware that probabilistic analyses submitted by the company provided mean undiscounted life years for chemotherapy of more than 24 months (ranging from 6 to 87 months). Other model outputs that used different matching analyses also estimated that chemotherapy would result in a mean life expectancy of at least 2 years. Therefore, many of the company's own analyses predicted a life expectancy longer than 24 months. The committee was aware that, given the concerns it had about the economic model, any outputs would have to be interpreted with a high level of caution. In the absence of robust modelled data, the committee considered expert clinical opinion. The company reported that, according to the clinical community, life expectancy would be less than 12 months once a patient's cancer becomes refractory to chemotherapy. A letter (submitted by the company) gathering the opinion of 15 clinical experts suggested that disease progression within 24 months of first relapse is associated with an increased risk of death. Additionally, a clinical expert estimated that life expectancy for patients having idelalisib could be 12 months on average. The committee was aware that no data were offered to support these figures. Therefore, the committee went on to consider the unadjusted data from the HMRN. It noted these data showed that a high proportion of patients survived for a median of 2 years, and that mean life expectancy would be longer. The committee weighed up all the available data and considered that, currently, the most reliable source of evidence from which to make its judgement was the HMRN data. It concluded that the short life-expectancy criterion had not been met.

- The committee discussed the criterion of whether the technology provided an extension to life, with a mean of at least 3 months of life compared with usual care. The committee noted that the modelled undiscounted incremental gains were more than 3 months with idelalisib compared with chemotherapy in all the analyses provided. It noted that a clinical expert suggested that it could be more than 3 months. However, no explanation or further data were given to support this figure. The committee also noted that the estimated undiscounted incremental life years in all analyses were more than 3 months with idelalisib compared with chemotherapy. However, given its concerns about the modelling and model inputs, the committee did not consider that the model generated a valid estimate of the mean extension to life with idelalisib. The committee was aware that NICE's guide to the methods of technology appraisal notes state that the appraisal committee must be satisfied that the 'the estimates of the extension to life are sufficiently robust'. Therefore, it concluded that, although idelalisib might prolong life, the magnitude of this is highly uncertain. Because the short life-expectancy criterion had not been met, the committee concluded that idelalisib could not be considered a life-extending treatment at the end of life.

Cancer Drugs Fund

Idelalisib is not a candidate for the Cancer Drugs Fund

3.27 Having concluded that idelalisib could not be recommended for routine use, the committee then considered whether it could be recommended for treating follicular lymphoma within the Cancer Drugs Fund. The committee discussed the arrangements for the Cancer Drugs Fund agreed by NICE and NHS England in 2016, noting NICE's Cancer Drugs Fund methods guide (addendum). It noted that, at the second meeting, the company had expressed an interest in providing idelalisib through the Cancer Drugs Fund. The company confirmed that there are no ongoing comparative trials to provide more robust, controlled evidence. The committee highlighted that, because no evidence showed that idelalisib improved length or quality of life, it encouraged research comparing idelalisib with individual chemotherapeutic regimens. The committee appreciated that the Cancer Drugs Fund is designed to resolve uncertainties, and that the key uncertainty in this appraisal was about the assumptions surrounding the comparative effectiveness of idelalisib with chemotherapy. It was aware that the Cancer Drugs Fund could collect data on the intervention (idelalisib), but not the comparator (chemotherapy), so would not resolve this uncertainty. Also, the committee had not seen reliable evidence that demonstrated idelalisib had 'plausible potential' to be cost effective (see section 3.24). It concluded that

idelalisib could not be recommended for use within the Cancer Drugs Fund as an option for follicular lymphoma that had not responded to 2 previous lines of therapy in adults.

The Cancer Drugs Fund would not address the key uncertainty in the idelalisib appraisal

3.28 In its response to the appraisal consultation document, the company referred to NICE's technology appraisal guidance on daratumumab monotherapy, which was recommended through the Cancer Drugs Fund to treat relapsed and refractory multiple myeloma. It argued there was a parallel between the 2 appraisals because clinical evidence for daratumumab was limited to single-arm trials and an unanchored matching adjusted indirect comparison. The committee recalled that it had recommended daratumumab through the Cancer Drugs Fund because an Early Access programme is collecting data that have the potential to reduce several uncertainties identified by the committee (such as trial generalisability and place in the treatment pathway). The committee reiterated that the key uncertainty in the idelalisib appraisal was about the assumptions and analyses surrounding the comparative effectiveness, and the effectiveness and adverse effects associated with individual chemotherapeutic regimens, and that the Cancer Drugs Fund could not address these issues.

Innovation

Idelalisib is potentially innovative

3.29 The committee noted that idelalisib is innovative in that it has a different mechanism of action to other available treatments for 'double-refractory' follicular lymphoma, and addresses an area of unmet clinical need (see section 3.1). The committee noted comments from the clinical experts that this different mechanism of action might bring psychological benefits for people reluctant to have more chemotherapy. It also noted that no evidence had been submitted to support this. The committee agreed that it could not determine whether the model captured quality-of-life improvements. It also agreed that an oral drug was preferable to an intravenous one. The committee concluded that idelalisib potentially reflects an innovative step change in treatment, but that it had not been presented with evidence to establish this.

4 Appraisal committee members and NICE project team

Appraisal committee members

The 4 technology appraisal committees are standing advisory committees of NICE. This topic was considered by committee B.

Committee members are asked to declare any interests in the technology to be appraised. If it is considered there is a conflict of interest, the member is excluded from participating further in that appraisal.

The minutes of each appraisal committee meeting, which include the names of the members who attended and their declarations of interests, are posted on the NICE website.

NICE project team

Each technology appraisal is assigned to a team consisting of 1 or more health technology analysts (who act as technical leads for the appraisal), a technical adviser and a project manager.

Aminata Thiam and Adam Brooke
Technical lead

Carl Prescott and Ahmed Elsada
Technical adviser

Jeremy Powell
Project manager

ISBN: 978-1-4731-3551-2

Accreditation