Single Technology Appraisal (STA)

Midostaurin for untreated acute myeloid leukaemia [ID894]

ERG's commentary on the addendum submitted by the company

Produced by	CRD and CHE Technology Assessment Group, University of York, Heslington, York YO10 5DD
Date	31/10/2017

Note on the text

All commercial-in-confidence (CIC) data have been highlighted in blue and underlined, all academic-in-confidence (AIC) data are highlighted in yellow and underlined.

Table of Contents

Table of Tables

1 Introduction

The evidence review group (ERG) was requested by NICE to provide a critique of additional evidence submitted by the company in response to ERG's concerns regarding the representativeness of the population recruited to the RATIFY trial .

The company's addendum included:

- Evidence to support the CS statement that FLT3-positive AML has homogenous biology regardless of age group;
- Evidence to support the CS assumption that prognosis of patients with FLT3-positive AML is similarly poor across differing age groups ;
- Evidence that older patients will be eligible for midostaurin therapy;
- Evidence supporting the effeteness of midostaurin in older patients;
- A new propensity score matched analysis of historical control data estimating the clinical effectiveness of midostaurin in older patients.
- New economic evidence incorporating a new age adjustment scenario in which the cost-effectiveness of midostaurin in older patients is estimated.

2 Critique of the additional evidence submitted by company after FAC to support "Age and Outcomes in Patients with FLT3-positive AML given Midostaurin"

2.1 Evidence 1. FLT3-positive AML has homogenous biology regardless of age group.

The argument presented here by the company is that the higher mortality seen in elderly AML patients is due at least in part to unfavourable cytogenetics, but that these unfavourable cytogenetics are rare in FLT3+ve patients. Hence older FLT3+ve patients would not be expected to have a more unfavourable prognosis than younger ones (based on age alone).

The ERG cannot comment in detail on this, but would point out that it does not resolve the uncertainty that arises from the lack of older (aged > 60 years of age) patients in the RATIFY trial.

2.2 Evidence 2. The prognosis of patients with FLT3-positive AML is similarly poor across differing age groups.

The company addendum presents additional evidence to support the argument that the prognosis of patients with FLT3-positive AML is uniformly poor across differing age groups and that disease risk and biology is unrelated to age. The ERG notes that the graphs presented are from unrelated cohorts: that for the ≤60 years is of a UK cohort (n=1609) (Linch et al.2014), the > 60 years cohort is for the

USA (n=243)(Whitman et al. 2010). The company did not confirm that these cohorts are essentially similar apart from age. The ERG notes that the Linch et al. 2014 cohort is divided into different % of FLT 3ITD, whilst the Whitman et al. 2010 cohort is not. The company addendum does not make clear how these cohorts should be compared. Assuming that the > 60 years FLT3-ITD cohort is to be compared with the highest % FLT3 cohort in the study of younger patients, the ERG estimates the following % alive numbers (Table 1). (Note these estimates were generated by a manual reading from the graphs, and the ERG fully acknowledges these are rough estimates.) If other % ITD data are used, then the % alive in the younger cohort is higher and the difference between the younger and older groups is greater.

Table 1: Overall survival in the <60 and >60 cohorts

	≤60 years cohort (Linch et al.2014)	**≥60 years cohort (Whitman et al. 2010)**
	FLT3 ITD >50%	FLT3 ITD
1 year	40%	33%
2 years	25%	20%
5 years	22%	10%

The ERG's interpretation of the data presented in the two graphs does not strongly support the conclusion drawn by the company that there is no fundamental change in disease risk and biology based on age alone among patients with FLT3-positive AML. Further, the ERG notes that rates of OS are uniformly lower in the >60 cohort. Therefore, while the ERG agrees that the prognosis of patients with FLT3-positive AML is poor in all ages, it does not agree that age is not a prognostic factor.

2.3 Evidence 3: Due to shifts in clinical practice, age is no longer the only factor for eligibility in chemotherapy.

The ERG made this point in their report: older patients are not necessarily unfit and some will be eligible to receive treatment. The ERG concurs with this statement and this is why the ERG questions the generalisability of the RATIFY population results: RATIFY excluded patients aged > 60 years.

2.4 Evidence 4: Data from the original submission showed that Midostaurin is effective in patients more than 60 years of age

The ERG agrees that the data in the original submission from a single-arm Phase II study showed that midostaurin was effective in older patients. However, the results also showed that younger FLT3 positive AML patients (≤60 years old) appeared to have better median overall survival ([redacted]), complete remission rate ([redacted]) median event-free survival ([redacted]

[redacted]), and median relapse free survival ([redacted]) than those who were older (>60 year old).

In the addendum the company present additional data from an expanded version of the Phase II study. The first patient entered the original study in June 2012 and in April 2014, after recruitment of n=147 pts, the study was amended including a sample size increase to 284 pts and a dose reduction to 12.5% of the initial dose of midostaurin in case of co-medication with strong CYP3A4 inhibitors (e.g. posaconazole). This study is available only as a conference abstract which focuses on age and the comparison between the first (n=147) and the second cohort (n=137) of the study in terms midostaurin dose-adaptation.

Unlike the results of the original study, the results from the expanded cohort were similar in patients aged younger and older than 60 years. Overall response to induction was the same for patients < 60 and ≥ 60 years old, at 76% (p = 0.81. The cumulative incidence of relapse and death after transplant in both age groups were also without differences, at 13% (p = 0.97) and 16% (p = 0.41) respectively. For patients < 60 years, median OS was 26 months, while for patients ≥ 60 years, it was 23 months, with no statistical difference (p = 0.15). However, death in patients < 60 years was 4%, and 10% in patients ≥ 60 years.

It is unclear to the ERG how long the later recruited patients were followed up for (the median follow-up for the whole cohort was 18 months), or how many events those patients contributed to the analysis. In addition, this analysis was an interim analysis and so the results are uncertain. Furthermore, the ERG is uncertain about the significance of the dose reduction in the expanded cohort and whether this is more or less reflective of clinical practice than the full dose. Finally, it must be noted that the cohort does not include any patients aged over 70 years.

2.5 Evidence 5: When comparing trial data with propensity score-matched historical controls, midostaurin demonstrated efficacy in patients > 60 years old.

From the addendum the ERG understands that the 223 patients from the expanded Phase II study (interim data only) were compared and matched, using propensity score matching, to 415 patients with specifically FLT3 ITD mutations selected from 5 successive clinical trials.

The results of this analysis found OS hazards ratio (HR) (95% CI) = [redacted] and event-free survival (EFS) HR (95%CI) = [redacted] When analysed by age group the results suggest that the treatment effect of midostaurin is greater in those aged over 60 years than in younger patients: OS hazards ratio between those on midostaurin and those on standard of care was [redacted] in patients ≤ 60 years, and [redacted] in patients > 60 years.

The ERG attempted to check this analysis but were unable to do so fully because the citations given in the addendum do not relate to all 5 trials and therefore information pertaining to only two trials could be checked by the ERG: HD 98-B (Schlenk et al. 2004 and Schlenk et al. 2006) and AMLSG 07-04 (Schlenk et al. 2016).

The ERG has the following comments.

1. Information from the publications of these two studies suggests that the historical control may not be reflective of current clinical practice. The treatment regimen does not match that given in the RATIFY trial, in particular the historical cohort trials induction included a lower dose of cytarabine, but also included etoposide (Table 2).

Table 2: Summary of historical control data

	HD 98-B	**07-04**
Date of recruitment	Feb 1998 to Sept 2001 (Schlenk 2004) or Aug 1997 to April 2003(Schlenk 2006)	Aug 2004 to Jan 2006
Number of patients contributing to historical control	23	203
Age range of total cohort	61-84.5 years	≤60 years
Age range of patients contributing to historical control	unknown	unknown
Induction:	Idarubicin 12mg.m^2 IV days 1+3 Cytababine 100mg/m^2 days 1-5 Etoposide 100 mg days 1+3 If response achieved a 2nd cycle was given, if not a more intensive regimen was given.	2 cycles of Idarubicin 12mg/m^2 IV days 1+3+5 Cytababine 100mg/m^2 days 1-7 Etoposide 100 mg days 1-3
Consolidation	First consolidation cycle: Cytarabine 0.5 g/m2/12 h i.v. days 1–3, mitoxantrone 10 mg/m2 i.v. days 2 and 3). SCT was permitted if there was a HLA-identical family donor on the decision of the local investigator Second randomization was performed after completion of first consolidation	3 cycles of high dose cytarabine – 3g/m^2 bid on days 1, 3, 5 (or 1,2 3) High risk patients received HCT. NB from Dec 2006 all FLT3 patients given HSCT as consolidation (number of patients not known). If MRD available HSCT in 1st CR to most patients

	therapy for patients in CR. Patients were randomized to either a second intensive consolidation therapy IEiv (idarubicin 12 mg/ m2 i.v. days 1 and 3, etoposide 100 mg/m2 i.v. days 1–5) or to a 1-year oral maintenance therapy IEpo (idarubicin 5mg p.o. days 1, 4, 7, 10, 13, etoposide 100 mg p.o. days 1 and 13; repeated on day 29 for 12 courses).	

2. Neither the characteristics of the controls nor the 223 Phase II study patient cohort are reported, though as the Phase II cohort included patients up to age 70 only it can be assumed that this analysis is also limited to patients of 70 years or younger. From the K-M plots provided, 62 (15%) were aged over 60 (up to 70 years) and 353 (85%) were 60 years or younger. This does not reflect clinical practice where >55% of patients are diagnosed aged >70 years.

3. This analysis is an observational study and therefore, is subject to bias. Whilst the propensity score matching may improve the comparability of the cohorts on known factors, it cannot account for unknown confounders. Furthermore, the method of propensity score matching is best applied to very large observational studies, rather than single arm studies of limited size.

4. The ERG notes that it is unclear if the analysis of OS was censored for SCT (the main analysis of RATIFY was not).

5. These results are much more favourable to midostaurin than are the results of the RATIFY trial. This appears to be due to the poorer survival results in the historical cohort than in the control arm of RATIFY. The ERG rough estimates (from manual reading off graphs provided) for OS and EFS are given in Table 3.

Table 3: Comparison of EFS and OS in the RATIFY trial and in the historical controls

	Control arm RATIFY	**Historical control**
EFS		
EFS at 12 mths	36%	30%
EFS at 24 mths	28%	22%
EFS at 48 mths	27%	20%
OS		
OS at 12 mths	64%	60%
OS at 24 mths	50%	35%
OS at 48 mths	44%	30%

The ERG notes that the OS control in the >60 years comparison is very much poorer than that in the ≤60 years one and also notes that the sample size in the older age group comparison is very much smaller.

In summary, these additional analyses do indicate that there is uncertainty around the midostaurin treatment effect in older patients. However, the inherent uncertainty in the analysis of observational data, means that there is still no reliable estimate of the treatment benefit of midostaurin in patients aged up to 70 years. Furthermore, there is still no estimate of the treatment benefit of midostaurin in the full population who would be eligible for this treatment in the NHS (i.e. including patients aged over 70 years).

Given the limitations of the propensity score analysis the ERG does not believe it provides a more reliable estimate of treatment effect than does the results of the RATIFY trial.

2.6 Evidence 6: The ICER given the new age adjustments recommended by the ERG

The addendum presents additional economic evidence exploring the impact of a number of alternative assumptions. This additional evidence consists of a number revisions to the company base-case model based on responses to questions raised by the ERG at the points for clarification stage and the incorporation of an age adjustment scenario which seeks to respond to the ERG's concerns regarding the representativeness of the population recruited to the RATIFY trial. The ERG considers the revisions based on the previously presented analysis, a new company preferred base-case, and

discusses these changes to the model, before proceeding to assess the validity of the new age adjusted scenario.

2.6.1 Revised company base-case

Table 4 provides a brief summary of the assumptions made in the company's revised base-case and those made in the ERG's base-case analysis.

Table 4: Comparison of Company revised base-case and ERG base-case

	Company revised base-case	**ERG base-case**
ERG calculation corrections	Not included	Included
Model structure/health state costs	50% reduction to routine care costs after 26 cycles.	Zero health state costs after treatment. Utilities all equal to CR1L health state following treatment
Complete response data	As per the original company base-case	As per the original company base-case
Time on treatment	As per the original company base-case	As per the original company base-case
Mortality: trial period	Updated OS data cut	Updated OS data cut
Mortality: post-trial period	General population	Four fold multiplier applied
Age adjusted utilities	Not included	Included
GVHD complications	Included	Included
SCT costs	Based on NHS blood and transplant 2014 cost	Based on NHS reference costs as per the company base-case.
Maximum number of cycles of monotherapy	Set to 12 as per the market authorisation	Set to 18 as per RATIFY trial.

As can be seen from Table 4, in its revised base-case the company rejects a number of changes made by the ERG in its base-case analysis. The company presents no additional evidence or argument relating to these issues, but presumably contests the appropriateness of these assumptions. For completeness the ERG presents a brief restatement of the ERG's position with regard these ERG base-case assumptions.

ERG calculation errors: The ERG identified a small number of inconstancies and calculation in the original company model. These corrections were not adopted by the company in its addendum submission. It is not clear whether the company is in disagreement with regards of the validity of these corrections, but the notes that the company did not raise any concerns with respect to the ERG's calculation corrections in its factual accuracy report. The ERG considers the company's decision to not include these calculation corrections to be highly problematic as it means there is not a mutually agreed model in which alternative assumptions can be explored.

Model structure health state costs: In the ERG's original report a number of significant issues with the company's base-case model structure were noted. Of primary concern was the fact that the model assumes significant ongoing routine care costs for patients in the relapse, CR 1L (patients in remission post discontinuation of treatment) and post-SCT recovery health states. The ERG made a number changes to the company's base-case model to ameliorate the impact of these assumptions by assuming zero routine care costs after discontinuation of treatment. However, the company chose not to adopt any of these changes and instead applied a 50% reduction in routine care costs after 26 cycles (months). As noted in our original report, the ERG consider this scenario inappropriate as it does not address the underlying issues identified with the company base-case and still implies significant ongoing health state costs for patients; ~£3000 per annum in the CR1L Post SCT recovery health states and ~£30,000 in the relapse health state. This scenario also implies that patients who relapse will continue to experience lower HRQoL even if they are successfully treated.

Mortality in the post-trial period: In the post-trial period the company base-case assumed that all surviving patients have general population mortality after the cure point (6.2 years). Existing epidemiological evidence, however, suggests that the mortality risk for patients surviving at least five years after SCT, without relapse, remains considerably higher than that for the general population (between 4 to 9 times higher, irrespective of age). The ERG therefore considered that the company's base-case was overly optimistic.

Age adjusted utilities: When utility values are considered over a 60-year lifetime horizon, it is evident that the utility values assigned to the CR 1L and post-SCT recovery states may eventually exceed general population utility estimates, which naturally decline with age. The ERG thus considers that utilities in the CR 1L and post-SCT recovery state should be further adjusted for declining HRQoL with age.

SCT costs: The company's original base-case included the costs of SCT from NHS reference costs. The ERG, however, noted alternative costs were available from NHS blood and transplant (2014). At the clarification stage, the company provided a scenario analysis implementing the values from the NHS Blood and Transplant (2014), which have now been adopted in its revised base-case. As part of their response the company, however, stated a preference for using NHS reference costs noting the following: "SCT process has evolved substantially since 2002 and is currently far more common. Therefore, inflating a 2002 cost to 2017 may overestimate the SCT cost. On the other hand, however, the NHS reference cost potentially excludes some of the SCT costs, though the more recent cost source is likely to be more accurate than the inflation of 2002 costs used in the NHS publication." The ERG considered the arguments provided by the company in its response persuasive and therefore retained the assumptions used in the original company base-model (NHS reference costs). It is not clear why the

company chose to adopt NHS blood and transplant (2014) as source of SCT costs in its revised base-case given their previously stated preference for using NHS reference costs.

2.6.2 Age related adjustment

To address the ERG's concerns regarding the representativeness of the modelled population the company presents a new age adjustment scenario. This new analysis makes two separate changes to the company's base-case analysis, these consist of changes to the OS data and changes to the average age of the population modelled.

Revisions to OS data: The age adjustment scenario incorporates new OS data into the model based on analysis of the propensity score matched historical controls described above (Evidence 5). To incorporate this new OS data the company splits the modelled population into two groups young (<60) and old (>60). Overall survival data for the young group are sourced from the RATIFY trial as per the original base-case. Overall survival data from the old cohort are sourced from the new comparison with historical controls (Evidence 5). To estimate cost-effectiveness for the combined cohort the OS data are weighted assuming 41% of patients are young and 59% are old. The weights are derived from age-specific incidence rates. Only the OS data used in the model is changed and all other clinical data are sourced from the RATIFY trial as per the original base-case.

Average age of cohort: In the post-trial period the overall survival of patients in the model is assumed to follow that of the general population with mortality rates determined by the mean age of the cohort. To account for the fact that this scenario assumes a greater number of older patients the model assumes that the mean age of the patients receiving midostaurin is 65 years.

The ERG has a number of substantive concerns regarding the age adjustment scenario. These issues concern the reliability of the data used to model OS; inconsistencies in the clinical data used; the proportion of older patients assumed; and the average age of the cohort assumed. These are discussed in turn below.

- **Reliability of OS data:** As described above the ERG has significant concerns relating to the reliability of the new propensity score matched historical control analysis noting that this comparison estimates that OS benefits are much larger than suggest by the RATIFY trial.
- **Inconstancies in clinical data:** While the age adjustment scenario incorporates new OS data into the model it does adjust the other clinical data used in the model including the response/relapse, rate of SCT or time on treatment. It is highly likely that are significant differences between younger and older patients with respect to these clinical parameters. For example, older patients are significantly less likely to receive a SCT than younger patients. These inconstancies are likely to have a significant impact on the cost-effectiveness estimates

and therefore caution should be used in interpreting the results of this analysis as they are subject to very significant uncertainty.

- **Proportion of older patients:** The proportion of older patients assumed by the company in the scenario is based on the incidence of AML in patients over the age of 60. The ERG considers that this approach is likely to overestimate the proportion of older patients. This is because not all older patients will be eligible for treatment with midostaurin treatment due to the requirement that patients be able to tolerate intensive chemotherapy. In the company model increasing the proportion of older patents acts to increase the ICER. This effect is, however, reversed in the ERG base-case model as a result of the ERG's changes to the model structure.
- **Mean age of cohort:** The mean age of the cohort is used to determine the mortality of patients post the cure point. In the company's age adjustment scenario the mean age of is assumed to be 65. The ERG considers this assumption to be inconsistent with the OS data used as only 59% of patients are assumed to be between the ages of 60 and 70, with the remaining 41% having a mean age of 45. If the mean age of the older age cohort is assumed to be 65 and the appropriate weighting applied, the mean age of the whole cohort is instead 56.8. The impact of reducing the mean age of the population modelled is to reduce the ICER

In summary while the ERG acknowledges the difficulties of modelling the effectiveness of midostaurin in more representative populations of young and older patients, the ERG considers that the additional age adjustment scenario is subject to a number of significant limitations. Particularly, the ERG considers it likely that the new OS data incorporated into the model is likely to significantly overestimate the benefits of midostaurin.

2.6.3 Additional ERG analysis

To allow the committee to understand the impact of the new age adjustment scenario the ERG present additional analysis in which this scenario is incorporated into the ERG's base-case. Results of this additional analysis are presented in Table 5. The results of this analysis show that incorporating the new age adjustment scenario into the ERG base-case substantially reduces the ICER (note this is the reverse of incorporating it into the company's revised base-case where the ICER increases). The reduction in the ICER observed in the ERG base-case is because the propensity score matched analysis estimates that midostaurin provides much greater OS benefits in the older cohort than in the younger cohort.

Table 5: Incorporating the new age adjustment scenario in ERG's base-case

Scenarios	Treatments	Costs	QALYs	Inc. cost	Inc. QALY	ICER	Change in ICER
CS revised base case without age related adjustment	Midostaurin therapy	[redacted]	[redacted]	[redacted]	[redacted]	£13,588	n/a
	SOC	[redacted]	[redacted]	-	-	-	-
CS revised base case (with age related adjustment: mean age 57)	Midostaurin therapy	[redacted]	[redacted]	[redacted]	[redacted]	24,001	+£10,413
	SOC	[redacted]	[redacted]	-	-	-	-
CS revised base case (with age related adjustment: mean age 65)	Midostaurin therapy	[redacted]	[redacted]	[redacted]	[redacted]	£27,754	+£14166
	SOC	[redacted]	[redacted]	-	-	-	-
ERG's preferred base case (without age related adjustment)	Midostaurin therapy	[redacted]	[redacted]	[redacted]	[redacted]	£62,810	+£49,222
	SOC	[redacted]	[redacted]	-	-	-	-
ERG's preferred base case (with new adjustment: mean age 57)	Midostaurin therapy	[redacted]	[redacted]	[redacted]	[redacted]	£35,999	+£22,411
	SOC	[redacted]	[redacted]	-	-	-	-
ERG's preferred base case (with new adjustment: mean age 65)	Midostaurin therapy	[redacted]	[redacted]	[redacted]	[redacted]	£45,060	+£31,472
	SOC	[redacted]	[redacted]	-	-	-	-

3 References

Linch, D., Hills, R., & Brunett, A. (2014). Impact of FLT3(ITD) mutant allele level on relapse risk in intermediate-risk acute myeloid leukemia. *Blood, 124*, 273-276.

National Health Service Blood and Transplant. *Unrelated donor stem cell transplantation in the UK. A report from the UK stem cell strategy oversight committee*; 2014. Available from: nhsbt.nhs.uk

Schlenk, R., Frohling, S., & Hartmann, F. (2004). Phase III study of all-trans retinoic acid in previously untreated patients 61 years or older with acute myeloid leukemia. *Leukemia, 18*, 1798-1803.

Schlenk, R., Frohling, S., & Hartmann, F. (2006). Intensive consolidation versus oral maintenance therapy in patients 61 years of older with acute myeloid leukemia in first remission: results of second randomization of the AML HD98-B treatment trial. *Leukemia, 20*, 748-750.

Schlenk, R., Fiedler, W., & Salih, H. (2016). Impact of age and midostaurin-dose on response and outcome in acute myeloid leukemia with FLT3-ITD: interim-analyses of the AMLSG 16-10 trial. *Blood, 128*, 449.

Whitman, S., Maharry, K., & Radmacher, M. (2010). FLT3 internal tandem duplication associates with adverse outcome and gene- and microRNA-expression signatures in patients 60 years of age or older with primary cytogenetically normal acute myeloid leukemia: a Cancer and Leukemia Group B study. *Blood, 116*, 3622-3626.